UNIVERSITÉ CLERMONT AUVERGNE
ÉCOLE DOCTORALE SPI
SCIENCES POUR L'INGENIEUR

T H È S E

pour obtenir le titre de

Docteur en Sciences

Présentée et soutenue par

Kilian CHANDELON

Contributions au recalage et au suivi visuel monoculaire et stéréoscopique en conditions synchrones et asynchrones pour la chirurgie minimalement invasive

Applications en chirurgie du rein, de l'utérus et du foie, perspectives de généralisation aux organes raisonnablement rigides

Thèse dirigée par Adrien BARTOLI

soutenue le 19 Décembre 2025

Jury :

Président :	**Nicolas PADOY**, Professeur des Universités, Université de Strasbourg, France
Rapporteurs :	**Daniel PIZARRO**, Professeur des Universités, Université d'Alcalá, Espagne
	Sergi PUJADES, Maître de Conférences, INRIA Université Grenoble Alpes, France
Examinatrice :	**Nathalie PHAM DANG**, Professeur des Universités, UCA – CHU Clermont-Ferrand, France
Directeur de thèse :	**Adrien BARTOLI**, Professeur des Universités, UCA – IP – CHU Clermont-Ferrand, SURGAR, France
Co-encadrant :	**Nicolas BOURDEL**, Professeur des Universités, UCA – IP – CHU Clermont-Ferrand, SURGAR, France
Invité :	**Prasad SAMARAKOON**, Docteur en Sciences, SURGAR, France

Résumé

La Chirurgie Mini-Invasive (CMI) et plus particulièrement la cœlioscopie, avec ou sans assistance robotique, a profondément transformé la pratique chirurgicale en réduisant les traumatismes peropératoires, la douleur, les complications postopératoires ainsi que la durée d'hospitalisation, comparativement à la chirurgie ouverte. Parmi ses applications en chirurgie urologique figure la Néphrectomie Partielle (NP) et sa variante robot-assistée la néphrectomie partielle robot-assistée, aujourd'hui majoritaire en France. Cette intervention constitue la référence pour le traitement des tumeurs rénales localisées. L'un des principaux défis de la NP réside dans la localisation peropératoire précise des structures anatomiques, notamment la tumeur, les branches vasculaires et le système collecteur. L'objectif est de réaliser l'exérèse tumorale tout en préservant au maximum le parenchyme rénal sain et par conséquent la fonction rénale post-opératoire.

L'imagerie préopératoire occupe une place centrale dans ce contexte : les acquisitions par tomodensitométrie ou imagerie par résonance magnétique permettent de générer un jumeau numérique tridimensionnel du rein et de ses structures internes. Ce modèle peut être visualisé en réalité virtuelle au bloc opératoire ou recalé sur la vue peropératoire en Réalité Augmentée (RA). Ce recalage 3D-2D repose sur des méthodes de vision par ordinateur et constitue une étape clé du guidage chirurgical augmenté. Les travaux antérieurs de la communauté scientifique et de l'équipe académique EnCoV (Endoscopy and Computer Vision) ont posé les fondations d'un processus computationnel de RA pour la chirurgie utérine. Toutefois, son adaptation à la NP soulevait des défis spécifiques, liés à la nature rétro-péritonéale du rein, à la nécessité de phases de dissection avant exposition et aux variations techniques de la chirurgie robotique offrant des vues très localisées et une forte dynamique dans les mouvements relatifs caméra-organe cible.

Cette thèse propose et évalue un ensemble de contributions méthodologiques et technologiques visant à fiabiliser et automatiser ce processus de RA en conditions opératoires.

La première contribution concerne la synchronisation automatique des flux vidéo stéréoscopiques, élément essentiel pour garantir la cohérence temporelle et géométrique de la perception tridimensionnelle. Une méthode de comparaison des vues gauche et droite, fondée sur l'utilisation conjointe des composantes spéculaires et diffuses, permet une synchronisation en temps réel, robuste aux perturbations optiques et adaptée aux environnements chirurgicaux à dynamique variable.

La seconde contribution porte sur le recalage 3D-2D automatique du jumeau numérique préopératoire sur les images peropératoires. Un modèle générique du rein, combiné à un réseau de neurones convolutionnel entraîné de manière supervisée sur données cliniques, permet d'estimer directement la pose tridimensionnelle de l'organe sans recours à des primitives de recalage, inexistantes en NP. La faisabilité technique de cette approche a été validée sur sept cas cliniques, avec une précision comparable au recalage manuel, tout en maintenant des temps d'exécution compatibles avec la pratique chirurgicale.

Enfin, la troisième contribution introduit un suivi visuel d'organe en temps réel, fondé sur une base d'images clés dynamique capable de s'adapter aux variations morphologiques et d'apparence au cours de l'intervention. Ce mécanisme améliore la stabilité de l'affichage augmenté et assure la continuité du guidage visuel lors des phases critiques de la NP.

L'ensemble de ces développements s'inscrit dans la continuité des travaux sur la RA chirurgicale appliquée à la CMI et démontre la faisabilité d'un processus computationnel complet, de l'acquisition vidéo à l'affichage augmenté, validé en contexte opératoire réel. Ce travail ouvre la voie à un guidage peropératoire automatisé et potentiellement généralisable à d'autres organes dits « raisonnablement rigides ».

Mots clés : chirurgie mini-invasive, chirurgie guidée par l'image, réalité augmentée, néphrectomie partielle, chirurgie robotique, recalage, suivi automatique, synchronisation

Abstract

Minimally Invasive Surgery (MIS) and more specifically laparoscopy, with or without robotic assistance, has profoundly transformed surgical practice by reducing intraoperative trauma, postoperative pain, and complications, as well as the length of hospital stay, when compared with open surgery. Among its applications in urological surgery is Partial Nephrectomy (PN) and its robot-assisted variant, robot-assisted partial nephrectomy, which currently represents the majority of procedures performed in France. This operation is the gold standard for the treatment of localised renal tumours. One of the main challenges of PN lies in the accurate intraoperative localisation of anatomical structures, particularly the tumour, vascular branches, and the collecting system. The objective is to achieve complete tumour excision while preserving as much healthy renal parenchyma as possible, thereby maintaining postoperative renal function.

Preoperative imaging plays a central role in this context : computed tomography or magnetic resonance imaging acquisitions enable the generation of a three-dimensional digital twin of the kidney and its internal structures. This model can be visualised in virtual reality in the operating room, or registered onto the intraoperative view in Augmented Reality (AR). This 3D-2D registration relies on computer vision methods and constitutes a key step in augmented surgical guidance. Previous work by the scientific community and the EnCoV (Endoscopy and Computer Vision) academic team laid the groundwork for a computational AR pipeline in uterine surgery. However, adapting it to PN raised specific challenges related to the retroperitoneal nature of the kidney, the requirement for dissection phases prior to exposure, and the technical variability of robotic surgery, which offers highly localised views and strong dynamics in the relative motion between the camera and the target organ.

This thesis proposes and evaluates a set of methodological and technological contributions aimed at improving the reliability and automation of this AR pipeline under real operative conditions.

The first contribution concerns the automatic synchronisation of stereoscopic video streams, a key element for ensuring temporal and geometric coherence in three-dimensional perception. A method comparing left and right views, based on the combined use of specular and diffuse components, enables real-time synchronisation that is robust to optical disturbances and suited to the variable dynamics of surgical environments.

The second contribution addresses the automatic 3D-2D registration of the preoperative digital twin onto intraoperative images. A generic kidney model, combined with a convolutional neural network trained in a supervised manner on clinical data, allows the direct estimation of the organ's three-dimensional pose without relying on registration primitives, which are absent in PN. The technical feasibility of this approach was validated on seven clinical cases, achieving accuracy comparable to manual registration while maintaining execution times compatible with surgical workflow.

Finally, the third contribution introduces a real-time visual tracking mechanism for the organ, based on a dynamic keyframe database capable of adapting to morphological and appearance changes during the procedure. This mechanism enhances the stability of the augmented display and ensures continuous visual guidance during the critical phases of PN.

Together, these developments build upon previous research on surgical AR applied to MIS and demonstrate the feasibility of a complete pipeline from video acquisition to augmented display, validated in real operative settings. This work paves the way for automated intraoperative guidance potentially generalisable to other so-called "reasonably rigid" organs.

Keywords : minimally invasive surgery, image-guided surgery, augmented reality, partial nephrectomy, robotic surgery, registration, automatic tracking, synchronisation

Remerciements

Je souhaite tout d'abord exprimer ma profonde gratitude à Adrien Bartoli, mon directeur de thèse, pour m'avoir offert l'opportunité d'explorer le domaine fascinant de la vision par ordinateur, qui m'était initialement inconnu. Merci pour votre encadrement scientifique exigeant, votre disponibilité, votre rigueur et la confiance que vous m'avez accordée tout au long de ces années.

J'adresse également mes sincères remerciements à Nicolas Bourdel, CEO de SURGAR, pour son implication précieuse sur le plan médical en tant que chirurgien, ainsi qu'à Prasad Samarakoon, CTO de SURGAR, pour son accompagnement dans le cadre de ce projet.

Je remercie chaleureusement l'ensemble des équipes SURGAR et EnCoV pour l'environnement de travail stimulant, les discussions scientifiques enrichissantes et l'esprit d'équipe qui ont rendu ces années particulièrement formatrices.

Ma reconnaissance s'étend également aux équipes cliniques et chirurgicales, qui m'ont permis de mieux appréhender les enjeux concrets de l'application médicale de mes travaux. Je remercie tout particulièrement le Pr. Jean-Christophe Bernhard, le Pr. Nathalie Pham Dang, le Pr. Laurent Devoize, le Dr. Laurent Savareux, le Dr. Abderrahmane Khaddad, le Dr. Gaëlle Margue, le Dr. Alice Pitout, ainsi que l'ensemble des professionnels ayant contribué à cette collaboration interdisciplinaire.

Je tiens à exprimer une gratitude particulière envers Julie D. et Mathieu S., collaborateurs étroits sur le projet, pour leur investissement, leurs échanges constructifs et leur constante bonne humeur. Mes remerciements s'adressent également à Mathieu L., Ivan M., Justine M., Laura O., Benjamin R., Romain M., Richard M., Rasoul S., Mina Z., Nicolas M. et Nicolas G., pour leur soutien et les nombreux moments de partage, tant professionnels qu'amicaux. Ces belles rencontres humaines ont grandement contribué à mon équilibre et à la richesse de ces années de doctorat.

Je souhaite également remercier mes collègues de promotion d'école d'ingénieur, devenus depuis de véritables amis, les Pingouins, qui se reconnaîtront pour leur amitié fidèle, leur humour et leur soutien tout au long de ces années. Leur présence, même à distance, a souvent apporté un souffle d'énergie et de légèreté bienvenu.

Je souhaite exprimer toute ma reconnaissance à ma famille, pour son soutien indéfectible. À mes parents, Magali et Lionel, pour m'avoir offert les conditions idéales d'étude et transmis des valeurs essentielles de persévérance, de curiosité et de bienveillance. À mon frère Melvyn, à ma tante Myriam, à mon oncle Christophe, à mes cousins Maëlly et Loan, ainsi qu'à mon grand-père Denis, pour leur affection et leurs encouragements constants.

Une pensée affectueuse va à celles et ceux qui ne sont plus parmi nous : Arlette, ma grand-mère, ainsi que Ginette et Arsène, mes arrière-grands-parents. Leur souvenir m'accompagne et continue d'inspirer mon parcours.

Enfin, je ressors de cette expérience avec une conviction forte : mettre mes compétences au service de l'amélioration de la santé de demain, en plaçant la technologie au service de l'humain.

Valorisations scientifiques

Articles publiés dans des revues internationales

- **Landmark-free Automatic Digital Twin Registration in Robot-assisted Partial Nephrectomy using a Generic End-to-end Model**
  K. Chandelon, A. Pitout, M. Souchaud, J. Desternes, G. Margue, J. Peyras, N. Bourdel, J.-C. Bernhard and A. Bartoli
  International Journal of Computer Assisted Radiology and Surgery, special issue : selected papers from CARS 2025, Juin 2025

- **Robotic-Assisted Partial Nephrectomy in Augmented Reality : Reliability of a Neural Network Training Methodology for Automated Renal Parenchyma Recognition (UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, M. Jaffredo, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  The Journal of Urology, Mai 2025

- **Development of an Augmented Reality Technique for Robot-Assisted Partial Nephrectomy (3D-AR-IGRAPN-UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  The Journal of Urology, Mai 2025

- **Synchronising a Stereoscopic Surgical Video Stream using Specular Reflection**
  K. Chandelon and A. Bartoli
  International Journal of Computer Assisted Radiology and Surgery, special issue : selected papers from IPCAI 2024, Juin 2024

- **A Deep-learning Approach to Renal Parenchyma Detection : A First Step Towards Augmented Reality to Guide Robot-assisted Partial Nephrectomy**
  A. Khaddad, A. Bartoli, K. Chandelon, G. Margue, J. Desternes, N. Bourdel and J.-C. Bernhard
  The Journal of Urology, 211(5S) :925-926, Mai 2024

- **Kidney Tracking for Live Augmented Reality in Stereoscopic Mini-invasive Partial Nephrectomy**
  K. Chandelon, R. Sharifian, S. Marchand, A. Khaddad, N. Bourdel, N. Mottet, J.-C. Bernhard and A. Bartoli
  Computer Methods in Biomechanics and Biomedical Engineering : Imaging and Visualization, special issue : selected papers from AE-CAI 2022, 11(4) :1251-1260, Juin 2023

- **Elliptical Specularity Detection in Endoscopy with Application to Normal Reconstruction**
  K. Makki, K. Chandelon and A. Bartoli
  International Journal of Computer Assisted Radiology and Surgery, special issue : selected papers from IPCAI 2023, 18(5) :1323-1328, Mai 2023

- **A Survey of Augmented Reality Methods to guide Minimally Invasive Partial Nephrectomy**
  A. Khaddad, J.-C. Bernhard, G. Margue, C. Michiels, S. Ricard, K. Chandelon, F. Bladou, N. Bourdel and A. Bartoli
  World Journal of Urology, 41(2) :335-343, Février 2023

- **Tracking Better, Tracking Longer : Automatic Keyframe Selection in Model-based Laparoscopic Augmented Reality**
  K. Chandelon and A. Bartoli
  International Journal of Computer Assisted Radiology and Surgery, special issue : selected papers from IPCAI 2022, 17(8) :1507-1511, Août 2022

- **A Proof-of-Concept Augmented Reality System in Oral and Maxillofacial Surgery**
  N. Pham-Dang, K. Chandelon, I. Barthélémy, L. Devoize and A. Bartoli
  Journal of Stomatology, Oral and Maxillofacial Surgery, 122(4) :338-342, Septembre 2021

Articles publiés dans des revues nationales

- **Faisabilité de la réalité augmentée 3D dans la néphrectomie partielle robot-assistée guidée par l'image : expérience clinique prospective et résultats préliminaires de 3D-AR-IGRAPN (UroCCR 112)**
  A Pitout, G Margue, K Chandelon, J Desternes, A Khaddad, M Souchaud, F Rubat Baleuri, M Pattou, M Jaffredo, L Richert, A Bartoli, N Bourdel, C Fleury and JC Bernhard
  Progrès en Urologie - FMC, Volume 35, Issue 7, Supplement, Novembre 2025

- **Néphrectomie partielle robot-assistée en réalité augmentée : fiabilité d'une méthodologie d'entraînement d'un réseau de neurones pour reconnaissance automatisée du parenchyme rénal (UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Progrès en Urologie - FMC, Volume 34, Issue 7, Supplement, Novembre 2024

- **Mise au point d'une technique de réalité augmentée pour la néphrectomie partielle robot-assistée**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Progrès en Urologie - FMC, Volume 34, Issue 7, Supplement, Novembre 2024

- **Néphrectomie totale avec thrombectomie cave guidée par réalité augmentée : technique et partage d'expérience**
  G. Margue, A. Pitout, J. Desternes, K. Chandelon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Progrès en Urologie - FMC, Volume 34, Issue 7, Supplement, Novembre 2024

- **Reconnaissance automatisée de la surface rénale par intelligence artificielle au cours de néphrectomies partielles robot-assistées (NPRA) : un pas vers la réalité augmentée**
  A. Khaddad, A. Bartoli, K. Chandelon, G. Margue, J. Desternes, N. Bourdel and J.-C. Bernhard
  Progrès en Urologie - FMC, special issue : Congrès Français d'Urologie, 33(3) :S46-S47, Novembre 2023

Communications orales

- **Landmark-free Automatic Digital Twin Registration in Robot-assisted Partial Nephrectomy using a Generic End-to-end Model**
  K. Chandelon, A. Pitout, M. Souchaud, J. Desternes, G. Margue, J. Peyras, N. Bourdel, J.-C. Bernhard and A. Bartoli
  CARS - Computer Assisted Radiology and Surgery, 2025

- **Development of an Augmented Reality technique for robot-assisted partial nephrectomy (3D-AR-IGRAPN - UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  AUA2025 - American Urological Association annual meeting, 2025

- **Vena cava thrombectomy and robot assisted radical nephrectomy using augmented reality guidance : Technique and experience sharing**
  G. Margue, A. Pitout, J. Desternes, K. Chandelon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  EAU25 - European Association of Urology congress, 2025

- **Development of an augmented reality technique for robot-assisted partial nephrectomy (3D-AR-IGRAPN - UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  EAU25 - European Association of Urology congress, 2025

- **Synchronising a Stereoscopic Surgical Video Stream using Specular Reflection**
  K. Chandelon and A. Bartoli
  IPCAI - International Conference on Information Processing in Computer-Assisted Interventions, 2024

- **Néphrectomie partielle robot-assistée en réalité augmentée : fiabilité d'une méthodologie d'entraînement d'un réseau de neurones pour reconnaissance automatisée du parenchyme rénal (UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Congrès Français d'Urologie, 2024

- **Mise au point d'une technique de réalité augmentée pour la néphrectomie partielle robot-assistée**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Congrès Français d'Urologie, 2024

- **Néphrectomie totale avec thrombectomie cave guidée par réalité augmentée : technique et partage d'expérience**
  G. Margue, A. Pitout, J. Desternes, K. Chandelon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  Congrès Français d'Urologie, 2024

- **Deep-learning approach to renal parenchyma detection : a first step towards augmented reality to guide robot-assisted partial nephrectomy**
  A. Khaddad, A. Bartoli, K. Chandelon, G. Margue, J. Desternes, N. Bourdel and J.-C. Bernhard
  EAU24 - European Association of Urology congress, 2024

- **Reconnaissance automatisée de la surface rénale par intelligence artificielle au cours de néphrectomies partielles robot-assistées (NPRA) : un pas vers la réalité augmentée**
  A. Khaddad, A. Bartoli, K. Chandelon, G. Margue, J. Desternes, N. Bourdel and J.-C. Bernhard
  Congrès Français d'Urologie, 2023

- **Elliptical Specularity Detection in Endoscopy with Application to Normal Reconstruction**
  K. Makki, K. Chandelon and A. Bartoli
  IPCAI - International Conference on Information Processing in Computer-Assisted Interventions, 2023

- **Kidney Tracking for Live Augmented Reality in Stereoscopic Mini-invasive Partial Nephrectomy**
  K. Chandelon, R. Sharifian, S. Marchand, A. Khaddad, N. Bourdel, N. Mottet, J.-C. Bernhard and A. Bartoli
  Joint AE-CAI | CARE | OR 2.0 Workshop at MICCAI, 2022

- **Tracking Better, Tracking Longer : Automatic Keyframe Selection in Model-based Laparoscopic Augmented Reality**
  K. Chandelon and A. Bartoli
  IPCAI - International Conference on Information Processing in Computer-Assisted Interventions, 2022

- **Preuve de concept en réalité augmentée pour la chirurgie maxillo-faciale**
  N. Pham-Dang, K. Chandelon, I. Barthélémy, L. Devoize and A. Bartoli
  Congrès de la Société Française de Stomatologie, Chirurgie Maxillo-Faciale et Chirurgie Orale, 2021

Communications affichées

- **Robotic-assisted partial nephrectomy in augmented reality : reliability of a neural network training methodology for automated renal parenchyma recognition (UroCCR 112)**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, M. Jaffredo, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  AUA2025 - American Urological Association annual meeting, 2025

- **3-D imaging and beyond - what does it add ?**
  A. Pitout, G. Margue, K. Chandelon, J. Desternes, A. Khaddad, M. Jaffredo, J. Peyras, M. Souchaud, E. Jambon, N. Bourdel, A. Bartoli and J.-C. Bernhard
  EAU25 - European Association of Urology congress, 2025

- **Synchronising a Stereoscopic Surgical Video Stream using Specular Reflection**
  K. Chandelon and A. Bartoli
  IPCAI - International Conference on Information Processing in Computer-Assisted Interventions, 2024

- **Kidney Tracking for Live Augmented Reality in Stereoscopic Mini-invasive Partial Nephrectomy**
  K. Chandelon, R. Sharifian, S. Marchand, A. Khaddad, N. Bourdel, N. Mottet, J.-C. Bernhard and A. Bartoli
  Joint AE-CAI | CARE | OR 2.0 Workshop at MICCAI, 2022

- **Development of a real-time augmented reality system for laparoscopic kidney surgery**
  K. Chandelon, A. Bartoli, P. Samarakoon and N. Bourdel
  JSD ED SPI 2021 - Journées scientifiques des doctorants, 2021

Chirurgie en direct

- **Radical nephrectomy with thrombus level 1-2 with augmented reality**
  J.-C. Bernhard and G. Margue (Surgical team) | K. Chandelon, J. Desternes, M. Souchaud, A. Bartoli and N. Bourdel (Tech team)
  ERUS24 - European Robotic Urology Section, 2024

Brevets

- **Titre de l'invention : procédé de détermination d'une métrique de divergence entre deux images et procédé de synchronisation, dispositif et système associés**
  Inventeurs : Kilian Chandelon et Adrien Bartoli
  Déposants : SURGAR, UNIVERSITÉ CLERMONT-AUVERGNE, CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, CLERMONT AUVERGNE INP
  Pays : France & Polynésie Française + PCT
  Numéro de la demande FR : FR2311679 déposée le 26 octobre 2023
  Numéro de la demande PCT : PCT/EP2024/080071 déposée le 24 octobre 2024
  Statut : rapport de recherche préliminaire (RRP) reçu avec une opinion écrite favorable

- **Titre de l'invention : procédé et dispositif de suivi en temps réel de la pose d'une caméra par gestion automatique d'une base d'images-clés**
  Inventeurs : Kilian Chandelon et Adrien Bartoli
  Déposants : SURGAR, UNIVERSITÉ CLERMONT-AUVERGNE, CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, CLERMONT AUVERGNE INP
  Pays : France & Polynésie Française, Europe, Chine, Canada, États-Unis, Inde et Japon

Numéro de la demande FR : FR2113591 déposée le 15 décembre 2021 et délivrée le 29 décembre 2023
Numéro de la demande PCT : PCT/EP2022/085905 déposée le 14 décembre 2022
Numéro de la demande CN : CN118451463 publiée le 6 août 2024
Numéro de la demande CA : 3,240,591 déposée le 10 juin 2024
Numéro de la demande US : 18/721,107 déposée le 17 juin 2024
Numéro de la demande IN : 202447053186 déposée le 11 juillet 2024
Numéro de la demande JP : 2024-532229 déposée le 28 mai 2024
Statut : rapport de recherche internationale (RRI) reçu avec une opinion écrite favorable

Logiciels

- **K-SURGAR prototype** (logiciel de réalité augmentée pour la néphrectomie partielle robot assistée), 2021 - 2025
- **RPU, Registration Process Understanding** (logiciel d'aide à la compréhension du recalage mental effectué par le chirurgien en peropératoire), 2023
- **U-SURGAR prototype** (logiciel de réalité augmentée pour la myomectomie), extension du module de suivi par ajout automatique d'images clés, 2021

Ensembles de données

- Recalage et suivi rigide du rein en Néphrectomies Partielles Robot-Assistées (NPRA), 7 chirurgies soit 299 659 images annotées, 2024 - 2025
- Synchronisation de flux vidéo en chirurgie robotique, 44 séquences vidéo stéréoscopiques synchrones acquises lors de Néphrectomies Partielles Robot-Assistées (NPRA) et de Résections Hépatiques Robot-Assistées (RHRA), 2023
- Segmentation des structures anatomiques en Néphrectomies Partielles Robot-Assistées (NPRA), 52 chirurgies soit 48 140 images annotées, 2020 - 2025

Vulgarisation scientifique

- **Intelligence artificielle et réalité augmentée, vers une nouvelle ère en chirurgie orale**
  K. Chandelon
  Congrès de la Société Française de Chirurgie Orale, Juin 2024

Encadrement et co-encadrement

- S. Marchand, thèse de médecine, CHU St Etienne, service d'urologie, 2020 - 2022 (avec A. Bartoli)
- A. Khaddad, M2 Recherche, CHU Bordeaux, service d'urologie, 2021 - 2022 (avec A. Bartoli)
- K. Zehnder, M1, ISIMA, 2022 (avec N. Pham Dang)
- Q. Chevillard, M2, ViBot, 2022 (avec A. Bartoli)
- G. Margue, M2 Recherche, CHU Bordeaux, service d'urologie, 2022 - 2023 (avec A. Bartoli et J. Desternes)
- A. Pitout, M2 Recherche, CHU Bordeaux, service d'urologie, 2023 - 2024 (avec A. Bartoli et J. Desternes)
- T. Gregoire, M1 médecine, 2023 (avec N. Pham Dang et L. Devoize)
- L. Jourjon, M1 médecine, 2023 (avec N. Pham Dang et L. Devoize)
- A. Egly, M1 médecine, 2023 - 2024 (avec N. Pham Dang et L. Devoize)
- E. Pallois, M1 médecine, 2023 - 2024 (avec N. Pham Dang et L. Devoize)
- L. Feral, P1 médecine, 2024 (avec N. Pham Dang et L. Devoize)
- M. Feral, P1 médecine, 2024 (avec N. Pham Dang et L. Devoize)
- A. Samadi, M1, CentraleSupélec, 2024 - 2025
- M. Pattou, M2 Recherche, CHU Bordeaux, service d'urologie, 2024 - 2025 (avec A. Bartoli et J. Desternes)
- F. Rubat, M2 Recherche, CHU Bordeaux, service d'urologie, 2024 - 2025 (avec A. Bartoli et J. Desternes)

Expertise / Lecture critique

- CBM – Computers in Biology and Medicine
- Scientific reports – Nature
- BMC Musculoskeletal Disorders – Nature
- JORS – Journal of Robotic Surgery
- IJCARS – International Journal of Computer Assisted Radiology and Surgery
- JORMAS – Journal of Stomatology oral and Maxillofacial Surgery
- AE-CAI, MICCAI Workshop – Augmented Environments for Computer-Assisted Interventions

- IPCAI – International Conference on Information Processing in Computer-Assisted Interventions
- DCAMI, CVPR Workshop – Data Curation and Augmentation in Enhancing Medical Imaging Applications

Table des matières

Liste d'abréviations **xix**

1 Introduction **1**
1.1 Contexte des travaux . . . 3
1.2 Contexte clinique . . . 3
1.2.1 Généralités . . . 3
1.2.2 Contexte chirurgical . . . 8
1.2.3 Chirurgie ouverte . . . 17
1.2.4 Chirurgie mini-invasive . . . 18
1.3 Contexte technique . . . 40
1.3.1 Outils de planification préopératoire . . . 40
1.3.2 Outils de guidage peropératoire . . . 46
1.3.3 Réalité augmentée en chirurgie . . . 50
1.4 Contributions et structure du manuscrit . . . 53

2 Contexte : Néphrectomie partielle et méthodes de guidage par réalité augmentée **55**
2.1 Contexte clinique de la chirurgie du rein . . . 57
2.1.1 Anatomie et physiologie du rein . . . 57
2.1.2 Pathologies et cancer du rein . . . 67
2.1.3 Traitements du carcinome à cellules rénales . . . 73
2.1.4 Néphrectomie partielle . . . 79
2.2 Guidage par réalité augmentée en chirurgie du rein . . . 85
2.2.1 Généralités . . . 85
2.2.2 Approche d'ingénierie logicielle . . . 90
2.2.3 Approche utilisateur . . . 92
2.2.4 État de l'art . . . 94
2.2.5 Acquisition vidéo endoscopique . . . 101
2.2.6 Étalonnage géométrique d'une caméra endoscopique . . . 104

3 Synchronisation de flux vidéo gauche-droit en chirurgie endoscopique **115**
3.1 Introduction . . . 117
3.2 Contexte technique général . . . 118
3.3 Interactions lumière-matière en cœlioscopie . . . 121
3.4 État de l'art . . . 127
3.4.1 Méthodes de synchronisation vidéo actives . . . 127
3.4.2 Méthodes de synchronisation vidéo passives . . . 128
3.4.3 Perception visuelle tridimensionnelle . . . 130
3.4.4 Détection automatique des taches spéculaires . . . 131

3.5 Mécanisme général de synchronisation . . . 133
3.6 Scores de comparaison Gauche-Droit . . . 136
3.6.1 Exploitation de la composante spéculaire . . . 136
3.6.2 Exploitation des composantes diffuse et spéculaire . . . 139
3.7 Génération de données d'évaluation semi-synthétiques . . . 139
3.8 Expériences . . . 142
3.8.1 Jeu de données pour l'évaluation . . . 142
3.8.2 Mesure de la distance temporelle . . . 143
3.8.3 Évaluation de la performance des scores proposés . . . 145
3.8.4 Évaluation de l'impact de la segmentation des taches spéculaires sur la synchronisation . . . 151
3.8.5 Évaluation de l'impact du score sur la synchronisation . . . 154
3.8.6 Applications en chirurgie assistée par ordinateur . . . 155
3.9 Conclusion . . . 160

4 Estimation automatique d'une pose globale par apprentissage profond 163
4.1 Introduction . . . 165
4.2 Contexte technique général . . . 167
4.3 État de l'art . . . 170
4.3.1 Méthodes de recalage 3D-2D en chirurgie abdomino-pelvienne 170
4.3.2 Recalage 3D-2D en néphrectomie partielle . . . 176
4.3.3 Récupération de la pose humaine à partir d'images monoculaires 178
4.4 Analyse de la méthodologie de recalage . . . 179
4.5 Approches de recalage 3D-2D basées sur des primitives . . . 184
4.5.1 Méthodes de recalage 3D-2D manuelles . . . 184
4.5.2 Méthodes de recalage 3D-2D semi-automatiques . . . 189
4.5.3 Méthodes de recalage 3D-2D automatiques . . . 190
4.6 Approche de recalage 3D-2D basée sur l'apprentissage automatique . 193
4.6.1 Généralités . . . 193
4.6.2 Collecte des données . . . 194
4.6.3 Labellisation des données . . . 195
4.6.4 Préparation des données . . . 197
4.6.5 Entraînement du réseau de neurones . . . 201
4.6.6 Inférence . . . 203
4.7 Expériences . . . 203
4.7.1 Évaluation de la performance brute . . . 203
4.7.2 Évaluation dans le contexte clinique . . . 204
4.7.3 Application . . . 206
4.8 Conclusion . . . 208

5 Suivi visuel d'organe en temps réel par points clés utilisant une base d'images clés dynamique 211
5.1 Introduction . . . 213
5.2 Contexte technique général . . . 214
5.3 État de l'art . . . 216
5.3.1 Mise en correspondance d'images . . . 216
5.3.2 Méthodes de suivi visuel en chirurgie assistée par ordinateur . 219
5.3.3 Méthodes de réduction de la dérive en suivi visuel . . . 228
5.3.4 Correction des distorsions de la perspective . . . 230
5.4 Segmentation automatique de l'organe dans les images peropératoires 231
5.4.1 Généralités . . . 231
5.4.2 Collecte des données . . . 232
5.4.3 Labellisation des données . . . 235
5.4.4 Entraînement du réseau de neurones . . . 240
5.4.5 Inférence . . . 242
5.5 Suivi monoculaire via une base d'images clés statique . . . 243
5.6 Suivi monoculaire via une base d'images clés dynamique . . . 247
5.7 Méthode étendue à la stéréovision . . . 250
5.8 Méthode avec correction automatique de la perspective . . . 251
5.9 Méthode spécialisée à la néphrectomie partielle robot assistée . . . 254
5.10 Méthode de suivi basée sur l'apprentissage profond . . . 257
5.11 Expériences . . . 259
5.11.1 Segmentation automatique du parenchyme rénal . . . 259
5.11.2 Apport de la méthode dynamique . . . 266
5.11.3 Apport de la stéréovision . . . 269
5.11.4 Apport de la correction de la perspective . . . 270
5.11.5 Évaluation en conditions peropératoires . . . 273
5.12 Conclusion . . . 276

6 Conclusion et travaux futurs 279
6.1 Conclusions . . . 281
6.1.1 Objectifs et trajectoire des travaux . . . 281
6.1.2 Synchronisation de flux vidéo stéréoscopiques en chirurgie endoscopique . . . 281
6.1.3 Estimation automatique d'une pose globale par apprentissage profond . . . 282
6.1.4 Suivi visuel d'organe en temps réel par points clés . . . 283
6.2 Travaux futurs . . . 283
6.2.1 Perspectives sur la synchronisation de flux stéréoscopiques . . 283
6.2.2 Perspectives sur le recalage 3D-2D automatique . . . 284
6.2.3 Perspectives sur le suivi visuel d'organe . . . 285

Annexes 287

A Liste des Diplômes d'Études Spécialisées (DES) définis par l'arrêté du 21 avril 2017 289

B Classification d'Altemeier 291

C Liste des solutions de chirurgie robotique sur le marché et en cours de développement en 2024 293

D Classification de Bosniak 2019 des tumeurs kystiques du rein 305

E Score de néphrométrie RENAL 307

F Score de néphrométrie PADUA 309

G Classification Tumour Node Metastasis (TNM) 2017 du cancer du rein 311

H Répartition des néphrectomies réalisées en 2023 selon la CCAM 313

I Structures anatomiques exposées et instrumentation employée lors des phases d'une NPRA 315

J Prototypes de réalité augmentée en NPRA 317

K Cas cliniques de NPRA assistée par réalité augmentée 319

L Liste non exhaustive de 123 méthodes de la littérature pour la détection et la description de points clés 321

M Visualisation des jumeaux numériques recalés sur un prototype de forme 325

Bibliographie 329

Liste d'abréviations

ACP Analyse en Composante Principale . . . 97
AdaLAM Adaptive Locally-Affine Matching . . . 218
ADN Acide DésoxyriboNucléique . . . 42
AFU Association Française d'Urologie . . . 73
AG Assistants au Guidage . . . 28
AGAST Adaptive and Generic Accelerated Segment Test . . . 145
AKAZE Accelerated-KAZE . . . 145
ALARA As Low As Reasonably Achievable . . . 42
ALIKE Accurate and LIghtweight KEypoint detection and descriptor . . . 145
ALIKED A LIghter Keypoint and descriptor Extraction network via Deformable transformation . . . 145
AML AngioMyoLipome . . . 70
ANRT Association Nationale de la Recherche et de la Technologie . . . 3
ANVISA Agência Nacional de VIgilância SAnitária . . . 120
API Application Programming Interface . . . 184
ASA American Society of Anesthesiologists . . . 16
BEBLID Boosted Efficient Binary Local Image Descriptor . . . 145
BFM Brute Force Matching . . . 99
BIC Base d'Images Clés . . . 211
BM Block Matching . . . 109
BNC Bayonet Neill–Concelman . . . 101
BoW Bag of Words . . . 222
BRDF Bidirectional Reflectance Distribution Function . . . 122
BRIEF Binary Robust Independent Elementary Features . . . 145
BRISK Binary Robust Invariant Scalable Keypoints . . . 145
BSDF Bidirectional Scattering Distribution Function . . . 123
BSSRDF Bidirectional Surface Scattering Reflectance Distribution Function . . . 123
BSU Bed Side Unit . . . 33
CAO Chirurgie Assistée par Ordinateur . . . 38
Carare Cancers rares du rein de l'adulte . . . 73
CBAM Convolutional Block Attention Module . . . 132
CBCT Cone-Beam Computed Tomography . . . 95

CCAFU Comité de Cancérologie de l'Association Française d'Urologie . 72
CCAM Classification Commune des Actes Médicaux 79
CCFE Contextual Contrast Feature Extraction 132
CCR Carcinomes à Cellules Rénales 68
CCU Camera Control Unit 20
CD Critères de Déclenchement 249
CE Cross Entropy 241
CETON Chirurgie Endoscopique Transluminale par Orifice Naturel . . . 22
CFR Constant Frame Rate 102
CGI Chirurgie Guidée par l'Image 38
CHU Centre Hospitalier Universitaire 3
CIFRE Convention Industrielle de Formation par la REcherche 3
CLIP Contrastive Language Image Pre-training 225
CM Co-Manipulateurs 27
CMD Centre de Mouvement Distant 33
CMI Chirurgie Mini-Invasive 3
CMOS Complementary Metal-Oxide-Semiconductor 22
CNN Convolutional Neural Network 167
CNRS Centre National de Recherche Scientifique 3
CO Chirurgie Ouverte 3
COAM CO-Attention for conditioned image Matching 146
CoCoLoT Combining Complementary trackers in Long-term visual Tracking 229
CODE COherence based DEcision boundaries 218
CP Correction de la Perspective 211
CPD Coherent Point Drift 97
CPU Central Processing Unit 90
CS Critères de Sélection 248
CTA Centrale de Traitement d'Air 14
DBSCAN Density Based Spatial Clustering method of Applications with Noise 218
DCT Discrete Cosine Transform 129
DDD Domain-Driven Design 92
DDL Degrés de Liberté 32
DDN Drift Determination Network 230
DeDoDe Detect, Don't Describe - Describe, Don't Detect 146

DefSLAM Deformable Simultaneous Localisation and Mapping 223
DeIT Data-efficient Image Transformer 225
DEL Diode ÉlectroLuminescente 20
DELTA Dense Efficient Long-range 3D Tracking for Any video 224
DeP Descripteur Profond 146
DES Diplômes d'Études Spécialisées 4
DeS Digital eye Strain 130
DFG Débit de Filtration Glomérulaire 72
DINO knowledge DIstillation with NO labels 225
DISK DIScrete Keypoints 146
DP DisplayPort 101
DR Differentiable Rendering 226
DSO Direct Sparse Odometry 223
DSP DataSet Patient 232
DSPT DataSet Patient Test 232
DTW Dynamic Time Warping 129
DUSt3R Dense and Unconstrained Stereo 3D Reconstruction 190
DVI Digital Visual Interface 101
DWA Dilated Window Attention 132
EAU European Association of Urology 73
ECC Enhanced Correlation Coefficient 128
ED Embedded Deformation 223
EDID Extended Display Identification Data 102
EKF Extended Kalman Filter 221
EnCoV Endoscopy and Computer Vision 3
EP Échographie Peropératoire 46
EPO Érythropoïétine 58
EVSAC Extreme Value SAmple Consensus 218
EWMA Exponentially Weighted Moving Average 135
FAST Features From Accelerated Segment Test 99
FBDSD Feature-Based Deformable Surface Detection 218
FBP Filtered Back Projection 41
FCN Fully Convolutional Network 136
FDA Food and Drug Administration 120
FFD Free Form Deformation 186

FLOPS FLoating-point Operations Per Second . . . 240
FPGA Field-Programmable Gate Array . . . 103
FREAK Fast Retina Keypoint . . . 145
GaLAM Globally-and-Locally-Affine Matching . . . 219
GANet Graph Attention Network . . . 219
GFC Gaussian Field Criterion . . . 218
GFTT Good Feature To Track . . . 145
GLOF Guided Local Outlier Factor . . . 218
GMM Gaussian Mixture Model . . . 97
GMS Grid-based Motion Statistics . . . 99
GNN Graph Neural Network . . . 219
GPIO General Purpose Input/Output . . . 127
GPMM Gaussian Process Morphable Model . . . 175
GPS Global Positioning System . . . 127
GPU Graphics Processing Unit . . . 85
GRBF Gaussian Radial Basis Function . . . 97
GT Ground Truth . . . 151
GTC Gestion Technique Centralisée . . . 14
HDCP High-bandwidth Digital Content Protection . . . 102
HDMI High-Definition Multimedia Interface . . . 101
HL Harris-Laplace . . . 145
HMR Human Mesh Recovery . . . 178
HSV Hue-Saturation-Value . . . 132
IA Intelligence Artificielle . . . 3
IADE Infirmière Anesthésiste Diplômée d'État . . . 11
IBODE Infirmière de Bloc Opératoire Diplômée d'État . . . 11
IC Images Clés . . . 195
ICF Identifying point Correspondence Function . . . 218
ICG IndoCyanine Green . . . 48
ICP Iterative Closest Point . . . 97
IEF Iterative Error Feedback . . . 178
IES Illuminating Engineering Society . . . 121
IHM Interface Homme-Machine . . . 30
IMDC International Metastatic Renal Cell Carcinoma Database . . . 72
IMU Inertial Measurement Unit . . . 226

INCa Institut National du Cancer . . . 73
IP Institut Pascal . . . 3
IRCT Insuffisance Rénale Chronique Terminale . . . 67
IRM Imagerie par Résonance Magnétique nucléaire . . . 40
IROCK International Radiosurgery Oncology Consortium for Kidney . . . 77
IU Interface Utilisateur . . . 32
JP Jumeau du Patient . . . 193
KLT Kanade-Lucas-Tomasi . . . 221
LATCH Learned Arrangements of Three Patch Codes . . . 145
LCT Long-term Correlation Tracking . . . 228
LEAP Long-term Effective Any Point tracking . . . 225
LIFO Last In First Out . . . 249
LMR Liver Mesh Recovery . . . 174
LMRe Learning Mismatch Removal . . . 219
LoFTR Local Feature TRansformer . . . 146
LOGO LOcality-guided Global-preserving Optimisation . . . 218
LOPO Leave-One-Patient-Out . . . 203
LPM Locality-Preserving Matching . . . 218
LRSM Low Rank Shape Model . . . 222
LSCM Least-Squares Conformal Mapping . . . 231
LSTM Long Short-Term Memory . . . 230
LSV-ANet Local Structure Visualisation-Attention Network . . . 219
LUCID Locally Uniform Comparison Image Descriptor . . . 145
LUT Look-Up Table . . . 41
MAE Mean Absolute Error . . . 202
MAEn Masked AutoEncoders . . . 225
MASt3R Matching And Stereo 3D Reconstruction . . . 190
MBConv Mobile inverted Bottleneck . . . 240
MFGA Multi-Facet Guided Attention . . . 225
MLESAC Maximum Likelihood Estimation SAmple Consensus . . . 218
MLP MultiLayer Perceptron . . . 98
MPR MultiPlanar Reconstruction . . . 41
MRC Maladies Rénales Chroniques . . . 67
MSD Maximal Self-Dissimilarity . . . 145
MSER Maximally Stable Extremal Regions . . . 145

N3Net Neural Nearest Neighbors Networks . . . 219
NAPSAC N Adjacent Points SAmple Consensus . . . 218
NASE Neighborhood Affine Subspace Embedding . . . 218
NIR Near-InfraRed . . . 48
NIRF Near-InfraRed Fluorescence imaging . . . 48
NMRC Neighborhood Manifold Representation Consensus . . . 218
NOTES Natural Orifice Transluminal Endoscopic Surgery . . . 22
NP Néphrectomie Partielle . . . 53
NPRA Néphrectomies Partielles Robot-Assistées . . . 45
NR-ICP Non-Rigid Iterative Closest Point . . . 175
NRSfM Non-Rigid Structure-from-Motion . . . 222
NR-SLAM Non-Rigid Simultaneous Localisation and Mapping . . . 223
NT Néphrectomie Totale . . . 73
OANet Order-Aware Network . . . 219
OCR Optical Character Recognition . . . 143
OMS Organisation Mondiale de la Santé . . . 73
ORB Oriented FAST and Rotated BRIEF . . . 99
PATS Patch Area Transportation with Subdivision . . . 146
PBPS Projection-Based Partial Selection . . . 97
PBR Physically-Based Rendering . . . 122
PCI Produits de Contraste Iodés . . . 70
PCIn Peripheral Component Interconnect . . . 103
PF Prototype de Forme . . . 193
PIC Potentielle Image Clé . . . 248
PIP Persistent Independent Particles . . . 225
PMD Progressive Mirror Detection . . . 132
PnP Perspective-n-Point . . . 97
PointCN Point net with Context Normalisation . . . 219
PTP Precision Time Protocol . . . 127
PTQ Post Training Quantisation . . . 243
PROSAC PROgressive SAmple Consensus . . . 218
QAT Quantisation-Aware Training . . . 243
QoE Quality of Experience . . . 130
RA Réalité Augmentée . . . 3
RANSAC RANdom SAmple Consensus . . . 97

RBF Radial Basis Functions . . . 175
RCM Remote Center of Motion . . . 33
RCP Réunions de Concertation Pluridisciplinaire . . . 73
ReLU Rectified Linear Unit . . . 136
ReTRo Real-Time Rotated descriptor . . . 220
RF RadioFréquences . . . 43
RGB-D Red Green Blue Depth . . . 133
RHRA Résection Hépatique Robot-Assistée . . . 118
RHU Recherche Hospitalo-Universitaire . . . 3
RMN Résonance Magnétique Nucléaire . . . 42
RNA Réseaux de Neurones Artificiels . . . 183
RoMa Robust dense feature Matching . . . 146
RPU Registration Process Understanding . . . 181
RV Réalité Virtuelle . . . 49
RVB Rouge Vert Bleu . . . 53
SA Surveillance Active . . . 75
SAM Segment Anything Model . . . 225
SAM-HQ Segment Anything Model in High Quality . . . 238
SAMo Spatial Attention Module . . . 132
SCGD Score de Comparaison Gauche-Droite . . . 53
SCRAMSAC Spatially Consistent RAndoM SAmple Consensus . . . 218
SD-DefSLAM Semi-Direct - Deformable Simultaneous Localisation and Mapping . . . 223
SDI Serial Digital Interface . . . 101
SDK Software Development Kit . . . 103
SfM Structure from Motion . . . 171
SfS Shape-from-Shading . . . 222
SfT Shape-from-Template . . . 221
SGBM Semi-Global Block Matching . . . 97
SGD Stochastic Gradient Descent . . . 146
SIC Single Image Calibration . . . 106
SIFT Scale Invariant Feature Transform . . . 128
SiLK Simple Learned Keypoints . . . 145
SLAM Simultaneous Localisation And Mapping . . . 99
SP SuperPoint . . . 146

SSF Static Specular Flow . . . 132
SSIM Structural Similarity Index Measure . . . 146
SSM Self-Similarity Matrix . . . 128
SSMo Statistical Shape Modelling . . . 175
SSMod State Space Model . . . 225
SSPI Salle de Surveillance Post-Interventionnelle . . . 12
SSS SubSurface Scattering . . . 123
SURF Speeded-Up Robust Features . . . 97
SURGAR SURGical Augmented Reality . . . 3
SurgIPC Surgical Image Perspective Correction . . . 231
TA Techniques Ablatives . . . 74
TC Télé-Chirurgie . . . 27
TDM TomoDensitoMétrie . . . 40
TAPIR Tracking Any Point with per-frame Initialisation and temporal Refinement . . . 225
TAPIP3D Tracking Any Point In Persistent 3D geometry . . . 224
TAPNext Tracking Any Point as Next token prediction . . . 224
TE Temps d'Écho . . . 43
TLE Target Localisation Error . . . 100
TOF Time Of Flight . . . 191
TPS Thin-Plate Splines . . . 221
TR Temps de Répétition . . . 43
TRE Target Registration Error . . . 94
TRL Test de Ratio de Lowe . . . 138
TSV Teinte-Saturation-Valeur . . . 132
TVE Target Visualisation Error . . . 96
UCA Université Clermont Auvergne . . . 3
UH Unités Hounsfield . . . 40
USAC Universal SAmple Consensus . . . 218
USB Universal Serial Bus . . . 127
UUPF Universal Update-Pacing Framework . . . 229
VAC Vergence-Accommodation Conflict . . . 130
VCI Veine Cave Inférieure . . . 78
VDF Volume Distribution Function . . . 123
VFC Vector Field Consensus . . . 218

VFM Vision Foundation Models 225
VFR Variable Frame Rate 102
VGG Oxford Visual Geometry Group 145
VGGT Visual Geometry Grounded Transformer 190
ViT Vision Transformer 225
VO Visual Odometry 225
VSLAM Visual Simultaneous Localisation and Mapping 222
WLS Weighted Least Squares 158
ZAC Zone à Atmosphère Contrôlée 12

Chapitre 1

Introduction

Dans ce premier chapitre, nous introduisons le cadre des travaux ainsi que certains aspects médicaux et technologiques nécessaires à la compréhension du contexte et de l'objectif de cette thèse. Tout d'abord, nous présentons le contexte clinique centré sur les applications chirurgicales et plus spécifiquement la chirurgie mini-invasive. Ensuite, nous abordons le contexte technique par la mise en lumière des différents outils technologiques d'aide au geste chirurgical, en commençant par les outils de planification jusqu'au guidage peropératoire. Nous poursuivons en montrant comment la réalité augmentée peut aider à surmonter les limites cliniques des outils actuels. Enfin, nous présentons un résumé des contributions que nous apportons pour améliorer la réalité augmentée en chirurgie mini-invasive des tissus mous de l'utérus et du rein.

Sommaire

1.1 Contexte des travaux 3
1.2 Contexte clinique 3
1.2.1 Généralités 3
1.2.1.1 Pratiques diagnostiques et thérapeutiques 4
1.2.1.2 Spécialités médicales et chirurgicales 4
1.2.1.3 Démographie médicale en France 4
1.2.1.4 Démographie de la chirurgie urologique en France . 6
1.2.2 Contexte chirurgical 8
1.2.2.1 Histoire de la chirurgie 8
1.2.2.2 Environnement chirurgical 10
1.2.2.3 Intervention chirurgicale 16
1.2.3 Chirurgie ouverte 17
1.2.4 Chirurgie mini-invasive 18
1.2.4.1 Chirurgie endoscopique 19
1.2.4.2 Chirurgie cœlioscopique 23
1.2.4.3 Chirurgie robotique 25
1.2.4.4 Chirurgie guidée par l'image 38
1.3 Contexte technique 40
1.3.1 Outils de planification préopératoire 40
1.3.1.1 Images tridimensionnelles en coupe 40
1.3.1.2 Reconstruction tridimensionnelle à partir de l'imagerie en coupe 44
1.3.1.3 Fabrication additive 45
1.3.2 Outils de guidage peropératoire 46
1.3.2.1 Échographie peropératoire 46
1.3.2.2 Imagerie par fluorescence 48
1.3.2.3 Réalité virtuelle en chirurgie 49
1.3.3 Réalité augmentée en chirurgie 50
1.4 Contributions et structure du manuscrit 53

1.1 Contexte des travaux

Cette thèse s'inscrit dans le cadre d'un partenariat scientifique public-privé établi entre l'équipe de recherche académique Endoscopy and Computer Vision (EnCoV) et la société SURGical Augmented Reality (SURGAR) dans le cadre d'une Convention Industrielle de Formation par la REcherche (CIFRE) mise en œuvre par l'Association Nationale de la Recherche et de la Technologie (ANRT).

L'objectif principal de cette collaboration est de développer et d'industrialiser des solutions de Réalité Augmentée (RA) destinées à la Chirurgie Mini-Invasive (CMI) des tissus mous, notamment pour les organes tels que l'utérus, le foie et le rein. Ce projet s'appuie sur une synergie scientifique et technologique issue d'une collaboration historique entre plusieurs acteurs clermontois comme l'Université Clermont Auvergne (UCA), l'Institut Pascal (IP) et le Centre Hospitalier Universitaire (CHU) de Clermont-Ferrand. Ce partenariat local est renforcé par des collaborations nationales avec des institutions comme le Centre National de Recherche Scientifique (CNRS), ainsi que par des projets de recherche tel que le programme de Recherche Hospitalo-Universitaire (RHU) Digital Urology 3D (ANR-21-RHUS-0015) coordonné par le CHU de Bordeaux.

Ce cadre interdisciplinaire constitue un terreau fertile pour l'innovation, combinant expertise clinique, compétences techniques et exploitation de données massives, avec une ambition centrale d'intégrer les outils de RA et d'Intelligence Artificielle (IA) au service de la chirurgie.

Dans les sections suivantes, le contexte clinique de cette thèse sera approfondi, en partant des fondements généraux de la pratique médicale et chirurgicale pour ensuite aborder les spécificités de la CMI, tout en intégrant une analyse comparative avec la Chirurgie Ouverte (CO) afin de contextualiser les défis actuels. Par la suite, le cadre technique sera exploré, avec une présentation détaillée des outils de planification préopératoire et des techniques de guidage peropératoire, menant à une discussion sur l'intégration et les apports de la RA en chirurgie. Enfin, les motivations scientifiques ainsi que les contributions majeures de cette thèse à l'avancement des applications de la RA dans le domaine de la CMI seront exposées.

1.2 Contexte clinique

1.2.1 Généralités

Le contexte clinique évolue continuellement, stimulé par des avancées significatives en recherche médicale et technologique. Cette dynamique influence directement les pratiques diagnostiques et thérapeutiques, enrichissant ainsi la pratique médicale et contribuant à l'amélioration de la qualité des soins.

1.2.1.1 Pratiques diagnostiques et thérapeutiques

Lors de la prise en charge des patients, le diagnostic constitue une étape importante et repose sur un ensemble de méthodes et d'outils, allant de l'examen clinique (anamnèse et examen physique) à des investigations sophistiquées comme l'imagerie médicale (tomodensitométrie, scintigraphie, etc.) ou les tests en laboratoire (biochimie sanguine, génétique, etc.).

Au regard du diagnostic établi, les traitements peuvent être de nature médicale ou chirurgicale. L'option médicale comprend l'administration de médicaments, les thérapies physiques et diverses interventions non invasives. Tandis que les traitements chirurgicaux impliquent des interventions invasives qui peuvent être indiquées lorsque les traitements médicaux ne suffisent pas ou en absence d'efficacité (en deuxième intention). Le choix thérapeutique est guidé par divers facteurs, tels que la nature et la sévérité de la pathologie, l'état général du patient et les recommandations issues des sociétés savantes et de la littérature scientifique.

1.2.1.2 Spécialités médicales et chirurgicales

A l'échelle internationale, l'activité clinique est de nature hétérogène tant en termes de formation que de pratique. En Europe, la Directive 2005/36/EC [Parlement Européen 2005] vise à harmoniser les qualifications professionnelles. En France, le cadre de formation est structuré autour des Diplômes d'Études Spécialisées (DES), définis par l'arrêté du 21 avril 2017 [Ministère de l'éducation nationale 2017], regroupant 31 spécialités médicales et 13 spécialités chirurgicales, listées en annexe A. Ces diplômes peuvent être acquis dans le cadre du troisième cycle des études de médecine [Ministère des solidarités et de la santé 2021].

Les spécialités sont généralement classées en trois groupes : généralistes, médicales ou chirurgicales [Conseil National de l'Ordre des Médecins 2024a]. En majorité, elles sont centrées sur un organe ou un système ; ensemble des organes qui ont une structure analogue comme le système vasculaire, le système musculaire et le système nerveux [Académie nationale de médecine 2024]. Certaines, comme la radiologie et l'imagerie médicale, font cependant exception et sont basées sur un ensemble de techniques spécifiques.

1.2.1.3 Démographie médicale en France

Du point de vue démographique, au 1er janvier 2024, 329 464 médecins sont inscrits au Tableau de l'Ordre des Médecins français [Conseil National de l'Ordre des Médecins 2024a]. Parmi eux, 237 300 sont en activité, soit une augmentation de l'effectif de plus de 10 % par rapport à 2010 et de plus de 1,4 % par rapport à 2023, comme illustré en figure 1.1. Les régions les plus densément peuplées, comme l'Île-de-France et Auvergne-Rhône-Alpes, concentrent la majorité des médecins en activité, voir figure 1.2. Enfin, comme le montre la

figure 1.3, les généralistes représentent 42,7 % des effectifs en activité, suivis par les spécialistes médicaux (44,9 %) et les spécialistes chirurgicaux (12,3 %).

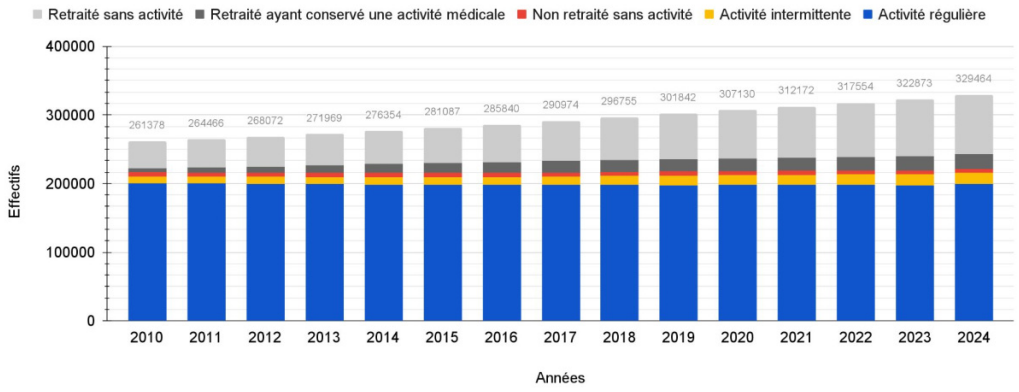


FIGURE 1.1 – Évolution des effectifs des inscrits au Tableau de l'Ordre des Médecins français par statut d'inscription depuis 2010. Source : générée par l'auteur à partir de [Conseil National de l'Ordre des Médecins 2024a].

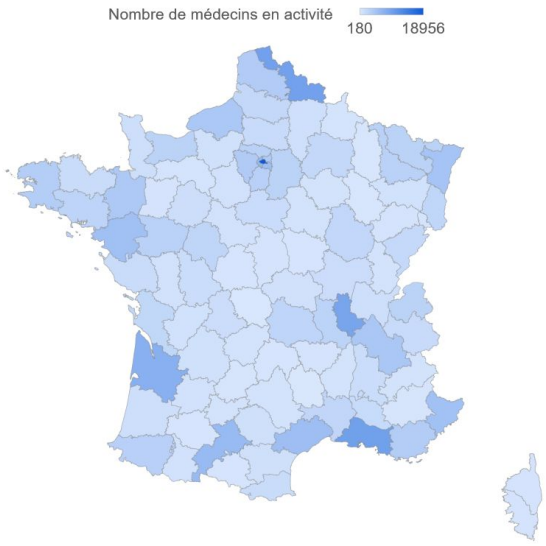


FIGURE 1.2 – Cartographie de la répartition en France métropolitaine des médecins en activité au 1er janvier 2024. Source : générée par l'auteur à partir de [Conseil National de l'Ordre des Médecins 2024a].

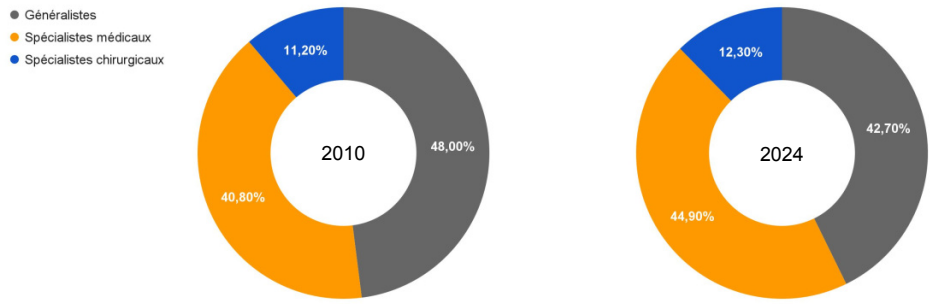


FIGURE 1.3 – Répartition des médecins en activité en France selon les groupes de spécialités en 2010 et 2024. (en gris) Généralistes. (en orange) Spécialistes médicaux. (en bleu) Spécialistes chirurgicaux. Source : générée par l'auteur à partir de [Conseil National de l'Ordre des Médecins 2024a].

1.2.1.4 Démographie de la chirurgie urologique en France

Dans le cas spécifique de la chirurgie urologique, spécialité d'intérêt pour ce manuscrit, les données d'effectifs en France au 1er janvier 2024, visibles en figure 1.4, montrent un taux de masculinisation de 88 %. Comme indiqué par le tableau 1.1, les chirurgiens urologues exercent pour la majorité d'entre eux en libéral (48 %), avec des pratiques mixtes (28 %) et salariées (24 %). La densité moyenne en France entière et en France métropolitaine est de 2,1 chirurgiens urologues pour 100 000 habitants avec une répartition départementale plutôt hétérogène, ainsi visible en figure 1.5.

Hommes			Femmes		
Libéraux	**Mixtes**	**Salariés**	**Libéraux**	**Mixtes**	**Salariés**
677	396	297	74	36	82
1370 *(87,7 %)*			192 *(12,3 %)*		
1562					

TABLE 1.1 – Effectifs par sexe et mode d'exercice en chirurgie urologique en France au 1er janvier 2024. Source : généré par l'auteur à partir de [Conseil National de l'Ordre des Médecins 2024b].

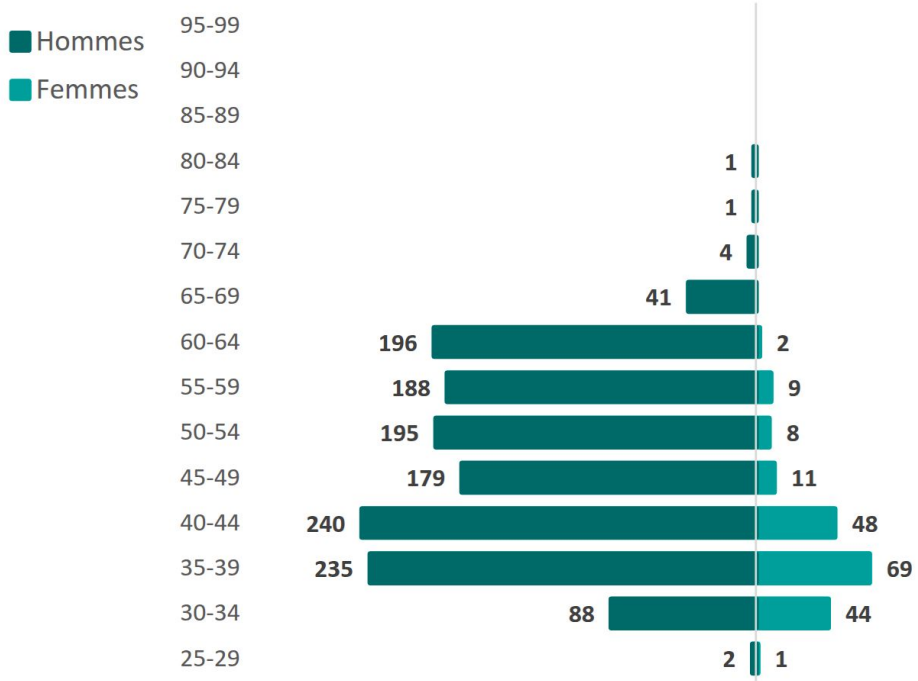


FIGURE 1.4 – Pyramides des âges en chirurgie urologique en France au 1er janvier 2024. Source [Conseil National de l'Ordre des Médecins 2024b].

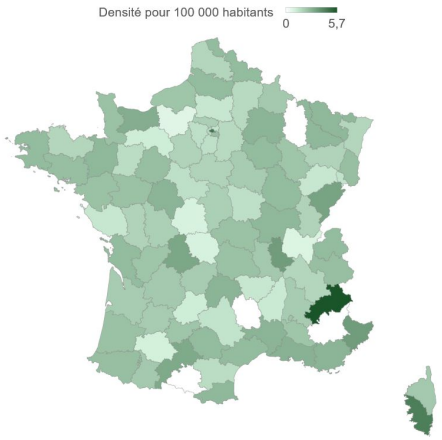


FIGURE 1.5 – Cartographie de la densité des chirurgiens urologues en activité au 1er janvier 2024 en France métropolitaine. Source générée par l'auteur à partir de [Conseil National de l'Ordre des Médecins 2024b].

1.2.2 Contexte chirurgical

1.2.2.1 Histoire de la chirurgie

La chirurgie, du grec keirourgia signifiant "travail manuel", englobe historiquement les méthodes, techniques et procédés comportant une action manuelle directe sur le corps humain, avec ou sans instrument, nécessitant une effraction du revêtement cutané [Académie nationale de médecine 2024]. Pratiquée depuis la Préhistoire (vers 3000 av. J.-C.), des preuves archéologiques (plus de 160 crânes trépanés retrouvés dans les Grandes Causses, France) attestent de gestes chirurgicaux comme les craniotomies et amputations des membres [Wauthy 2023]. Longtemps limitée par l'ignorance de l'anatomie, de la physiologie, l'hémorragie, la douleur et l'infection, la chirurgie a évolué parallèlement aux avancées des connaissances médicales et techniques.

Pendant l'Antiquité (3200 avant J.-C. à 476 après J.-C.), l'écriture a permis à certaines connaissances de traverser les siècles. Les traités égyptiens Ebers sur les plaies et Edwin Smith sur la neurochirurgie en sont les parfaits exemples [Wauthy 2023]. En Inde, la chirurgie oculaire et plastique est très développée. L'enseignement médico-chirurgical s'effectuait en école et sur animaux [Wauthy 2023]. En Grèce antique, les pratiques chirurgicales étaient sommaires jusqu'à l'époque d'Hippocrate (460 avant J.-C - 370 avant J.-C.) lorsque la chirurgie est intégrée à la médecine. À Rome, des techniques telles que la césarienne émergent. Galien (c. 129 - c. 201) est l'un des personnages clés de la chirurgie romaine du fait de ses découvertes anatomiques et ses innovations techniques. Par conviction religieuse, il fait cependant interdire la dissection anatomique entraînant peu à peu une perte de connaissances anatomiques en Occident [Wauthy 2023].

Au Moyen Age (476 - 1492), deux dynamiques opposées s'observent. Dans le monde arabo-musulman, les universités prospèrent et la chirurgie connaît un essor important grâce à des figures comme Abulcassis de Cordoue (c. 940 - c. 1013), auteur du Tarsif, une encyclopédie influente décrivant plus de 200 instruments chirurgicaux et des techniques novatrices. Des avancées significatives en chirurgie oculaire (Alhazen c. 965 - c. 1040), en intubation (Avicenne 980 - 1037) ou encore en antiseptiques (Rhazes 865 - 925) marquent cette période. Ibn Al-Nafis (c. 1213 - 1288) structure les interventions chirurgicales en trois stades (pré, per et postopératoire) et Avenzoar (c. 1073 - 1162) pose les bases de la chirurgie expérimentale avec la dissection et l'expérimentation animale [Wauthy 2023].

En Europe chrétienne, la chirurgie décline. Pendant cette période d'obscurantisme, la chirurgie est souvent perçue comme une pratique barbare, limitée aux barbiers et empiristes. Toutefois, l'influence musulmane en Espagne et en Sicile permet la traduction d'ouvrages médicaux en latin, amorçant un renouveau au sud de l'Europe. Les premières écoles de chirurgie européennes apparaissent, comme à Salerne et Bologne en Italie [Wauthy 2023].

En France, la faculté de médecine de Montpellier, fondée en 1220, est le produit des figures comme Guy de Chauliac (c. 1300 - 1368) [Wauthy 2023]. Cependant, des conflits persistants entre chirurgiens et médecins, amplifiés par l'influence du

clergé, freinent la reconnaissance des chirurgiens. Ces tensions culminent avec la disparition de l'école de chirurgie de Paris au XVII^{e} siècle, témoignant des entraves à la pratique chirurgicale en Europe chrétienne comparée à son essor dans le monde arabo-musulman.

A la Renaissance, la chirurgie marque un tournant grâce aux évolutions de nombreux domaines techniques. L'imprimerie permet une meilleure diffusion des connaissances, notamment celles héritées du monde arabo-musulman. L'apparition des armes à feu introduit de nouvelles blessures graves causées par les projectiles nécessitant des amputations fréquentes.

En Italie, des figures comme Léonard de Vinci (1452 - 1519), André Vésale (1514 - 1564) et Fallope (1523 - 1562) révolutionnent l'anatomie grâce à la pratique de dissections. En France, Ambroise Paré (1509 - 1590) se distingue comme un pionnier de la chirurgie moderne [Wauthy 2023].

À partir du XII^{e} siècle, la chirurgie s'éloigne progressivement de la médecine. Condamnée par l'Église et exclue du domaine des clercs, la chirurgie devient une pratique à part entière.

En 1660, la Confrérie de Saint-Côme est dissoute, privant les chirurgiens du droit d'exercer la médecine. Malgré cela, des chirurgiens comme Antoine Lambert (? - 1699) continuent de traiter des pathologies complexes. En 1686, la chirurgie regagne du prestige lorsque Charles-François Félix (1635 -1703), premier chirurgien de Louis XIV, opère avec succès une fistule anale du roi, rétablissant la confiance du public

En 1731, Louis XV fonde l'Académie royale de chirurgie, qui devient un centre de formation et d'innovation. En 1743, sous l'impulsion de Germain Pichault de La Martinière (1697 - 1783), chirurgiens et barbiers sont définitivement séparés et l'égalité entre médecins et chirurgiens est reconnue. La Martinière développe des écoles de chirurgie en France et introduit la pratique sur cadavres [Wauthy 2023].

À la Révolution, la distinction entre médecine et chirurgie disparaît, unifiant les enseignements et marquant une étape essentielle vers la médecine moderne. Des écoles de santé sont créées pour former des officiers de santé.

Après la chute de l'Empire, la chirurgie progresse passant d'individuelle (au domicile de l'opéré) à hospitalière avec des figures marquantes comme Guillaume Dupuytren (1777 - 1825) spécialiste des résections osseuses et anévrismes, Joseph Récamier (1774 - 1852) fondateur de la gynécologie moderne [Wauthy 2023]. Des collaborations entre chirurgiens et industriels naissent pour fabriquer des instruments chirurgicaux (exemple de la maison Charrière en France). Cependant, au XIX^{e} siècle, les infections provoquent une régression dramatique, notamment lors des guerres de Crimée et d'Italie, où les amputations crurales atteignent 85 % de mortalité.

L'apparition de l'anesthésie révolutionne les interventions chirurgicales avec l'utilisation de l'éther dès 1846 par William Morton (1819 - 1868). Les progrès antiseptiques initiés par Semmelweis (1818 - 1865) et développés par Lister (1827 - 1912) réduisent considérablement la mortalité [Wauthy 2023]. La fin du XIX^{e} siècle voit l'avènement de l'asepsie, avec l'introduction des gants, de l'autoclave et des blocs opératoires, posant les jalons de la chirurgie moderne.

Avant la Première Guerre mondiale, les grandes techniques chirurgicales prennent forme, notamment celles concernant l'abdomen, le tube digestif, le thorax et la gynécologie. Pendant la guerre, les blessures des "gueules cassées" stimulent le développement de la chirurgie réparatrice. De nouvelles technologies apparaissent avec l'électrification comme le bistouri électrique, la diathermie chirurgicale et la thérapie par laser.

Dans la seconde moitié du XX^{e} siècle, la chirurgie connaît une révolution avec l'essor des greffes d'organes : reins, moelle osseuse, foie, puis cœur et poumons, avec la première transplantation cardio-pulmonaire réalisée en 1982 par Christian Cabrol (1925 - 2017). Ces avancées s'appuient sur le développement des techniques d'imagerie qui transforment la pratique chirurgicale moderne.

En 2001, l'opération Lindbergh marque une avancée majeure avec la première télé-chirurgie réussie, réalisée entre New York et Strasbourg. En 2008, la première transplantation de deux bras entiers est effectuée en Allemagne par une équipe dirigée par Christoph Hijhnke et Edgar Biemer. Deux ans plus tard, en 2010, l'Espagne marque une nouvelle étape avec la première greffe complète de visage, menée par Joan Pere Barret et son équipe.

En neurochirurgie, les progrès sont significatifs grâce à la "chirurgie éveillée", qui permet des interventions plus précises sur des patients conscients, réduisant les risques de séquelles postopératoires.

En d'autres termes, l'essor de la chirurgie moderne n'est pas seulement une histoire technique mais aussi une histoire sociale dans le contexte plus général des sociétés modernes et industrialisées. Couramment les termes de "révolution chirurgicale des trois A" sont employés pour Anesthésie (à partir de 1846) avec Horace Wells (1815 - 1848) et William Morton (1819 - 1868), Antisepsie (à partir de 1867) avec Joseph Lister (1827 - 1912) et Asepsie (à partir de 1886) avec Louis Pasteur (1822 - 1895). Cette révolution a permis aux chirurgiens d'opérer plus largement, sans être limité par la douleur ou les risques infectieux. Le chirurgien peut alors prendre son temps et acquérir d'autres qualités comme la minutie et la précision.

1.2.2.2 Environnement chirurgical

Du point de vue infrastructurel, au XVIII^{e} siècle, une intervention chirurgicale se déroulait avec un risque infectieux élevé, où le danger des mains sales et des instruments infectés était ignoré. L'opération s'effectuait à domicile et pour les premiers chirurgiens hospitaliers au milieu de la salle commune des malades.

Avec la découverte des microbes, le contrôle aseptique du champ opératoire (stérilisation du matériel chirurgical) devient crucial, il s'étend aux chirurgiens eux-mêmes avec le port de gants puis de blouse et du masque.

Fort de ces connaissances, le bloc opératoire voit le jour en Allemagne et en Suisse à la fin du XIX^{e} siècle. Il se compose de plusieurs salles d'opérations indépendantes, reliées à des pièces adjacentes pour les activités de microbiologie et de stérilisations-désinfections. En France, le premier bloc opératoire est construit à l'Hôpital Necker à l'initiative d'Antonin Gosset (1872 - 1944) qui crée aussi celui de la Salpêtrière en 1912.

Comme illustré sur la figure 1.6, jusqu'en 1950 l'activité des salles d'opération est dépendante de la lumière du jour. L'éclairage opératoire aussi appelé scialytique va révolutionner le bloc opératoire et permettre de travailler en espace clos.

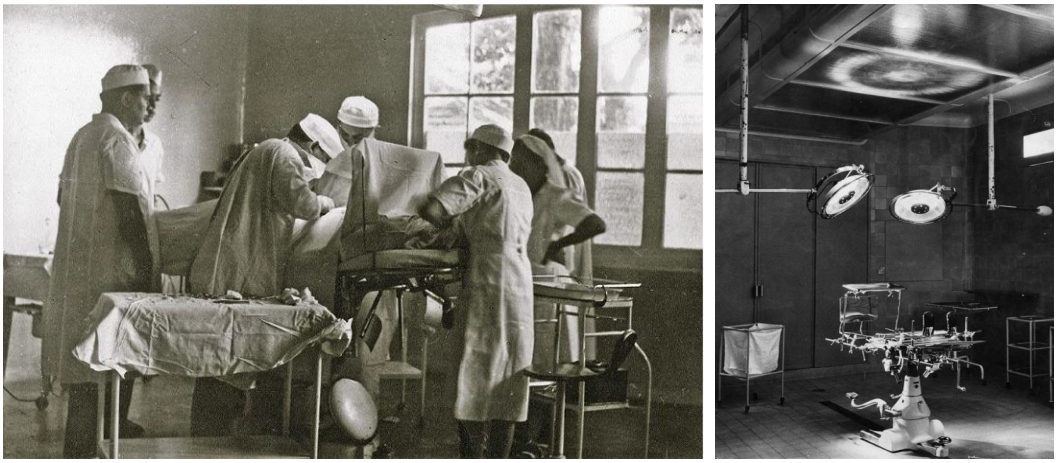


FIGURE 1.6 – (à gauche) Salle d'opération avant 1950 éclairée à la lumière du jour. (à droite) Salle d'opération après 1950 éclairée avec de la lumière artificielle, le scialytique. Source [Dreyer 2009].

Aujourd'hui, les spécialités chirurgicales s'exercent majoritairement dans des structures aux moyens humains et techniques importants (cliniques, centres hospitaliers et centres hospitaliers régionaux et universitaires), avec des objectifs diagnostiques (biopsie) ou thérapeutiques (excision de tumeurs, réparation de fracture ou transplantation). La chirurgie moderne est très dépendante de l'imagerie préopératoire et dans certains cas peropératoire. Elle nécessite un environnement spécifique ainsi qu'une équipe chirurgicale multidisciplinaire composée sans s'y restreindre du chirurgien, de son assistant, d'un anesthésiste, d'une Infirmière Anesthésiste Diplômée d'État (IADE) et d'une Infirmière de Bloc Opératoire Diplômée d'État (IBODE). Le bloc opératoire représente un des secteurs majeurs en tant que pièce maîtresse du plateau technique d'un établissement hospitalier. Par exemple, le CHU de Clermont-Ferrand dispose d'un plateau technique composé de trois blocs opératoires avec un total d'une trentaine de salles d'opérations [Centre Hospitalier Universitaire de Clermont-Ferrand 2024].

Le bloc opératoire est une enceinte dédiée aux actes invasifs réalisés quelles que soient la modalité et la finalité, en ayant recours aux équipements adéquats et en regroupant toutes les compétences médicales et paramédicales requises pour assurer la sécurité des patients [Gandjbakhch 2009]. Il se compose de cinq zones principales, schématisées sur la figure 1.7 :

1. Aire d'arrivée et de contrôle des patients, du personnel et des fournitures médicales ;
2. Locaux du personnel comprenant vestiaires, services sanitaires et salles de repos ;
3. Bloc chirurgical qui englobe les salles d'opérations ;
4. La Salle de Surveillance Post-Interventionnelle (SSPI) ;
5. Les locaux de soutien comprennent les salles d'utilités, de dépôts, de lavage et de préparation des instruments, le local d'anesthésie et celui de l'entretien ménager.

Tout est mis en œuvre pour garantir la sécurité des patients et prévenir les infections superficielles (peau et tissus sous-cutanés) et profondes (fascias, muscles et organes) du site opératoire. Selon l'enquête Prévalence 2022 [Santé publique France 2023], les infections du site opératoire représentent la troisième cause d'infections nosocomiales (14,3 %) après les infections urinaires (28,0 %) et les infections pulmonaires (16,3 %). Elles touchent environ un patient sur dix-huit et entraînent une augmentation significative de la durée et du coût d'hospitalisation. Bien que rares, elles sont directement ou indirectement responsables de décès dans environ 4 % des cas en 2017 [Deniau 2020]. La classification d'Altemeier [Altemeier 1985], élaborée en 1984, reste encore aujourd'hui une référence pour catégoriser les interventions chirurgicales en quatre classes selon le risque de contamination et d'infection postopératoire, comme détaillé en annexe B.

Sur le plan technique, le bloc opératoire est défini comme une Zone à Atmosphère Contrôlée (ZAC), spécialement conçue pour mettre en œuvre le principe d'asepsie progressive. Cet environnement utilise une série de barrières visant à réduire le risque de contamination de la plaie opératoire. L'objectif est de prévenir l'introduction de particules et de micro-organismes grâce à des verrous aérauliques ou douanes positionnés entre les locaux contigus, tout en maintenant un gradient de pression décroissant de l'intérieur vers l'extérieur du bloc opératoire, figure 1.8. Chaque zone contiguë est maintenue en surpression d'environ 15 pascals ($1,5 \times 10^{-4}$ bar).

Au sein de la ZAC, les sources de contamination sont majoritairement de deux natures : le vivant principalement représenté par l'humain et l'inerte comprenant les surfaces, dispositifs et poussières. Ces contaminants nécessitent une gestion rigoureuse pour préserver un environnement propre et/ou stérile. L'accès au bloc opératoire est ainsi strictement contrôlé et soumis à des règles d'hygiène imposant le port d'une tenue réglementaire : tunique, pantalon, coiffe, masque et chaussures spécifiques. Ce protocole vestimentaire vise à limiter les risques de contamination aéroportée.

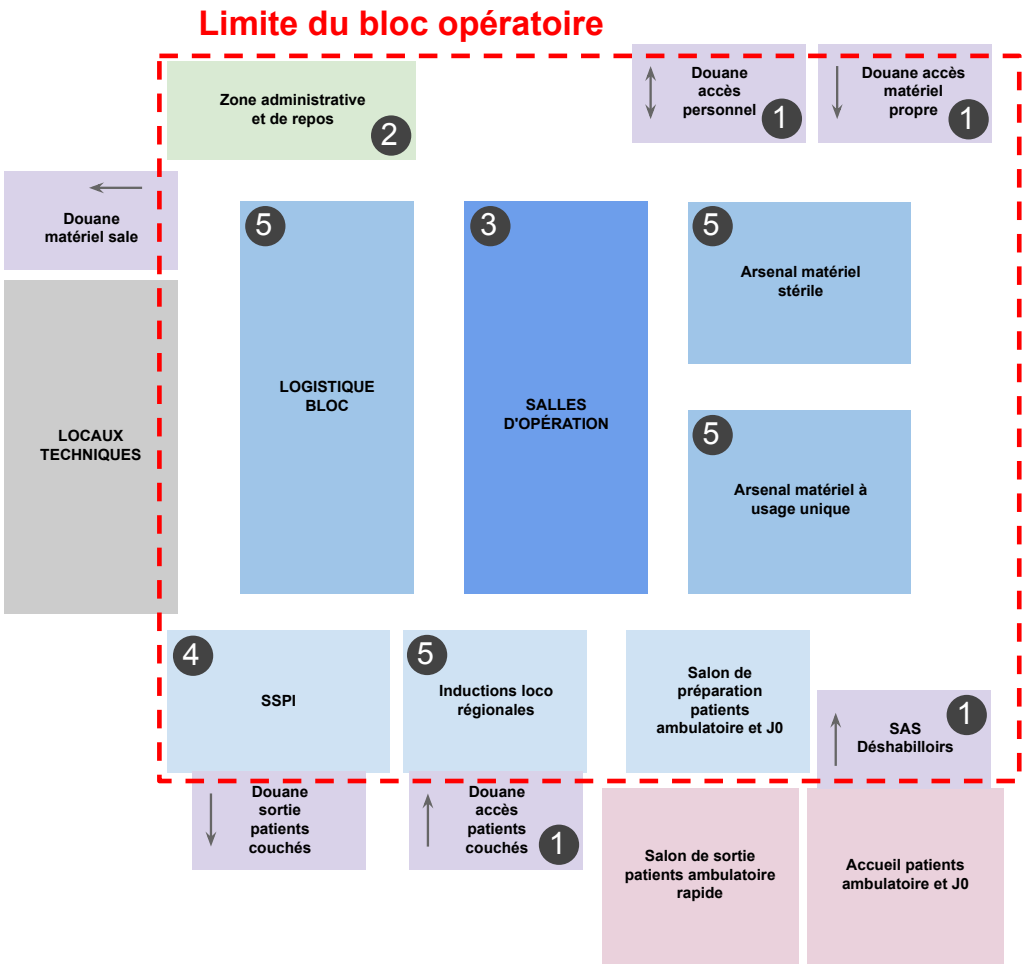


FIGURE 1.7 – Représentation schématique de l'agencement du bloc opératoire. Source : générée par l'auteur à partir de [Agence Régional de Santé Auvergne-Rhône-Alpes 2018].

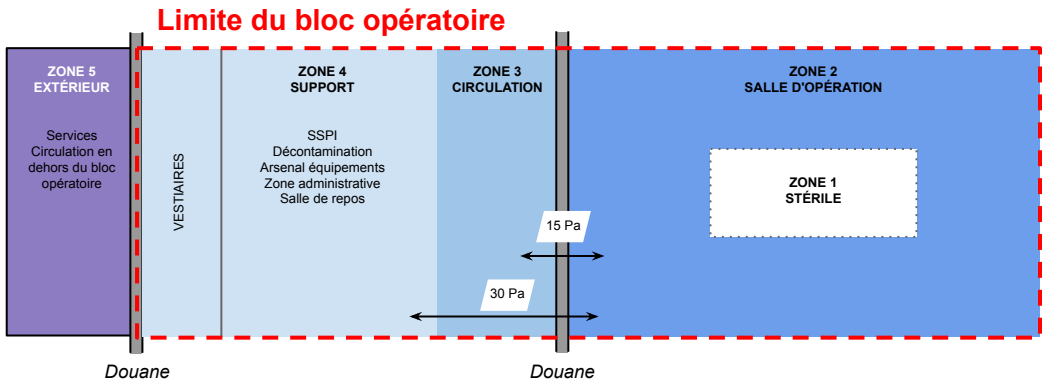


FIGURE 1.8 – Illustration de l'asepsie progressive du bloc opératoire. Les zones contiguës sont en surpression et dans certains cas séparées par des douanes. Source : auteur.

En complément, un système de traitement d'air aussi appelé Centrale de Traitement d'Air (CTA) est chargé du renouvellement constant en air neuf (environ six volumes par heure), du brassage homogène de l'air et du maintien de la suppression. Parallèlement, la CTA assure également la régulation des paramètres environnementaux, tels que la température (19-26 °C) et l'humidité (45-65 %) afin de garantir la protection du patient, le confort des occupants et l'absence de turbulences en zone stérile. Ces données environnementales sont reportées sur le système de Gestion Technique Centralisée (GTC). Cette surveillance continue permet une réactivité immédiate en cas de déviation des normes.

Les dispositifs techniques associés aux bonnes pratiques permettent de réduire efficacement l'aérobiocontamination, définie comme la contamination aéroportée due à la présence de particules en suspension. La cinétique d'aérobiocontamination évolue selon l'intensité de l'activité en salle et la nature de l'intervention. Une fois l'intervention terminée, un nettoyage approfondi est suivi d'un temps de repos (cinétique de décontamination), qui permet à la salle de retrouver son état nominal avant une nouvelle utilisation [Veyres 2014].

Sur le plan normatif, les infrastructures des blocs opératoires doivent répondre à des exigences strictes dans divers domaines techniques. Cela inclut la radioprotection (NF 15-160), les systèmes électriques (NF C15-100 et NF C15-211), les fluides médicaux (NF EN ISO 7396-1) et le traitement d'air (ISO 14644 et NF S90-351). Ces normes garantissent non seulement la sécurité des patients et du personnel mais aussi la pérennité des équipements et la qualité de l'environnement chirurgical.

Les salles d'opération se déclinent en deux grandes catégories selon le risque infectieux : les salles hyper-aseptiques ou zones de risque 4 (ISO 5 selon ISO 14644-1) et les salles aseptiques ou zones de risque 3 (ISO 7 selon ISO 14644-1). Les premières, dotées d'un flux laminaire unidirectionnel sont réservées aux interventions critiques telles que l'orthopédie ou l'ORL, où le risque de contamination doit être réduit

au minimum. Les secondes, équipées d'un flux turbulent maîtrisé conviennent aux interventions moins sensibles comme certaines opérations urologiques illustrées en figure 1.9.



FIGURE 1.9 – Salle d'opération aseptique (ISO 7 selon ISO 14644-1) lors d'une Néphrectomie Partielle Robot-Assistée (NPRA) au Centre Hospitalier Universitaire (CHU) de Bordeaux. Source : auteur.

Au sein de ces salles, l'organisation spatiale est optimisée pour assurer efficacité et sécurité. La zone centrale stérile est réservée à l'équipe chirurgicale et aux instruments, tandis que des zones périphériques accueillent des fonctions spécifiques comme l'anesthésie. Une salle d'opération typique d'une surface de plus de 40 m^2 [Gandjbakhch 2009] est pensée pour être facilement nettoyable, modulable et adaptable aux différents besoins des équipes chirurgicales.

Un principe fondamental de gestion de l'espace est celui de la "salle vide". En dehors des interventions, les salles sont dépourvues d'équipements non essentiels afin de réduire les sources potentielles de contamination. Ce concept influence également la conception des dispositifs médicaux qui doivent s'intégrer harmonieusement à cet environnement exigeant. La modularité et l'interaction avec les utilisateurs et l'environnement sont des facteurs déterminants dès les phases de conception.

Enfin, toutes ces mesures convergent vers un objectif commun : garantir la sécurité et l'efficacité des interventions chirurgicales. Chaque intervention repose sur une coordination fluide entre les équipes et une maîtrise totale de l'environnement.

Malgré les avancées technologiques et médicales, cinq milliards de personnes n'ont pas accès à des soins chirurgicaux et anesthésiques sûrs et abordables quand nécessaire [The Lancet Commission on Global Surgery 2015]. Ces carences sont particulièrement prononcées dans les populations rurales et marginalisées des pays à

revenu faible ou intermédiaire. Les conséquences humaines et économiques sont considérables, allant de taux élevés de mortalité et de morbidité à une incapacité économique prolongée.

Par exemple, le risque de mortalité pour une césarienne est de 0,3 pour 1 000 naissances en Europe et Asie centrale, mais il est 1,3 à 4,7 fois plus élevé en Amérique latine, 10,0 à 35,7 fois en Asie du Sud et jusqu'à 100 fois plus en Afrique subsaharienne, où la qualité des soins chirurgicaux et anesthésiques reste tragiquement inégale [Sobhy 2019]. Réduire ces disparités constitue un enjeu majeur pour les systèmes de santé mondiaux, soulignant l'importance d'un accès équitable à des soins de qualité.

1.2.2.3 Intervention chirurgicale

L'organisation rigoureuse de l'environnement chirurgical se reflète également dans la structuration temporelle de l'intervention chirurgicale, scindée en plusieurs phases ou "temps opératoires". Ces étapes permettent une gestion méthodique et sécurisée de chaque moment clé, englobant les soins préopératoires, peropératoires et postopératoires ainsi que la préparation spécifique des zones opérées. Elles s'intègrent dans divers parcours de soin, allant des séjours hospitaliers prolongés aux chirurgies ambulatoires, où le patient peut regagner son domicile le jour même.

Les soins préopératoires sont cruciaux pour préparer le patient et garantir des conditions optimales pour l'intervention. Ils comprennent un examen approfondi de l'état général du patient, souvent évalué par des scores standardisés tels que le score American Society of Anesthesiologists (ASA) [Saklad 1941] et la réalisation de tests biologiques pour détecter d'éventuelles anomalies ou risques. Le consentement éclairé du patient est également recueilli après une explication détaillée de l'intervention, des risques et des bénéfices.

Le jour de l'intervention, une attention particulière est portée à la préparation de la peau dans la zone opératoire. Cela inclut des protocoles rigoureux de désinfection pour minimiser le risque d'infection.

Les interventions peuvent suivre plusieurs voies, principalement la CO ou la CMI. La CO conventionnelle (cf. section 1.2.3), comme la laparotomie, exige une incision importante pour un accès direct à la zone ciblée, ce qui augmente généralement les temps de récupération et les risques d'infection.

À l'inverse, la CMI (cf. section 1.2.4) utilise de petites incisions ou des orifices naturels pour introduire des instruments miniaturisés, comme en cœlioscopie ou en angioplastie. Elle offre des avantages significatifs, notamment une récupération plus rapide, une douleur postopératoire réduite et des risques d'infection moindres.

Une alternative hybride combine les deux approches pour tirer parti des avantages de chaque méthode selon les besoins spécifiques de l'intervention.

Après l'intervention, le patient est généralement transféré en SSPI, où il est étroitement monitoré pour s'assurer de sa stabilité. Cette phase est cruciale pour prévenir et gérer les complications potentielles, tels que des troubles respiratoires, hémodynamiques ou infectieux.

1.2.3 Chirurgie ouverte

La CO représente l'approche conventionnelle et historique (cf. section 1.2.2.1), qui consiste à pratiquer une incision suffisamment large pour accéder directement à l'organe ou à la région à traiter. Cette méthode très largement documentée dans la littérature, offre au chirurgien une vision directe et un retour tactile essentiel à la manipulation des structures anatomiques.

L'un des principaux avantages de la CO réside dans le retour sensoriel qu'elle procure. Le toucher permet au chirurgien d'évaluer la dureté des tissus, facilitant ainsi la distinction entre tissus sains et tumoraux. De plus, la perception de la force appliquée aide à juger de l'amplitude des mouvements, réduisant ainsi le risque de lésions tissulaires ou vasculaires.

Les chirurgiens expérimentés peuvent ainsi discerner au toucher le degré d'invasion tumorale grâce à la sensibilité de leurs doigts. La vision binoculaire complète du site opératoire améliore également la perception spatiale et la réalisation des gestes.

La CO est particulièrement indiquée dans les interventions nécessitant un contrôle maximal ou une réaction immédiate, comme la neurochirurgie illustrée en figure 1.10 ou certaines chirurgies d'urgence [Zhao 2022c]. Elle demeure aussi l'alternative privilégiée en cas de complications survenant lors d'une CMI, notamment en cas d'hémorragie, c'est la conversion (de la CMI vers la CO).



FIGURE 1.10 – Installation au bloc opératoire de neurochirurgie. Source : Hôpital Lariboisière, Assistance Publique - Hôpitaux de Paris (AP-HP).

Bien qu'elle implique un temps de récupération plus long, un risque accru de complications et des conséquences esthétiques dues à des incisions souvent plus visibles par rapport aux techniques mini-invasives, la CO reste parfois privilégiée, notamment chez les patients présentant des comorbidités ou un risque anesthésique élevé [Zhao 2022c].

Par ailleurs, son apprentissage est essentiel et obligatoire dans de nombreux pays. Bien que la courbe d'apprentissage soit plus courte que pour la CMI, elle constitue un socle fondamental pour la formation des chirurgiens via une meilleure exposition des structures anatomiques et facile donc l'acquisition des compétences techniques.

Malgré les avancées technologiques et l'essor de la CMI, la CO reste un pilier incontournable de la pratique chirurgicale moderne. Dans certains cas, notamment pour traiter des pathologies complexes ou des tumeurs difficiles d'accès, elle demeure irremplaçable.

1.2.4 Chirurgie mini-invasive

Depuis le XX^{e} siècle, la CMI s'est largement développée, révolutionnant les pratiques chirurgicales en permettant des interventions précises et ciblées avec un impact réduit sur les tissus environnants. Contrairement à la CO, qui nécessite de larges incisions pour accéder à l'organe cible, la CMI repose sur l'utilisation de petites incisions ou sur l'accès par des orifices naturels. Cette approche repose sur des dispositifs d'imagerie, initialement conçus pour le diagnostic, tels que les imageurs à rayons X, l'échographie et les caméras endoscopiques. Ces outils permettent aux chirurgiens de compenser la perte de vision directe du site opératoire en affichant des images en temps réel sur un écran externe.

La CMI s'est adaptée à de nombreuses spécialités chirurgicales. En chirurgie vasculaire et radiologie interventionnelle, illustrée en figure 1.11, elle permet la pose de prothèses par voie endovasculaire, la neuro-embolisation cérébrale ou encore la chimio-embolisation de tumeurs hépatiques, le tout sous contrôle radiologique. Dans le domaine digestif, l'endoscopie facilite l'exploration et le traitement de diverses affections gastro-intestinales grâce à l'introduction d'une caméra endoscopique dans l'œsophage, l'estomac ou le côlon. En gynécologie, l'hystéroscopie permet d'intervenir directement dans l'utérus sans incision, tandis qu'en orthopédie, l'arthroscopie simplifie la prise en charge des lésions articulaires comme illustré en figure 1.11. La cystoscopie en urologie et la cœlioscopie en chirurgie abdominale, illustrent également l'étendue des applications de ces techniques mini-invasives.



FIGURE 1.11 – (à gauche) Exemple d'un acte en radiologie interventionnelle. Source : clinique ELSAN Santé Atlantique. (à droite) Chirurgie mini-invasive du poignet par arthroscopie. Source : clinique du sport de Bordeaux.

Les bénéfices de la CMI sont nombreux. Cette approche permet une réduction notable de la douleur postopératoire, ce qui limite le recours aux antalgiques puissants. Elle entraîne également une moindre perte sanguine, réduisant ainsi la nécessité de transfusions. L'un de ses atouts majeurs réside dans l'aspect esthétique, car les cicatrices laissées par ces interventions sont bien plus discrètes que celles des CO. De plus, la diminution des traumatismes tissulaires favorise un temps d'hospitalisation et une récupération plus courts, permettant aux patients de reprendre plus rapidement leurs activités quotidiennes. Enfin, le risque d'infections est réduit, puisque les organes internes sont moins exposés à l'environnement extérieur pendant l'intervention.

Toutefois, malgré ces nombreux avantages, la CMI présente certaines limites. Sa mise en œuvre nécessite une courbe d'apprentissage plus longue pour les chirurgiens, qui doivent suivre une formation spécifique. L'équipement requis, intégrant des technologies de pointe, engendre un coût élevé, ce qui restreint son accessibilité à certains établissements hospitaliers. De plus, certaines structures de santé, notamment dans les pays en développement, ne disposent pas des infrastructures adaptées pour permettre l'adoption généralisée de ces techniques. Enfin, malgré les avancées technologiques, certaines zones anatomiques demeurent difficilement accessibles par cette méthode, ce qui peut limiter son application à certains types d'interventions.

1.2.4.1 Chirurgie endoscopique

Parmi les différentes techniques de CMI, l'endoscopie occupe une place centrale. Cette approche permet d'explorer l'intérieur du corps humain en introduisant une caméra dans les cavités naturelles ou à travers de petites incisions. Dans sa conception la plus simple, un endoscope se compose d'un guide optique et d'un système d'éclairage.

Le terme “endoscope” a été employé pour la première fois en 1855 par l'ingénieur-opticien Charles Chevalier (1804 - 1859), en référence au premier endoscope optique développé en 1852 par le chirurgien urologue Antonin Jean Desormeaux (1815 - 1894). Au début de l'endoscopie, un défi majeur subsistait : l'éclairage. En effet, les premières tentatives utilisaient des bougies et des lampes à huile, ce qui limité considérablement l'efficacité de l'exploration.

Un tournant décisif a été franchi vers 1895 avec la mise au point du premier endoscope auto-éclairé par John Macintyre (1857 - 1928) à la Glasgow Royal Infirmary, l'un des premiers hôpitaux équipés de l'électricité. Par la suite, les années 1930 ont marqué une avancée significative avec le développement du premier gastroscope semi-flexible, conçu pour l'exploration de l'estomac. À la fin des années 1950, l'introduction des fibres optiques a révolutionné la discipline en donnant naissance aux fibroscopes, des endoscopes entièrement flexibles qui ont considérablement élargi le champ d'application de cette technique. Enfin, les progrès de la vidéo et la miniaturisation des caméras dans les années 1980 ont ouvert la voie à la chirurgie endoscopique moderne, en permettant une visualisation en temps réel des interventions sur écran.

Aujourd'hui, lors de l'acte de vidéo-endoscopie médicale, une colonne vidéo-endoscopique est utilisée, illustration en figure 1.12. Elle intègre une unité d'alimentation et de contrôle de la caméra (Camera Control Unit (CCU) en Anglais), un processeur vidéo chargé du traitement numérique du signal provenant de la caméra, incluant la correction des couleurs, l'amélioration du contraste et la gestion de la balance des blancs. Une source de lumière froide, généralement basée sur une lampe xénon ou une Diode ÉlectroLuminescente (DEL) haute intensité, assure un éclairage homogène du champ opératoire en limitant les risques de brûlure des tissus. En complément, la colonne intègre divers systèmes auxiliaires dédiés à l'insufflation de dioxyde de carbone (CO_2), à l'aspiration des fumées chirurgicales et liquides, ainsi qu'à l'enregistrement et la transmission des images en temps réel vers le dispositif d'affichage.

L'image captée peut être obtenue sous la forme d'une vue unique ou d'une paire d'images gauche/droite, correspondant respectivement à un endoscope 2D (monoculaire) ou 3D (stéréoscopique), en fonction du type de caméra et du système optique employé. L'imagerie 2D, standard en endoscopie, offre une résolution élevée et un rendu fidèle des structures anatomiques. L'imagerie tridimensionnelle, obtenue à l'aide de caméras binoculaires ou par reconstruction stéréoscopique, permet une perception en relief et une meilleure appréciation des distances et des volumes.

En complément de cette unité technique, l'endoscope constitue l'élément d'exploration propre ou stérile inséré dans le patient. La caméra peut également être stérile et doit, dans ce cas, être autoclavable à haute température ce qui impose une conception hermétique. Il en existe deux types principaux, visibles en figure 1.13 : l'endoscope rigide et l'endoscope souple.

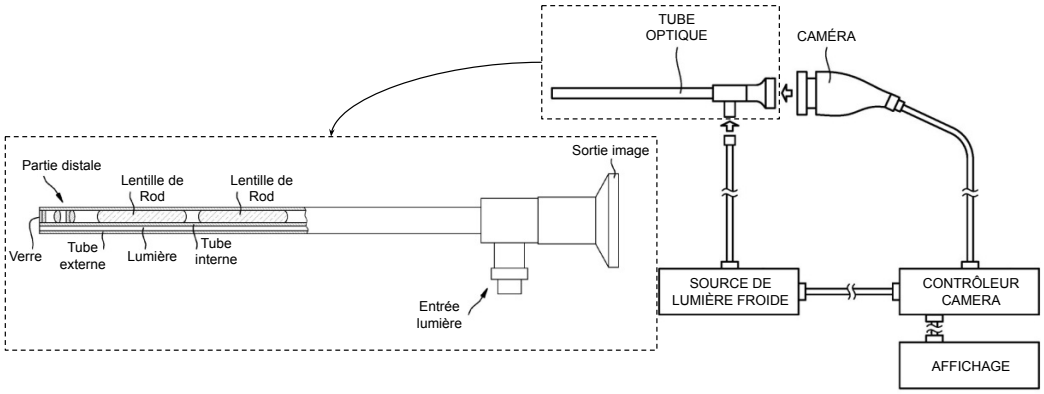


FIGURE 1.12 – Représentation schématique d'une colonne de vidéo-endoscopie équipée d'un endoscope rigide et de sa caméra amovible, d'une unité de contrôle de la caméra, d'une source de lumière froide et d'un système d'affichage. Source : générée par l'auteur à partir de [Amano 2018, Dahmen 2014].



FIGURE 1.13 – (à gauche) Un endoscope rigide composé de l'assemblage du tube optique et de la tête caméra. Source : Karl Storz. (à droite) Un endoscope flexible. Source : Olympus.

L'endoscope rigide est constitué d'un tube métallique d'un diamètre compris entre 1 et 12 millimètres et d'une longueur variant de 15 à 50 centimètres. Il remplit une double fonction : guider la lumière de la source vers la scène opératoire et collecter puis acheminer la lumière réfléchie vers le capteur d'imagerie. Pour ce faire, il intègre un premier canal dédié à la conduction de la lumière ainsi qu'un second contenant un train de lentilles de Rod dédié au transport de l'image, figure 1.12. Son angle de visée, généralement de 0 ou 30 degrés, influence l'exploration : le modèle 0 degré offre une vision axiale, tandis que le modèle 30 degrés permet une observation oblique, optimisant l'accès aux structures anatomiques complexes. Il est généralement couplé à une tête caméra équipée d'un capteur Complementary Metal-Oxide-Semiconductor (CMOS). Toutefois, certaines architectures, notamment en chirurgie robotique, intègrent directement un capteur d'imagerie au niveau distal, comme décrit dans la section 1.2.4.3. L'endoscope rigide est couramment utilisé pour les explorations articulaires (arthroscopie), l'examen de la vessie (cystoscopie) et l'exploration de la cavité abdominale (cœlioscopie).

L'endoscope souple, ou fibroscope, se distingue par l'intégration de fibres optiques qui assurent la transmission de la lumière et de l'image via un système de lentilles relayant l'information jusqu'au capteur. Il peut atteindre une longueur de 250 centimètres, ce qui le rend particulièrement adapté aux explorations des structures anatomiques profondes et sinueuses. Sa flexibilité, combinée à un contrôle mécanique ou électronique de son extrémité distale, permet une navigation précise à travers les conduits naturels du corps humain, comme les voies respiratoires (bronchoscopie), le tube digestif supérieur (gastroscopie) et le tube digestif inférieur (duodénoscopie, coloscopie).

La chirurgie endoscopique ne cesse d'évoluer, menant à des approches toujours plus innovantes, comme la Chirurgie Endoscopique Transluminale par Orifice Naturel (CETON) ou Natural Orifice Transluminal Endoscopic Surgery (NOTES) en Anglais. Mise en œuvre pour les premières fois entre 2006 et 2007, cette technique combine les avancées de l'endoscopie interventionnelle et de la chirurgie cœlioscopique pour accéder à la cavité abdominale par les orifices naturels du corps [Arulampalam 2009]. En éliminant les incisions externes, elle permet de réduire significativement les traumatismes postopératoires.

La CETON repose sur l'introduction d'un endoscope souple multicanaux par différentes voies : transgastrique, transvaginale, transvésicale ou transcolique. Cette approche présente de nombreux bénéfices, notamment une diminution des douleurs postopératoires, un temps de récupération plus court et l'absence quasi totale de cicatrices visibles [Boguet 2007, Kalloo 2007].

Toutefois, son développement requiert une collaboration étroite entre gastroentérologues et chirurgiens, ainsi qu'une amélioration continue des instruments et des protocoles de fermeture des incisions internes. De la CETON naissent de nombreuses attentes [Garud 2012] car elle représente une solution intéressante pour la CMI même si aujourd'hui la cœlioscopie reste la technique de référence pour accéder à la cavité abdominale.

1.2.4.2 Chirurgie cœlioscopique

La cœlioscopie constitue aujourd'hui une technique incontournable en CMI, offrant un compromis entre sécurité, efficacité et récupération rapide. Cette approche, illustrée en figure 1.14 permet d'intervenir sur divers organes abdominaux et thoraciques tout en réduisant l'impact chirurgical, tant pour le patient que pour l'établissement hospitalier.



FIGURE 1.14 – Installation de base en cœlioscopie. Le chirurgien opère à l'aide d'instruments insérés au travers de trocarts, en observant la scène au travers d'un écran externe. Source : B. Braun.

Depuis la réalisation de la première cholécystectomie cœlioscopique par Philippe Mouret (1938 - 2008) en 1987 [Litynski 1999], la cœlioscopie s'est imposée comme une technique incontournable en CMI, favorisant son adoption progressive par la communauté chirurgicale. Cette avancée s'inscrit dans une évolution plus large, précédée par d'autres réalisations majeures, notamment en 1962, 1972 et 1983 [Palmer 1962, Gomel 1974], qui ont posé les bases du développement de la cœlioscopie chirurgicale. En 1982, Kurt Semm (1927 - 2003) a réalisé la première appendicectomie cœlioscopique [Semm 1983], marquant un tournant décisif dans l'essor de la CMI et ouvrant la voie à son application à d'autres interventions complexes. L'expansion des indications de cette approche s'est poursuivie avec la première série de colectomies cœlioscopiques, publiée par Jacobs en 1991 [Jacobs 1991], démontrant ainsi son efficacité et son intérêt croissant.

L'amélioration continue des protocoles opératoires et des dispositifs chirurgicaux a contribué à une meilleure précision des dissections, garantissant des résultats oncologiques équivalents à ceux de la CO. Cette évolution a consolidé la place de la

cœlioscopie comme une alternative de choix dans de nombreuses spécialités chirurgicales [Himal 2002, St John 2020], renforçant ainsi son rôle dans la prise en charge chirurgicale moderne.

En 2025, la littérature scientifique témoigne de l'intérêt croissant pour cette approche, avec plus de 85 300 articles indexés sur PubMed contenant le terme "laparoscopy" dans leur titre. Cette multiplicité d'études atteste d'une adoption progressive de la cœlioscopie. Toutefois, son déploiement reste limité dans certaines spécialités par rapport à la CO et dans certains pays principalement en raison des coûts élevés du matériel requis.

Sur le plan technique, les interventions réalisées dans l'espace péritonéal nécessitent impérativement la création d'un espace de travail par insufflation gazeuse, généralement au dioxyde de carbone (CO_2), tandis que celles réalisées dans le thorax reposent sur l'affaissement pulmonaire. En chirurgie gynécologique, urologique et hépatique, la procédure repose sur un protocole bien défini [Hageman 2008, Larousse 2024] : après préparation du patient et installation des champs opératoires, une incision est réalisée pour permettre l'insertion d'un premier trocart. La cavité péritonéale est alors insufflée et maintenue en surpression (5 à 15 mmHg) afin de créer un pneumopéritoine qui éloigne la paroi des viscères et facilite l'accès chirurgical. L'endoscope relié à la colonne de cœlioscopie (cf. section 1.2.4.1) est ensuite introduit afin de réaliser une exploration initiale, puis des trocarts supplémentaires sont placés pour permettre l'introduction des instruments chirurgicaux. L'intervention est réalisée sous contrôle vidéo, après quoi les trocarts sont retirés, le gaz est évacué et les incisions sont suturées. Ces dernières, de taille réduite, limitent les traumatismes tissulaires et diminuent le risque de saignement et d'infection.

Les principaux atouts de la chirurgie cœlioscopique sont ceux de la CMI, ils résident dans la réduction du traumatisme opératoire, qui se traduisent par une récupération plus rapide et une diminution des complications postopératoires. La cicatrisation est accélérée, limitant ainsi la douleur et le risque d'infections. De plus, la réduction de la perte sanguine et la diminution de la morbidité postopératoire contribuent à améliorer le pronostic des patients [Okunrintemi 2016, Chen 2017, Carr 2019]. La diminution de la durée d'hospitalisation constitue un avantage supplémentaire, réduisant les coûts globaux de prise en charge. Certaines interventions mini-invasives peuvent même être réalisées en ambulatoire, en fonction de leur complexité. Malgré ces bénéfices, la CMI demeure techniquement exigeante. Sa pratique requiert une formation spécialisée et une expertise accrue du chirurgien, avec une courbe d'apprentissage longue. Les contraintes spatiales inhérentes à cette technique, telles que la restriction de l'amplitude des mouvements, l'inversion directionnelle due à l'effet fulcrum et l'absence de perception tactile, rendent certaines interventions plus complexes [Carr 2019]. De plus, la vision en deux dimensions ainsi que la modification de la continuité visuelle et de la proprioception altèrent la précision gestuelle, rendant plus difficile la reproduction des techniques de la CO. Ces limitations peuvent se traduire par un risque accru de lésions iatrogènes, notamment des lésions vasculaires, des hémorragies ou des résections tumorales incom-

plètes [Doddamani 2018]. Par ailleurs, la durée opératoire peut être prolongée, nécessitant une vigilance accrue du chirurgien afin de limiter les complications peropératoires [Kaplan 2016, Driessen 2017, Li 2021].

Malgré ces défis, la cœlioscopie continue de s'imposer comme une technique de référence dans de nombreuses spécialités. L'amélioration des outils chirurgicaux et des stratégies de formation permet aux praticiens d'optimiser leur apprentissage et de mieux gérer les complications potentielles. Le développement d'outils de guidage peropératoire (cf. section 1.3.2) facilite la réalisation des interventions en améliorant la visualisation et la précision des gestes [Robinson 2004].

Initialement limitée aux interventions gynécologiques et abdominales, la cœlioscopie s'est progressivement étendue à d'autres territoires anatomiques. En chirurgie digestive, elle est couramment utilisée pour la cholécystectomie, la colectomie ou la gastrectomie. Dans l'espace rétropéritonéal, elle permet d'effectuer des interventions telles que la néphrectomie ou la surrénalectomie. En thoracoscopie, elle est utilisée pour traiter les lésions pulmonaires ou réaliser des interventions sur l'œsophage. En chirurgie cardiovasculaire, elle a ouvert la voie à des gestes chirurgicaux complexes comme les pontages coronariens ou la réparation du canal artériel.

Ainsi, bien que la cœlioscopie présente certaines contraintes, elle représente une avancée majeure dans le domaine chirurgical. Son adoption croissante, soutenue par l'innovation technologique et l'optimisation des stratégies de formation, en fait aujourd'hui une alternative incontournable à la CO dans de nombreuses spécialités. La tendance actuelle laisse présager une démocratisation encore plus large de ces techniques, consolidant leur rôle dans l'amélioration des soins et du pronostic des patients. L'évolution constante des dispositifs endoscopiques et l'essor de la robotique chirurgicale contribuent à repousser les limites de la CMI.

1.2.4.3 Chirurgie robotique

La chirurgie robotique représente une avancée majeure en offrant la possibilité d'effectuer des interventions assistées par des systèmes robotisés. Ces technologies permettent soit une manipulation à distance par le chirurgien (télé-manipulation), soit une collaboration homme-machine pour réaliser des tâches spécifiques avec une précision accrue (chirurgie collaborative).

L'histoire de la chirurgie robotique remonte aux années 1980. En 1985, aux Etats-Unis, le Docteur Yin San Kwoh réalise l'une des premières interventions assistées par un robot lors d'une biopsie cérébrale mini-invasive sous tomodensitométrie [Kwoh 1988]. Parallèlement, le Docteur Patrick Kelly développe un cadre stéréotaxique motorisé et informatisé, ouvrant la voie aux futurs dispositifs robotiques intégrés en neurochirurgie [Kall 1985]. En mars 1989, à Grenoble, le Professeur Alim Benabid opère un patient en neurochirurgie stéréotaxique assistée par un robot [Lavallée 1989, Benabid 1992]. Cette intervention marque le début d'une nouvelle ère : le robot utilisé, développé au sein du laboratoire TIMC, devient par la suite le Neuromate, premier système robotisé à être industrialisé pour la neurochirurgie.

En 2024, environ 200 robots chirurgicaux sont référencés en annexe C, toutes spécialités confondues. Parmi eux, 50 % sont déjà commercialisés et 50 % sont encore en développement, comme illustré dans les figures 1.15 et 1.17. À l'échelle mondiale, les principaux acteurs de la robotique médicale sont localisés aux États-Unis (26 %), en Chine (23 %) et en France (13 %). Presque toutes les spécialités chirurgicales bénéficient désormais de solutions robotiques, témoignant d'une progression significative. En effet, entre 2022 et 2024, le nombre d'applications robotiques en chirurgie a enregistré une augmentation moyenne de 126 %, comme le montre la figure 1.16.

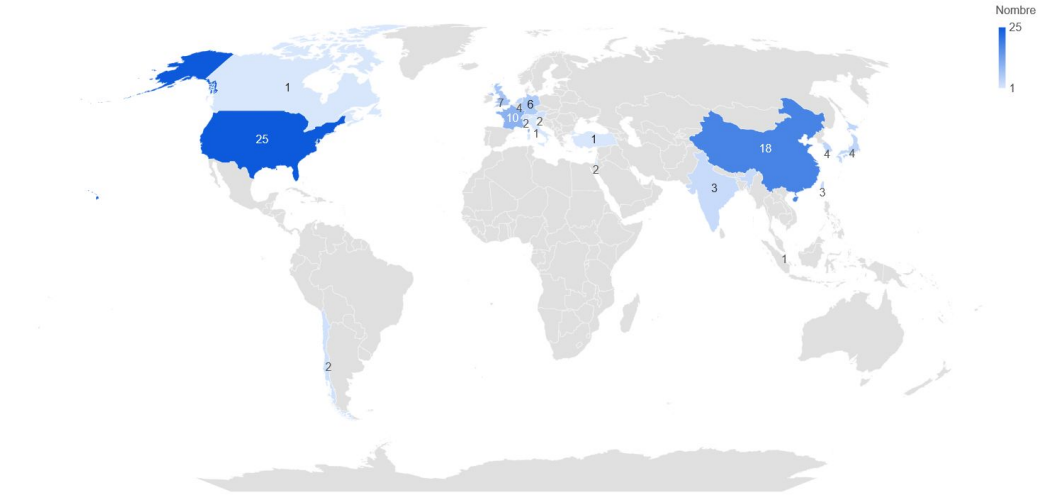


FIGURE 1.15 – Cartographie mondiale de la répartition des fabricants de robots chirurgicaux disponibles sur le marché en 2024. Source : auteur.

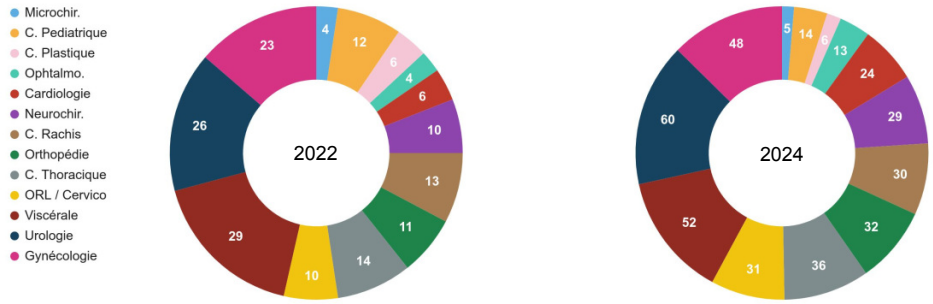
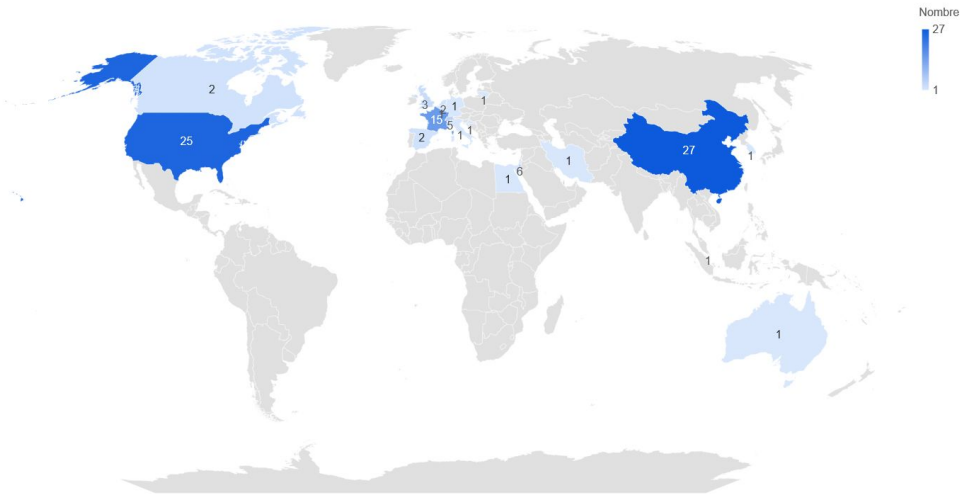


FIGURE 1.16 – Répartition des robots chirurgicaux disponibles sur le marché international selon les spécialités chirurgicales en 2022 et 2024. Source : auteur.


Nombre
27
1
25
2
27
1
1
1

main" pour le chirurgien ;

3. Les Assistants au Guidage (AG) facilitent le positionnement précis des instruments en suivant des plans préétablis, assurant une exécution fidèle à la planification chirurgicale.

Cette classification s'appuie sur les critères définis dans l'arbre décisionnel proposé par [Service d'Ingénierie Biomédicale 2022] et visible en figure 1.18. En 2024, la répartition des systèmes disponibles sur le marché entre les catégories TC, CM et AG est respectivement de 30,2 %, 14,6 % et 52,1 %. Les détails relatifs aux sous-groupes sont illustrés à la figure 1.19.

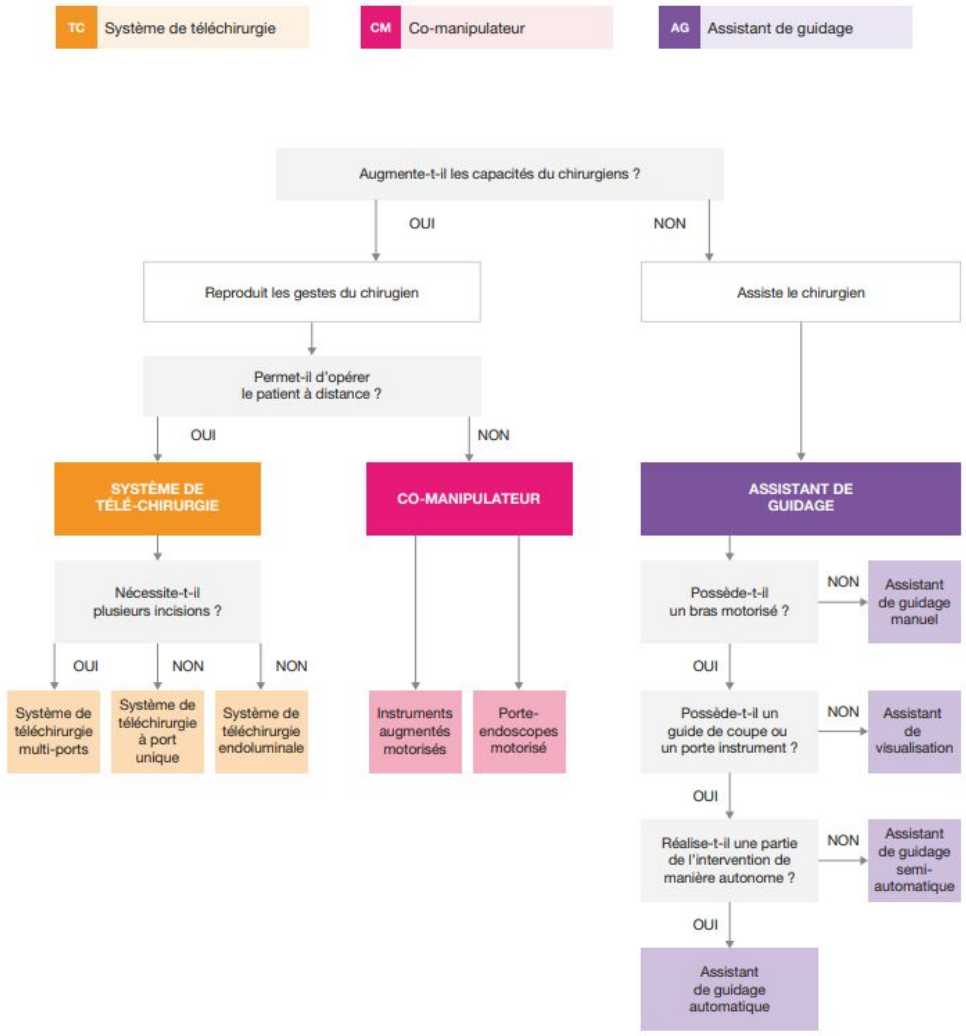


FIGURE 1.18 – Modalités de classification d'un robot chirurgical selon le type d'assistance apportée au chirurgien. Source : [Service d'Ingénierie Biomédicale 2024].

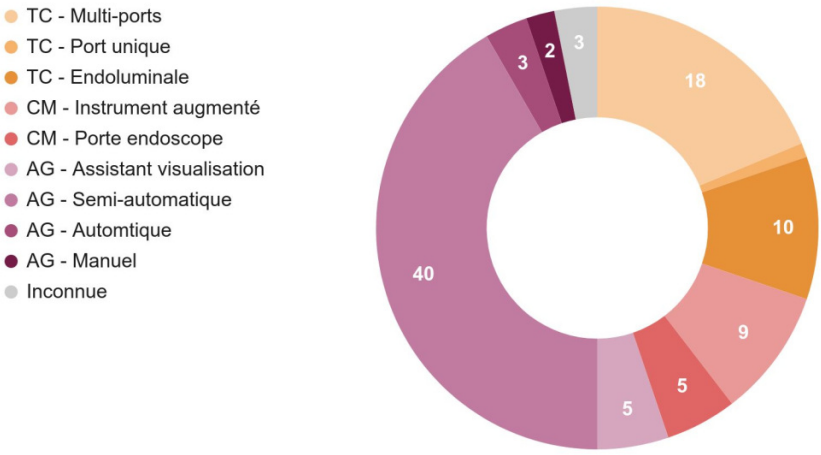


FIGURE 1.19 – Répartition des robots chirurgicaux disponibles sur le marché selon le type d'assistance apportée au chirurgien. Source : auteur.

En TC abdominale, l'utilisation d'un télémanipulateur robotique vise à surmonter les limites de cœlioscopie standard (cf. section 1.2.4.2). Le but est d'accroître la qualité et la précision des gestes chirurgicaux en réduisant les tremblements et en augmentant la dextérité du chirurgien [Ngu 2017, Haidegger 2022]. La cancérologie illustre parfaitement le besoin croissant de précision, amplifié par les avancées du diagnostic, qui permettent désormais le dépistage des formes infracliniques. Ainsi, la chirurgie d'exérèse d'un organe a progressivement laissé place à une approche plus ciblée, centrée sur l'exérèse de la tumeur [Martin 2020]. En outre, la TC robotique améliore l'ergonomie du poste de travail du chirurgien, qui passe d'une position de travail debout à une position assise, réduisant ainsi la fatigue.

Cependant, la chirurgie robotique engendre des défis organisationnels et impacte l'accès aux soins, notamment en raison de l'accessibilité limitée des équipements, de leur coût élevé et de leur répartition géographique hétérogène sur le territoire.

En TC abdominale, trois acteurs industriels dominent actuellement le marché avec des parts inégales : Intuitive Surgical, CMR Surgical, et Medtronic. Chacun proposant une plateforme robotique spécifique : Da Vinci, Versius, et Hugo respectivement, visible en figure 1.20.

En septembre 2024, les systèmes Da Vinci (Intuitive Surgical) ont réalisé plus de 16 millions d'interventions [Surgical 2024], contre un peu plus de 20 000 pour Versius (CMR Surgical) en mars, confirmant la domination de la société Intuitive Surgical sur le marché de la TC abdominale.

En d'autres termes, cela représente plus de 10 200 robots Intuitive Surgical en service dans le monde. Pour référence, le coût d'acquisition d'un Da Vinci varie entre 0,8 et 3,1 M $, avec des consommables facturés de 800 à 3 600 $ par intervention et des services annuels de 100 000 à 225 000 $ par robot [Surgical 2024]. Le ratio coût/bénéfice clinique est assez fragile.



FIGURE 1.20 – Plateformes de télé-chirurgie (en haut) Da Vinci Xi, Intuitive Surgical, (à gauche) Versius, CMR Surgical et (à droite) Hugo, Medtronic. Sources : Intuitive Surgical, CMR Surgical et Medtronic.

Sur le plan technique, une plateforme de TC abdominale, illustrée en figure 1.21, se compose généralement de trois éléments principaux placés en salle d'opération :

1. La console chirurgien permet de télé-manipuler les instruments et visualiser le site opératoire. Ce robot maître numérise les mouvements du chirurgien ;
2. Le robot patient porte les instruments et la caméra endoscopique. Ce robot esclave reproduit les gestes du chirurgien avec précision en adaptant l'échelle ;
3. La colonne robotique centralise et gère les équipements nécessaires à l'intervention (flux vidéo, lumière, insufflation de dioxyde de carbone, équipements d'électrochirurgie, etc.).

La console du chirurgien, illustrée en figure 1.22, intègre un affichage stéréoscopique, immersif ou ouvert, ainsi qu'une Interface Homme-Machine (IHM) assurant la transmission des mouvements du chirurgien aux instruments robotisés. Selon sa conception, un système audio intégré facilite la communication entre l'opérateur et l'assistant positionné au chevet du patient.

Le système de visualisation de l'image issue de la caméra endoscopique peut être positionné de manière à aligner le regard du chirurgien avec ses mains, optimisant ainsi la coordination main-œil (cf. figure 1.23). L'affichage peut nécessiter ou non le port de lunettes 3D en fonction de la technologie employée, qu'il s'agisse d'un écran stéréoscopique ou d'un casque de réalité virtuelle.

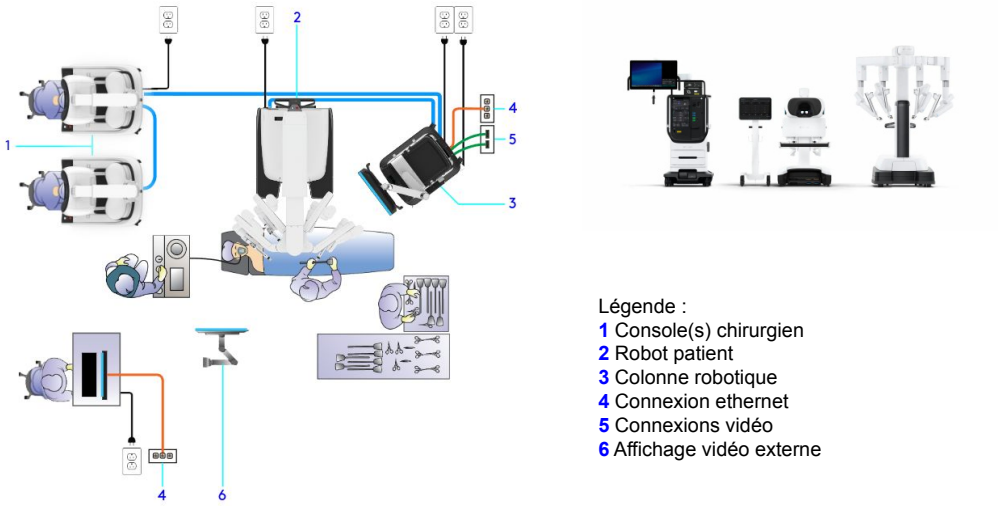


FIGURE 1.21 – Représentation schématique des éléments d'une plateforme de télé-chirurgie Da Vinci 5 (Intuitive Surgical) placée en salle d'opération autour du patient et interconnectés. Source : générée par l'auteur à partir de [Surgical 2025a].

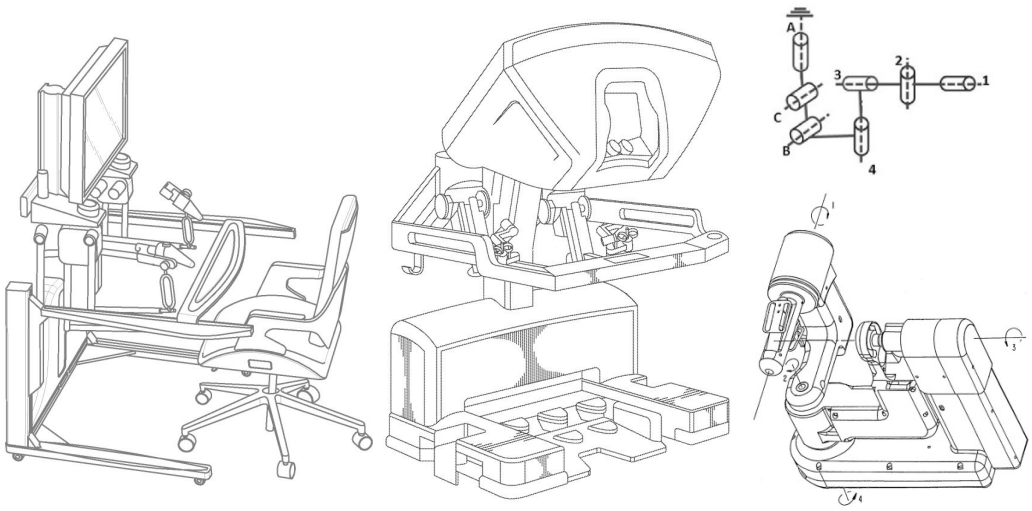


FIGURE 1.22 – Représentations schématiques des interfaces homme-machine : à gauche, une console ouverte (Versius, CMR Surgical) ; Au centre, une console immersive (Da Vinci Xi, Intuitive Surgical) ; À droite, une commande chirurgicale du Da Vinci Xi (Intuitive Surgical) accompagnée de son schéma cinématique. Source : générée par l'auteur à partir de [Marshall 2019, Liu 2024b, Niemeyer 1999, Freschi 2013].

L'IHM repose généralement sur deux interfaces de commande, chacune disposant de sept Degrés de Liberté (DDL) : trois en rotation, trois en translation ainsi qu'une fonction de préhension. Couplée à un dispositif de contrôle podal (pédales), elle permet la manipulation simultanée de deux instruments chirurgicaux ou de l'endoscope [Freschi 2013]. Par mesure de sécurité, toute sortie de la tête du chirurgien hors du champ de la console entraîne la désactivation des interfaces de commande [SCHWAB 2017].

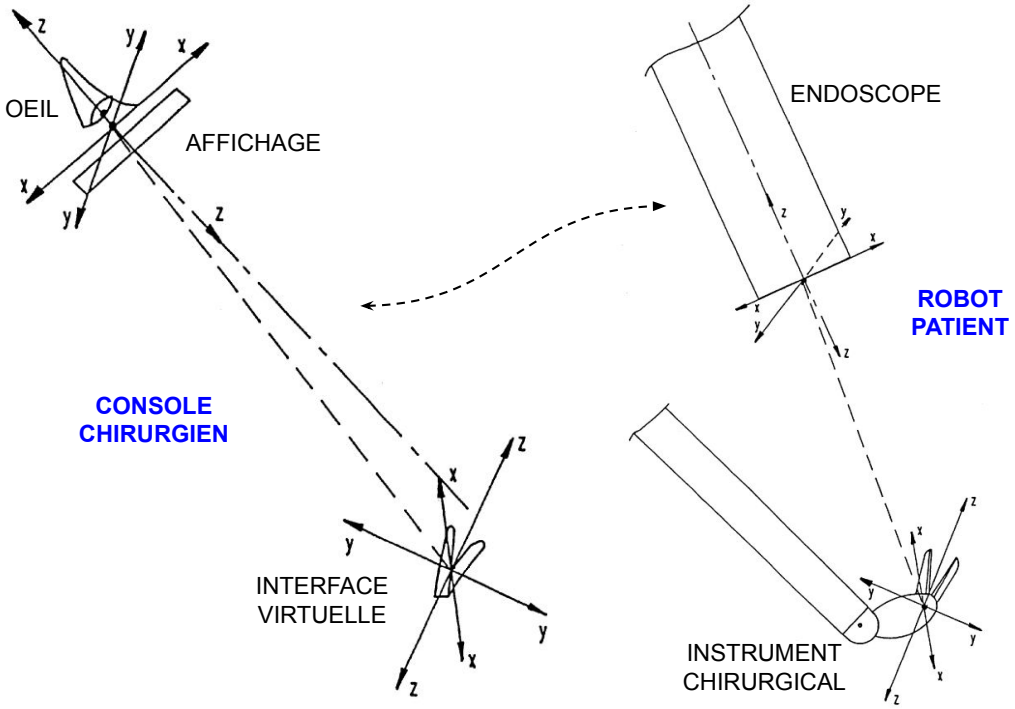


FIGURE 1.23 – Représentation schématique de la restauration de la coordination main-œil dans une interface homme-machine d'une console chirurgien. Source : générée par l'auteur à partir de [Niemeyer 1999].

L'Interface Utilisateur (IU), dont un exemple est illustré en figure 1.24, fournit des informations essentielles tels que l'état des instruments (présence, activation, mode d'électrochirurgie) et les paramètres de l'endoscope (orientation, luminosité, etc.). Un ou plusieurs écrans auxiliaires, tel que le système TilePro (Intuitive Surgical), peuvent enrichir l'affichage principal en intégrant des sources externes, comme des images échographiques.

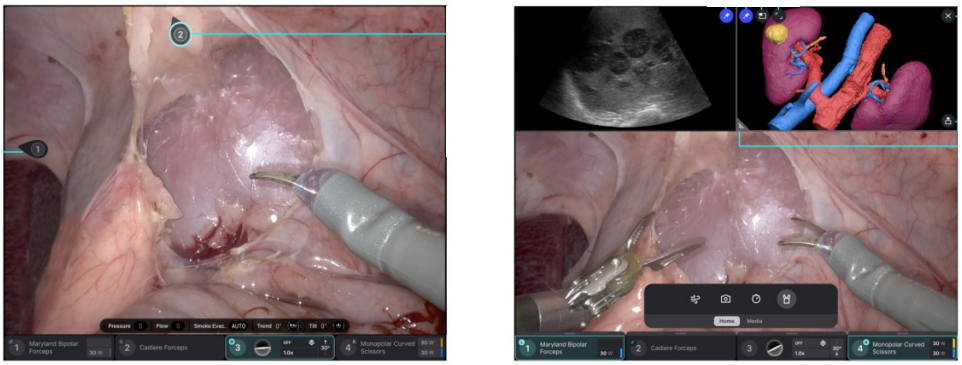


FIGURE 1.24 – Exemples d'interface utilisateur : (à gauche) interface standard ; (à droite) Interface avec affichages auxiliaires d'une image échographique et d'un rendu graphique de structures anatomiques. Source : générée par l'auteur à partir de [Surgical 2025a].

Dans le domaine de la TC abdominale, deux grandes approches se distinguent pour les robots patients (voir figure 1.25) : les robots monoblocs, qu'ils soient multi-ports ou à port unique, comme les systèmes développés par Intuitive Surgical et les robots distribués, tels que Versius (CMR Surgical) et Hugo (Medtronic). La flexibilité des unités de chevet ou Bed Side Unit (BSU) en Anglais offre une plus grande liberté dans le positionnement des trocarts [Halabi 2024].

Dans tous les cas, l'emplacement du robot et des trocarts impacte directement la dextérité du système et l'accessibilité du site chirurgical. Il est donc essentiel d'optimiser leur positionnement afin de minimiser les collisions entre les bras robotiques tout en garantissant un volume de travail intra abdominal suffisant.

Ce robot esclave peut être composé d'un ou plusieurs bras articulés. Certains de ces bras peuvent inclure des articulations passives, dont le rôle est de positionner la partie distale de la chaîne cinématique où se trouvent les articulations actives. Ces dernières peuvent être mobilisées manuellement ou commandées par le chirurgien via des interfaces maîtres.

La structure cinématique du bras robotique définit un Centre de Mouvement Distant (CMD) ou Remote Center of Motion (RCM) en Anglais [Freschi 2013]. Ce point d'appui, situé distalement par rapport à la structure du bras, correspond à la zone où l'instrument traverse la peau du patient. Le bras est conçu de manière à pivoter autour d'un axe de rotation aligné avec celui de la canule, comme illustré sur la figure 1.26. Cela permet de stabiliser le point d'insertion et de réduire les efforts de traction sur les tissus. Une configuration mécanique ou virtuelle en parallélogramme garantit que l'instrument se déplace le long d'un axe rigide [Schena 2016].

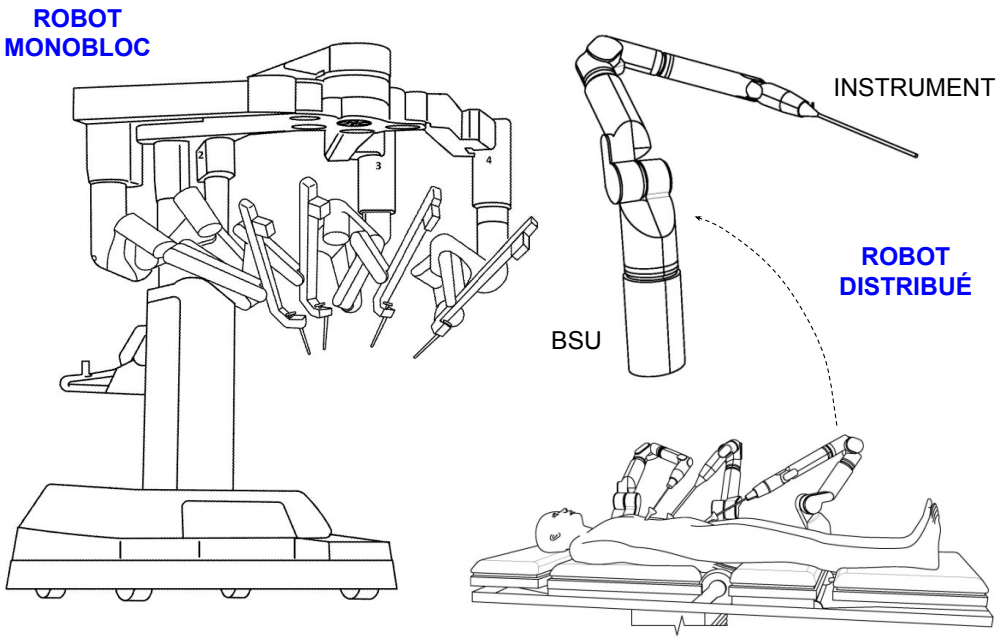


FIGURE 1.25 – Représentations schématiques des deux grandes approches pour les robots patient. (à gauche) Un robot patient monobloc (Da Vinci Xi, Intuitive Surgical) ; (à droite) Une unité de chevet d'un robot patient distribué (Versius, CMR Surgical). Source : générée par l'auteur à partir de [Shirazian 2019, Marshall 2019].

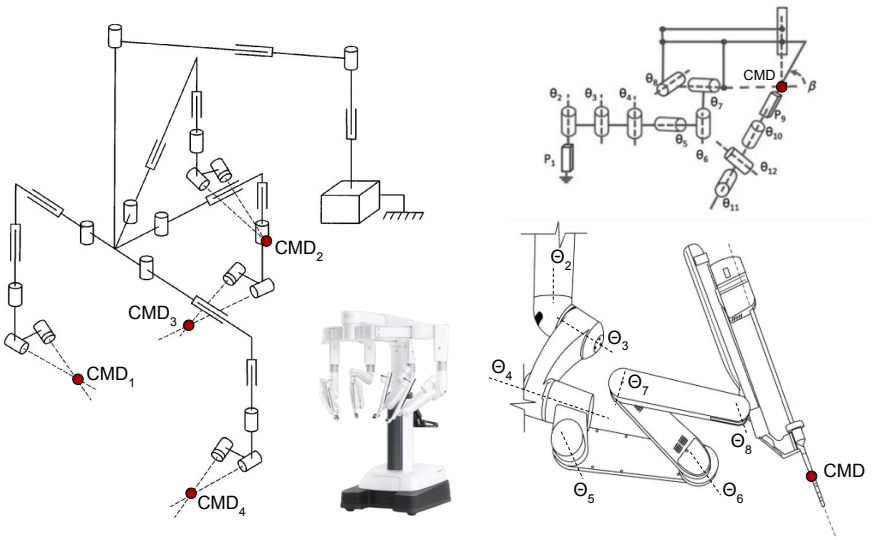


FIGURE 1.26 – Représentation schématique de la chaîne cinématique d'un robot patient monobloc (Da Vinci, Intuitive Surgical). Source : générée par l'auteur à partir de [Schena 2016, Shirazian 2019, Freschi 2013] et Intuitive Surgical.

La chaîne de commande entre maître et esclave, illustrée en figure 1.27, a pour but premier d'assurer la réplication précise des mouvements des interfaces maîtres vers les bras et instruments esclaves. Afin d'assurer un contrôle rigoureux des mouvements, un asservissement en boucle fermée est utilisé [Schena 2016, Prisco 2010].

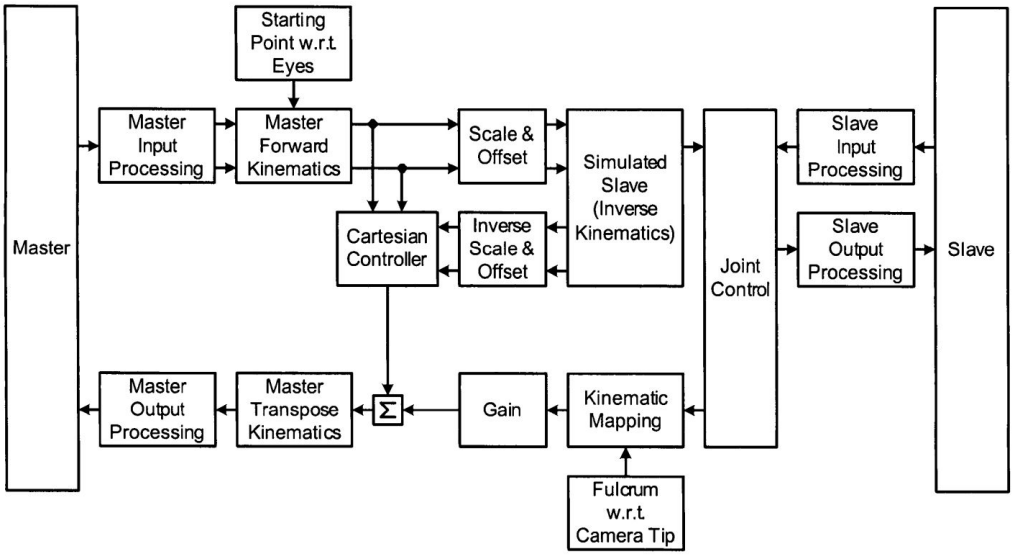


FIGURE 1.27 – Illustration simplifiée d'un schéma fonctionnel de commande pour un système de télé-chirurgie maître/esclave avec un asservissement en boucle fermée. Source : générée par l'auteur à partir de [Prisco 2010].

Dans la phase d'action, les mouvements des interfaces maître sont traduits en coordonnées cartésiennes via la cinématique directe et adaptés par un module de mise à l'échelle et décalage pour correspondre aux contraintes opérationnelles. Par exemple, si le chirurgien déplace l'interface de la pince virtuelle de un centimètre vers la gauche à l'écran, avec une échelle de mouvement 1 : 1, le point d'appui de l'instrument chirurgical se déplace également de un centimètre vers la gauche par rapport au repère de la caméra. Les translations sont mappées de manière relative pour maintenir une position neutre pour les bras du chirurgien, tandis que les rotations sont mappées en absolu. Les coordonnées transformées sont ensuite converties en commandes articulaires par la cinématique inverse du système esclave [Prisco 2010, Niemeyer 2005].

Dans la phase de rétroaction, les données issues des capteurs du système esclave sont traitées pour être retranscrites en informations cartésiennes et transmises au maître via la cinématique transposée. Cela permet une rétroaction tactile ou visuelle en temps réel, essentielle pour maintenir la précision du système [Prisco 2010, Niemeyer 2005].

Un filtre actif est intégré dans le chemin de commande de l'instrument pour atténuer les vibrations indésirables, en bloquant les fréquences au-delà d'une limite de coupure typique (environ 3 Hz), réduisant ainsi les tremblements naturels de la main de l'homme.

Les systèmes robotiques peuvent interagir avec les équipements du bloc opératoire, comme la table opératoire. Cela permet de mobiliser de manière synchrone la table et les bras du robot afin de repositionner le patient sans interrompre l'intervention [Itkowitz 2015].

Tout comme la chaîne de commande, la chaîne d'acquisition et de traitement numérique vidéo constitue un module central du système de TC, comme illustré en figure 1.28. La source du flux vidéo est généralement un endoscope stéréoscopique. Il est soit équipé de deux capteurs ou d'un unique capteur. Cette dernière configuration, compacte et performante, repose sur deux chemins optiques distincts (un pour chaque œil) guidés par des trains de lentilles et des prismes. Ces éléments projettent les images sur des zones spécifiques d'un capteur d'image continu, réduisant ainsi les erreurs d'alignement, éliminant le besoin de synchroniser deux capteurs et améliorant la qualité des images stéréoscopiques [Patwardhan 2017, Shafer 2006].

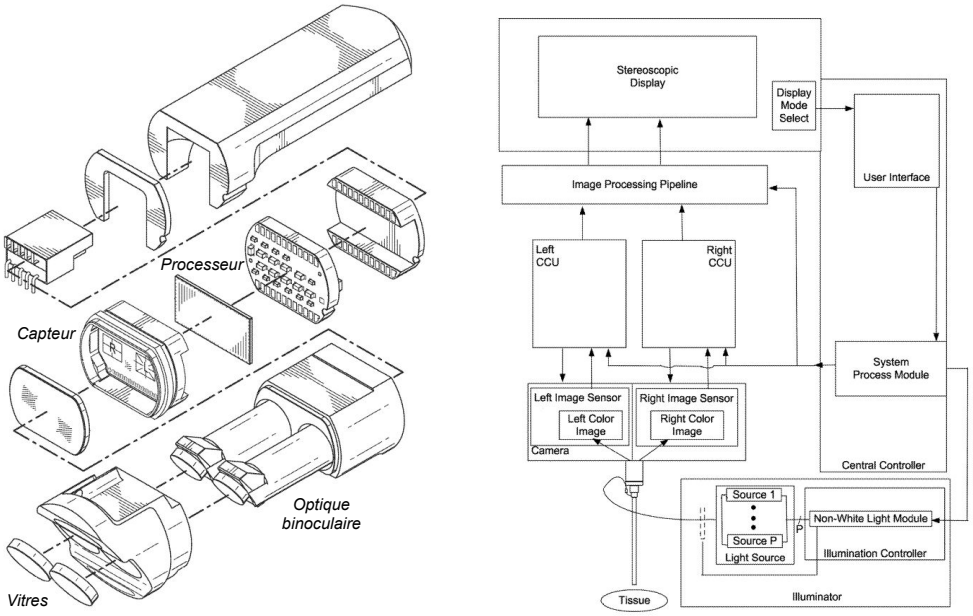


FIGURE 1.28 – (à gauche) Représentation schématique de la partie distale d'un endoscope stéréoscopique compact (Da Vinci SP, Intuitive Surgical) ; (à droite) Représentation schématique de la chaîne de traitement numérique du flux vidéo endoscopique. Source : générée par l'auteur à partir de [Patwardhan 2017, Dicarlo 2015].

Pour minimiser le chevauchement des images (cross-talk en Anglais), des séparateurs de champ physiques ou des revêtements non réfléchissants sont utilisés. Des filtres infrarouges et de balance des couleurs sont intégrés pour ajuster la qualité d'image selon les exigences chirurgicales. Le design compact privilégie un chemin optique court avec un capteur intégré dans la partie distale de l'endoscope, réduisant l'atténuation lumineuse et les distorsions optiques, mais impose des contraintes accrues sur la sensibilité et le traitement du signal [Shafer 2006].

La gestion de l'intensité lumineuse et de la colorimétrie s'effectue via un système en boucle fermée, qui adapte dynamiquement l'exposition et la couleur d'éclairage selon les structures anatomiques observées et les données de profondeur. Ce mécanisme garantit un éclairage homogène et optimal pour chaque scène opératoire [He 2022c, Dicarlo 2015, He 2022b, Richmond 2019, Levesque 2019]. Des techniques de traitement numérique, telles que l'adaptation automatique de l'exposition, la réduction des réflexions spéculaires et l'imagerie multispectrale par fluorescence du vert d'indocyanine, contribuent à améliorer le contraste et à révéler avec précision les structures anatomiques d'intérêt [McDowall 2017, Azizian 2015].

Le système de traitement d'image intègre également une correction numérique des distorsions optiques et de la disparité native des images stéréoscopiques. Pour cela, les paramètres d'un modèle géométrique de caméra sont déterminés à l'aide de cibles externes, souvent tridimensionnelles [Watson 2024, Morrissette 2015, Zhao 2023a, Mori 2020].

Enfin, une gestion thermique efficace, utilisant des conduits thermiques ou des sources de chaleur, prévient la condensation sur la vitre de l'endoscope. Cette solution assure une clarté visuelle constante [Deyanov 2020, Halderman 2023].

Sur le plan des accessoires, les systèmes de TC abdominale mettent en œuvre des instruments aux pointes articulées afin de surmonter les problèmes d'amplitude de mouvements et de dextérité couramment rencontrés en cœlioscopie. Les mouvements de l'extrémité distale de l'instrument sont générés par un système de câbles et d'enrouleurs ou des tringles [Rabindran 2024, Brisson 2023, Ortiz 2023, Gawaikar 2023], comme illustré en figure 1.29. Ces instruments permettent un mouvement dans au moins trois DDL et sont adaptés à des manipulations complexes comme saisir, couper et manipuler les tissus.

Le retour haptique est un aspect clé des nouveaux instruments de chirurgie robotique. Plusieurs approches sont utilisées pour fournir ce retour, notamment des capteurs placés au niveau distal et/ou proximal pour mesurer des forces, des déformations ou des déplacements [Ortiz 2023, Suresh 2023]. Ces capteurs doivent être conçus pour résister aux interférences électromagnétiques, comme celles générées par l'électrochirurgie, qui peuvent affecter les mesures lorsque le champ magnétique interagit avec le corps de l'instrument [Ortiz 2023]. Restaurer les sensations de toucher et de force chez le chirurgien garantit la sécurité du patient en évitant l'application de couples et de forces excessifs sur les tissus [Servais 2024].

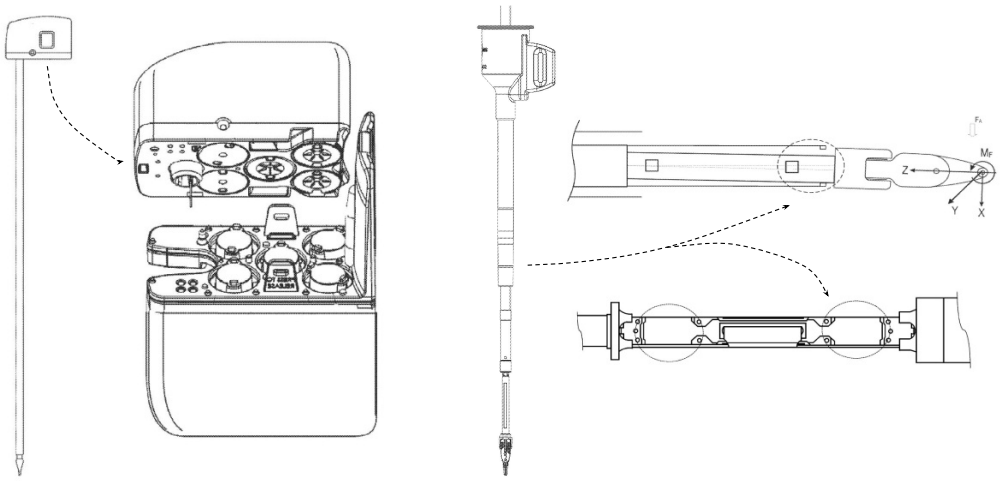


FIGURE 1.29 – (à gauche) Schéma d'un instrument robotique articulé à l'aide d'un système de câbles et d'enrouleurs pour contrôler les mouvements de l'extrémité distale ; (à droite) Schéma d'un instrument robotique intégrant une jauge de contrainte pour fournir un retour haptique. Source : générée par l'auteur à partir de [Brisson 2023, Suresh 2023].

1.2.4.4 Chirurgie guidée par l'image

En parallèle de l'évolution des techniques chirurgicales, allant de la CO aux approches mini-invasives classiques et robotiques, les outils d'assistance se sont considérablement développés, portés par les progrès en imagerie et en informatique. La Chirurgie Guidée par l'Image (CGI), sous-domaine de la Chirurgie Assistée par Ordinateur (CAO), désigne toute procédure au cours de laquelle le chirurgien s'appuie sur des instruments suivis et guidés en temps réel, en conjonction avec des images préopératoires ou peropératoires, afin d'augmenter la précision et la sécurité du geste chirurgical [Bucholz 1995, Nasab 2019, Azagury 2015, Lin 2023d].

La CGI repose sur l'intégration de capteurs, notamment des caméras optiques et des détecteurs de champs électromagnétiques, permettant la numérisation de l'anatomie du patient et la localisation dynamique de l'instrumentation chirurgicale. Cette technologie combine des méthodes de planification préopératoire avec des outils de guidage peropératoires, dont certains sont détaillés en section 1.3 [Azagury 2015, Lin 2023d]. Les informations de guidage peropératoire sont restituées au chirurgien via des interfaces variées, allant des écrans classiques aux dispositifs de réalité augmentée, voire à des systèmes de projection directe sur le patient [Azagury 2015, Lin 2023d].

L'adoption de la CGI en pratique clinique présente de nombreux avantages, parmi lesquels une meilleure maîtrise peropératoire, un retour d'information en temps réel sur l'effet du geste chirurgical, une réduction des traumatismes tissulaires et une minimisation des perturbations de l'accès aux structures anatomiques profondes.

En neurochirurgie, la CGI a initialement émergé dans le cadre du traitement des tumeurs cérébrales par chirurgie stéréotaxique et radiochirurgie [Kelly 2000]. Son utilisation permet de réduire significativement le risque de déficits neurologiques postopératoires et les complications inhérentes aux procédures ablatives. Elle a notamment rendu possible l'ablation de tumeurs autrefois considérées comme inopérables en raison de leur localisation critique ou de leur volume [Mezger 2013]. Aujourd'hui des dispositifs de navigation chirurgicale comme illustré en figure 1.30 sont utilisés.



FIGURE 1.30 – (à gauche) Salle hybride de chirurgie guidée par l'image pour le traitement par cryothérapie de l'endométriose. (à droite) Système de navigation chirurgicale utilisé en neurochirurgie et chirurgie du rachis. Sources : Centre Hospitalier Universitaire (CHU) Clermont-Ferrand et Brainlab.

En chirurgie abdominale, la CGI joue un rôle majeur dans la localisation des lésions hépatiques, pancréatiques et rénales en optimisant la navigation tumorale et en réduisant les marges d'exérèse. De même, en chirurgie orthopédique, elle facilite la pose d'implants et garantit un alignement précis des structures osseuses grâce aux reconstructions tridimensionnelles issues de l'imagerie préopératoire [Azagury 2015].

L'essor des salles d'opération hybrides, intégrant des dispositifs d'imagerie peropératoire comme l'imagerie par résonance magnétique et la tomodensitométrie, illustré en figure 1.30, a renforcé l'efficacité de la CGI en fournissant une acquisition en temps réel des données anatomiques et fonctionnelles. Cette avancée contribue à compenser les imprécisions induites par les déplacements tissulaires durant l'intervention [Lin 2023d].

La CGI constitue une avancée majeure pour de multiples spécialités chirurgicales dans la prise en charge moderne des patients, garantissant une précision et une sécurisation accrue des actes invasifs. Son développement, en synergie avec les progrès en imagerie médicale, en intelligence artificielle et en robotique, ouvre la voie à une médecine de plus en plus personnalisée, où chaque intervention est optimisée en fonction de la morphologie et de la physiopathologie spécifique du patient.

1.3 Contexte technique

En complément des outils standards, les avancées technologiques récentes se sont intégrées aux différentes étapes du parcours chirurgical des patients, optimisant ainsi leur prise en charge et contribuant à une amélioration des résultats cliniques.

1.3.1 Outils de planification préopératoire

La planification préopératoire constitue une étape essentielle de la prise en charge chirurgicale. Elle remplit une double fonction : d'une part, elle optimise la préparation cognitive du chirurgien en favorisant une modélisation mentale précise de l'acte opératoire ; d'autre part, elle améliore l'éducation thérapeutique du patient en facilitant sa compréhension des enjeux et du déroulement de l'intervention.

1.3.1.1 Images tridimensionnelles en coupe

Dans le domaine de la radiologie médicale, les images tridimensionnelles en coupe sont généralement de deux natures : la TomoDensitoMétrie (TDM) et l'Imagerie par Résonance Magnétique nucléaire (IRM). L'utilisation d'une de ces modalités est souvent incontournable en planification préopératoire.

La TDM repose sur l'acquisition et l'interprétation des coefficients d'atténuation du rayonnement électromagnétique ionisant X traversant les tissus biologiques. Leur conversion en Unités Hounsfield (UH), permet une quantification précise des densités tissulaires. L'évolution technologique des scanners a conduit à des performances accrues en termes de rapidité, de résolution spatiale et de gestion de la dose d'irradiation [Cunningham 2000, Bricault 2017, Jung 2021, Pontana 2025].

L'architecture des scanners modernes, illustrée en figure 1.31, repose sur une structure annulaire intégrant un ou plusieurs systèmes d'acquisition en rotation continue, couplés à une table motorisée assurant la translation du patient à travers le champ d'irradiation. Chaque système d'acquisition se compose d'un tube émetteur qui génère un faisceau X collimaté en éventail, dont l'atténuation après traversée des tissus est enregistrée par un réseau de détecteurs. L'adoption de détecteurs multicoupes a remplacé les systèmes mono-coupes, permettant l'acquisition simultanée de plusieurs sections anatomiques au cours d'une même rotation [Cunningham 2000, Jung 2021, Pontana 2025].

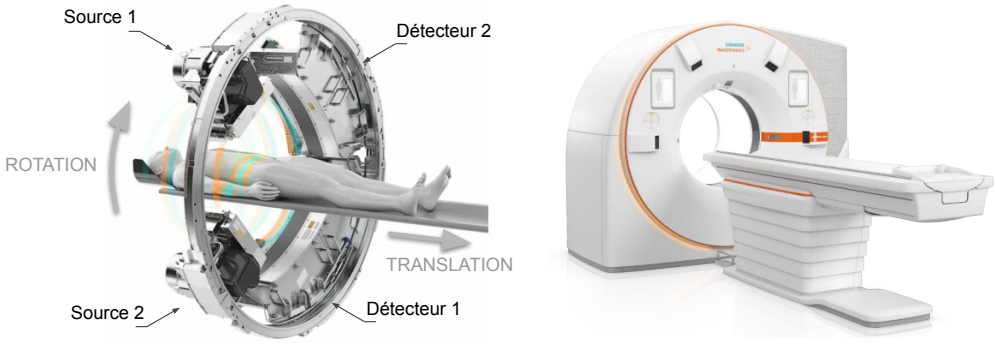


FIGURE 1.31 – (à gauche) Illustration des principaux composants d'un scanner bi-tube ; (à droite) Photographie d'un scanner à comptage photonique (NAEOTOM Alpha, Siemens Healthineers). Source : générée par l'auteur à partir d'illustrations Siemens Healthineers.

L'acquisition hélicoïdale constitue une avancée importante en garantissant un balayage continu du volume anatomique grâce à la synchronisation du déplacement du tube émetteur et de la table motorisée. Le paramètre de pitch, défini comme le rapport entre le déplacement de table et la collimation, influence la rapidité et la qualité d'acquisition. Une optimisation fine des paramètres d'acquisition permet une réduction des artefacts de mouvements et une meilleure synchronisation avec l'administration des agents de contraste radio-opaques [Bricault 2017].

La reconstruction tomodensitométrique repose sur la résolution d'un système d'équations reliant les mesures d'atténuation aux positions angulaires du faisceau X, organisées sous forme de sinogramme. Chaque coupe du volume reconstruit, représentée par une matrice de 512×512, soit 262 144 voxels, nécessite un algorithme de reconstruction performant compatible avec les contraintes cliniques.

Historiquement privilégiée pour sa rapidité, la rétroprojection filtrée (Filtered Back Projection (FBP) en Anglais) présente un bruit d'image élevé et des limites en résolution spatiale. L'émergence des algorithmes itératifs a permis une réduction significative du bruit et de la dose d'irradiation en ajustant les coefficients d'atténuation par comparaison entre projections simulées et données expérimentales. Aujourd'hui, l'intégration des méthodes d'apprentissage profond révolutionne la reconstruction des images, optimisant la qualité des acquisitions en faible dose, la reconstitution des données sous-échantillonnées et la réduction des artefacts métalliques [Pontana 2025].

L'amélioration de la résolution des détecteurs permet d'effectuer des reconstructions multiplanaires (MultiPlanar Reconstruction (MPR) en Anglais) selon n'importe quel axe de l'espace. Pour afficher les images reconstruites sur écran, un fenêtrage est appliqué afin d'adapter les niveaux de gris et d'optimiser le contraste des structures examinées. Toutefois, comme un écran standard ne peut afficher que 256 nuances de gris, une table de correspondance (Look-Up Table (LUT) en Anglais)

convertit les valeurs brutes en intensités visibles, proportionnelles à la densité des tissus [Pontana 2025].

L'utilisation des rayonnements ionisants est soumise au principe As Low As Reasonably Achievable (ALARA) en Anglais, garantissant l'exposition minimale nécessaire pour répondre à la question clinique posée. L'optimisation passe par la limitation du champ exploré, la réduction du nombre de séquences après injection de contraste et l'adaptation des paramètres d'acquisition. L'exposition aux rayons X, bien que contrôlée, peut induire des effets biologiques délétères, tels que des altérations de l'Acide DésoxyriboNucléique (ADN), un risque accru de pathologies radio-induites et des lésions tissulaires en cas d'irradiation excessive [Pontana 2025].

L'IRM repose sur l'exploitation des propriétés magnétiques des noyaux d'hydrogène omniprésents dans les tissus biologiques en raison de leur abondance dans l'eau et les lipides. Chaque proton d'hydrogène possède un moment magnétique intrinsèque, lié à son spin quantique, qui interagit avec un champ magnétique externe intense, généralement compris entre 1,5 et 3 Tesla en imagerie clinique. Ce phénomène, fondé sur la Résonance Magnétique Nucléaire (RMN), repose sur la quantification des niveaux d'énergie associés à l'orientation des moments magnétiques dans un champ statique.

Le dispositif d'IRM, illustré en figure 1.32, est constitué d'un aimant supraconducteur générant un champ magnétique principal, orienté selon l'axe longitudinal du tunnel dans lequel le patient est placé en position allongée. Sous l'influence de ce champ, les moments magnétiques des protons s'orientent préférentiellement selon deux états énergétiques distincts : parallèle ou antiparallèle. Cette répartition génère une aimantation macroscopique, qui constitue l'état d'équilibre magnétique du système [Tourdias 2025].

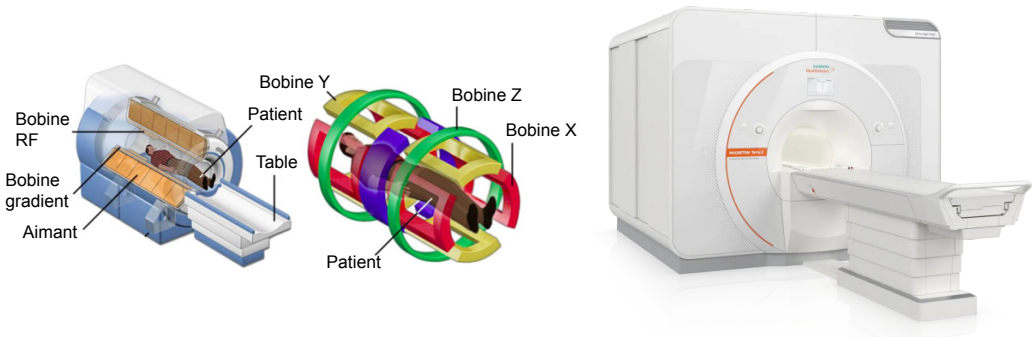


FIGURE 1.32 – (à gauche) Illustration des principaux composants d'une IRM ; (à droite) Photographie d'une IRM 7T (MAGNETOM Terra.X, Siemens Healthineers). Source : générée par l'auteur à partir de [Kurzhunov 2017] et d'illustrations Siemens Healthineers.

L'application d'une onde électromagnétique dans le domaine des RadioFréquences (RF), accordée à la fréquence de Larmor, induit une transition entre ces états énergétiques, provoquant une bascule de l'aimantation macroscopique vers le plan transverse. Cette excitation, réalisée par une impulsion RF, entraîne une mise en résonance des protons. Après l'arrêt de cette excitation, les protons retournent à leur état d'équilibre par deux mécanismes distincts : la relaxation longitudinale (T1) et la relaxation transversale (T2). Le signal RMN issu de cette précession est détecté grâce à l'induction électromagnétique dans les antennes de réception [Tourdias 2025, Berger 2002, Plewes 2012].

L'IRM ne se limite pas à la détection d'un signal global, mais requiert une localisation spatiale fine du signal. Cette spatialisation est obtenue par l'application de gradients de champ magnétique selon les trois axes de l'espace. Ces gradients modulent linéairement la fréquence de Larmor en fonction de la position, introduisant ainsi une relation fréquence de résonance et coordonnée spatiale. La sélection d'un plan de coupe spécifique est obtenue en combinant une impulsion RF et un gradient de champ magnétique, permettant d'exciter uniquement les spins situés dans une tranche donnée. Le codage spatial du signal au sein de cette coupe est ensuite assuré par l'application successive de gradients de lecture et de phase, permettant d'attribuer une position précise à chaque signal détecté [Tourdias 2025, Berger 2002, Plewes 2012].

L'IRM présente des avantages notables en imagerie médicale. Son principal atout réside dans l'absence de radiation ionisante, ce qui la distingue des techniques utilisant les rayons X comme la TDM. Elle offre également une excellente résolution en contraste, particulièrement en imagerie des tissus mous, ce qui permet une détection fine de nombreuses pathologies notamment neurologiques, hépatiques et musculo-squelettiques. L'ajustement des paramètres d'acquisition, comme le Temps d'Écho (TE) et le Temps de Répétition (TR), permet de moduler la pondération de l'image en fonction du temps de relaxation des tissus, optimisant ainsi la visualisation des structures anatomiques et des anomalies sous-jacentes [Tourdias 2025].

Toutefois, cette modalité d'imagerie présente certaines limitations. L'intensité du champ magnétique impose des contraintes strictes de compatibilité pour les dispositifs implantés, rendant l'examen contre-indiqué chez certains patients porteurs de prothèses ferromagnétiques ou de pacemakers non compatibles. Par ailleurs, la durée d'acquisition, souvent plus longue que celle d'un scanner, peut générer une gêne pour les patients claustrophobes ou ayant des difficultés à maintenir l'immobilité. Enfin, l'IRM demeure une technologie onéreuse, nécessitant des infrastructures adaptées, ainsi qu'une maintenance rigoureuse du système cryogénique assurant la supraconductivité de l'aimant principal [Tourdias 2025, Berger 2002].

1.3.1.2 Reconstruction tridimensionnelle à partir de l'imagerie en coupe

Les reconstructions tridimensionnelles volumiques et surfaciques issues de l'imagerie en coupe, illustrées en figure 1.33, fournissent une représentation détaillée de l'anatomie du patient. Ce type de modélisation est souvent désigné dans le domaine clinique sous le terme de jumeau numérique tridimensionnel [Sun 2023]. La précision du modèle reconstruit dépend étroitement de la qualité des images initiales. Sa génération repose sur une chaîne de traitement, allant de l'acquisition des données à la visualisation finale.

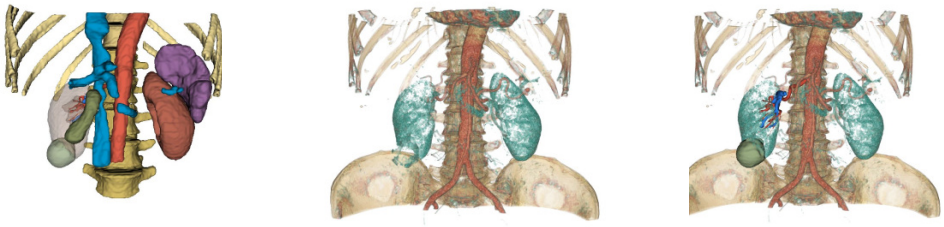


FIGURE 1.33 – (à gauche) Reconstructions tridimensionnelles obtenues à partir d'un uroscanner chez un patient présentant une tumeur rénale du pôle inférieur droit ; (à gauche) Mode surfacique ; (au centre) Mode volumique ; (à droite) Mode mixte. Source : auteur.

Dans le cas d'un examen multiphasique en TDM ou IRM, un recalage intra-modalité peut être nécessaire pour aligner spatialement les séquences d'images acquises à différents temps. Diverses techniques de recalage sont décrites dans la littérature scientifique. Certaines sont intégrées dans des bibliothèques logicielles comme elastix [Klein 2009] ou des solutions dédiées telles que 3D Slicer [Fedorov 2012] et Synapse 3D (Fujifilm) [Kojo 2024]. Ces méthodes se répartissent en deux grandes catégories : rigides et déformables, le choix étant conditionné par l'organe cible et l'application clinique. Le processus de recalage intra-modalité consiste a résoudre un problème d'optimisation qui permet d'estimer une transformation, le plus souvent affine, qui aligne une image sur une autre. Cette optimisation s'appuie sur des fonctions de coût qui exploitent des critères explicites tels que l'intensité des voxels et la forme des structures anatomiques [Flores-Mangas 2012, Heryan 2016]. Par ailleurs, des approches implicites utilisant l'apprentissage profond sont également développées [Chen 2024a].

Une fois les données alignées, la reconstruction volumique est obtenue par application d'une fonction de transfert qui associe aux valeurs de gris de chaque voxel des attributs visuels tels que la couleur et la transparence. Le rendu volumique repose principalement sur des algorithmes de ray-casting [Smelyanskiy 2009]. L'extension vers le rendu cinématique s'appuie sur des méthodes de path-tracing et de modélisation de l'illumination globale, améliorant le réalisme des représentations tridimensionnelles [Xu 2023, Dappa 2016, Eid 2017].

La reconstruction surfacique nécessite quant à elle une segmentation précise afin d'isoler et de caractériser les structures anatomiques d'intérêt. Des algorithmes de segmentation tels que graph-cuts sont couramment utilisés. A cela s'ajoutent des techniques basées sur des atlas anatomiques [Yang 2014] ou des réseaux de neurones profonds automatisant l'extraction des structures avec une précision surpassant les méthodes conventionnelles [Gibson 2018]. Un maillage tridimensionnel est alors généré à partir des données segmentées pour être ensuite affiché.

1.3.1.3 Fabrication additive

La fabrication additive s'impose aujourd'hui comme une technologie incontournable dans le domaine médical, facilitant la transition du monde virtuel au monde réel. Elle permet la production de structures anatomiques complexes et constitue une alternative aux méthodes conventionnelles, parfois inadaptées aux exigences de la fabrication médicale. Son application en santé repose sur la transformation des images en coupes en modèles tridimensionnels surfaciques.

En phase préopératoire, l'utilisation de cette technologie se décline autour de trois axes principaux : la planification chirurgicale, l'entraînement des chirurgiens et l'éducation thérapeutique du patient. Grâce à la précision des modèles tridimensionnels ainsi fabriqués, il devient possible d'améliorer la compréhension des relations anatomiques entre les structures, d'optimiser les stratégies interventionnelles et de réduire le risque peropératoire [Meyer-Szary 2022, Moldovanu 2024]. Elle joue un rôle essentiel dans la simulation de procédures rares ou techniquement complexes, permettant une personnalisation accrue de la prise en charge et une anticipation des défis chirurgicaux spécifiques à chaque patient [Moldovanu 2024].

En urologie, la fabrication additive permet de produire des modèles anatomiques aux propriétés métrologiques satisfaisantes, comparables à celles de l'imagerie source, notamment lorsqu'ils sont réalisés par impression PolyJet. Cette technologie repose sur la projection de photopolymères liquides en couches ultra-fines à l'aide de têtes d'impression à jet d'encre, suivi d'un durcissement immédiat par exposition à la lumière ultraviolette. Elle garantit une précision dimensionnelle élevée et permet de reproduire des détails anatomiques complexes avec une fidélité remarquable. De plus, l'impression PolyJet offre la possibilité d'attribuer des propriétés visuelles spécifiques aux structures modélisées (couleurs, opacité), tout en conservant une incertitude dimensionnelle de fabrication inférieure à 3 % par rapport à l'anatomie réelle [Michiels 2019].

L'entraînement chirurgical à l'aide de modèles physiques obtenus par fabrication additive a démontré une diminution significative des pertes sanguines, en particulier lors des Néphrectomies Partielles Robot-Assistées (NPRA) [Von Rundstedt 2017, Maddox 2018]. Sur le plan pédagogique, ces modèles favorisent une meilleure visualisation des structures anatomiques et facilitent l'apprentissage des techniques chirurgicales complexes [Moldovanu 2024].

Concernant l'éducation thérapeutique, l'impact des modèles tridimensionnels sur la compréhension du patient a été largement documenté. Les pratiques et outils de communication soignant-soigné tendent à évoluer comme le montre la figure 1.34. L'utilisation de modèles personnalisés a permis une amélioration de 16,7 % de la compréhension de la physiologie rénale, de 50 % de l'anatomie rénale, de 39,3 % des caractéristiques tumorales et de 44,6 % de la procédure chirurgicale envisagée [Bernhard 2016, Wake 2019, De Hauteclocque 2020, Michiels 2021a, Moldovanu 2024]. Ces bénéfices se traduisent par une meilleure acceptation de l'intervention et une réduction de l'anxiété préopératoire, favorisant ainsi une prise de décision plus éclairée par le patient.

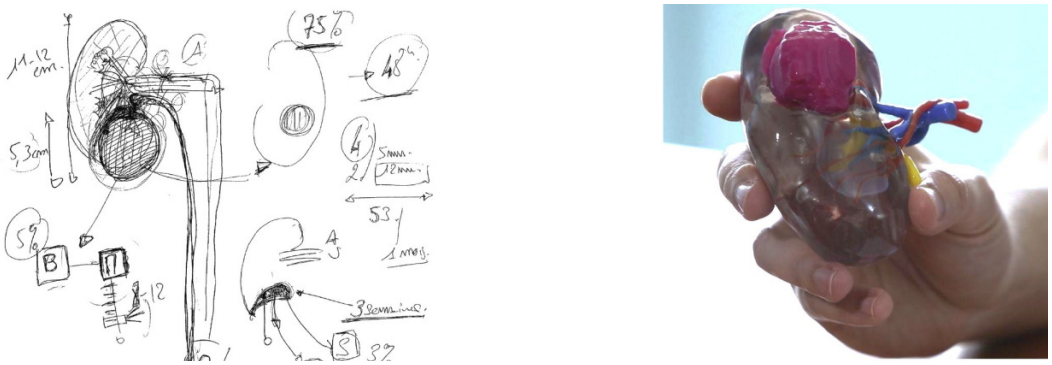


FIGURE 1.34 – Évolution des outils de la communication soignant-soigné avant néphrectomie partielle. (à gauche) Schéma explicatif réalisé classiquement par le chirurgien. (à droite) Modèle physique du patient obtenu par fabrication additive PolyJet. Source : [Michiels 2020].

La fabrication additive s'impose ainsi comme un outil incontournable, tant pour les chirurgiens que pour les patients. En intégrant cette technologie aux pratiques cliniques, il devient possible d'optimiser la phase préopératoire, d'améliorer la formation des praticiens et de renforcer l'interaction médecin-patient

1.3.2 Outils de guidage peropératoire

Le geste chirurgical en peropératoire peut également être assisté par divers outils technologiques.

1.3.2.1 Échographie peropératoire

L'Échographie Peropératoire (EP) est un outil conçu pour assister le chirurgien en temps réel dans son geste. En CMI, elle contribue à l'identification précise des structures, à l'évaluation de leurs positions relatives et de leurs interrelations, garantissant ainsi une meilleure sécurisation de l'acte chirurgical tout en minimisant les risques de complications peropératoires. Comme pour l'échographie diagnostique, le système comporte deux composants, illustrés en figure 1.35 : l'échographe qui

se présente généralement sous la forme d'un dispositif mobile équipé d'un écran et la sonde qui permet l'émission et la réception des ultrasons via un transducteur piézoélectrique convertissant une impulsion électrique en une onde acoustique et inversement [Patat 2025, Makuuchi 1998].



FIGURE 1.35 – (à gauche) Photographie de sondes échographiques L43K et L51K (Fujifilm Healthcare) utilisables en chirurgie robotique. (au centre) Photographie d'un échographe ARIETTA Precision (Fujifilm Healthcare). (à droite) Image échographique visualisée dans une console Da Vinci (Intuitive Surgical) lors d'une néphrectomie partielle robot-assistée. Sources : Fujifilm Healthcare et Centre Hospitalier Universitaire (CHU) de Bordeaux.

Lors d'une intervention en CMI pour l'exérèse d'une tumeur, une sonde stérile est insérée dans la cavité par un trocart. Manipulée par des instruments classiques ou robotisés, elle permet de délimiter avec précision les limites tumorales, d'identifier le plan de dissection et de localiser les structures sensibles avoisinantes, notamment les vaisseaux. La technologie repose principalement sur la technique du pulse-écho, dans laquelle des ultrasons de haute fréquence, généralement compris entre 2 et 15 MHz, sont émis à travers les tissus et réfléchis en fonction de leurs propriétés acoustiques, générant ainsi des images en niveaux de gris. Ces images sont obtenues grâce à l'analyse du temps aller-retour des ondes et de la variation d'amplitude des échos captés. En complément, la technique Doppler permet d'analyser la dynamique du flux sanguin en mesurant le décalage de fréquence entre l'onde ultrasonore incidente et les échos réfléchis par les éléments en mouvement, notamment les globules rouges. Cette approche permet d'évaluer la vitesse et la direction du flux sanguin, un paramètre clé pour la cartographie vasculaire peropératoire [Patat 2025, Uppal 2010, Marcal 2013].

Bien que l'EP soit aujourd'hui considérée comme le gold standard en CMI, elle présente certaines limitations. La qualité de visualisation est fortement dépendante des propriétés acoustiques des tissus traversés. En effet, l'échographie repose sur la réflexion des ondes ultrasonores aux interfaces des tissus ayant des impédances acoustiques différentes. Ainsi, les milieux osseux et gazeux constituent des barrières infranchissables pour les ultrasons, expliquant l'impossibilité d'obtenir des images à travers les structures osseuses ou les zones remplies d'air, comme les intestins. L'application d'un gel de couplage peut être nécessaire pour éliminer les artéfacts liés à l'air entre la sonde et la surface des tissus. En outre, l'EP est sujette à des artéfacts acoustiques, notamment les cônes d'ombre et le renforcement postérieur, qui peuvent altérer sa précision [Patat 2025].

De plus, la technique est fortement opérateur-dépendante, nécessitant un long apprentissage et une expertise avancée pour optimiser l'acquisition et l'interprétation des images. Enfin, bien que l'échographie offre une excellente résolution spatiale, elle demeure une modalité d'imagerie à champ restreint, contrairement aux techniques en coupe comme l'IRM ou la TDM, qui permettent une visualisation volumique plus complète.

Malgré ces limitations, l'EP continue d'évoluer grâce aux progrès technologiques récents. Le développement de sondes matricielles à haute fréquence et de modes avancés d'imagerie ultrasonore, tels que l'élastographie ou l'échographie de contraste, ouvre de nouvelles perspectives pour améliorer la précision des diagnostics peropératoires et affiner les stratégies chirurgicales en temps réel.

1.3.2.2 Imagerie par fluorescence

L'imagerie par fluorescence dans le proche infrarouge ou Near-InfraRed Fluorescence imaging (NIRF) en Anglais constitue un outil essentiel et complémentaire pour le guidage peropératoire. Cet outil repose sur l'administration de vert d'indocyanine ou IndoCyanine Green (ICG) en Anglais, un colorant hydrosoluble de la famille des cyanines, largement utilisé en imagerie médicale en raison de ses propriétés d'absorption et d'émission dans le spectre du proche infrarouge (∼800 nm). Initialement développée pour l'évaluation du débit cardiaque et de la perfusion sanguine [Nagaya 2017, Zheng 2019], son application en CMI cœlioscopique et robot-assistée s'est révélée particulièrement précieuse pour l'amélioration du contraste anatomique peropératoire [Krane 2012, Mitsui 2012, Watanabe 2016, Aoki 2018]. En oncologie rénale et hépatique, l'ICG présente une affinité spécifique pour les cellules tumorales, facilitant ainsi leur détection et leur résection chirurgicale ciblée [Branch 1982, Krane 2012, Mitsui 2012, Giorgio 2018]

Sur le plan physico-optique, l'excitation de l'ICG est réalisée par une source lumineuse spécifique, généralement un laser émettant dans le proche infrarouge, permettant l'activation de la fluorescence qui est ensuite numérisée par un capteur haute sensibilité filtrant les longueurs d'ondes non visibles. L'intégration de caméras Near-InfraRed (NIR) dans des endoscopes chirurgicaux permet une visualisation en temps réel des structures anatomiques, y compris les vaisseaux sanguins et les voies

urinaires. Cependant, cette technique présente certaines limitations, notamment une faible pénétration optique de la lumière infrarouge (5 à 10 mm), limitant ainsi la détection des lésions profondes. De plus, la métabolisation rapide de l'ICG et son excrétion nécessitent parfois des injections répétées afin de prolonger la visualisation des structures cibles.

Malgré ces contraintes, l'ICG demeure un agent de contraste peu onéreux et particulièrement efficace pour la cartographie vasculaire, la localisation de petites lésions non détectées en imagerie préopératoire et la sécurisation des résections chirurgicales [Alander 2012]. Comme illustré en figure 1.36, son utilisation en NPRA et en chirurgie hépatique illustre parfaitement son rôle croissant dans l'amélioration de la précision chirurgicale et de la sécurité des patients.



FIGURE 1.36 – Images de fluorescence au vert d'indocyanine lors d'une néphrectomie partielle robot-assistée. Source : Centre Hospitalier Universitaire (CHU) de Bordeaux.

1.3.2.3 Réalité virtuelle en chirurgie

L'intégration de la Réalité Virtuelle (RV) et des modèles tridimensionnels interactifs en peropératoire offre une assistance cognitive précieuse au chirurgien en réduisant la charge mentale associée à la reconstruction mentale des images en coupe. En fournissant une visualisation immersive, intuitive et manipulable des structures anatomiques et des lésions, ces outils optimisent la compréhension spatiale et favorisent une interprétation plus rapide et plus précise de l'anatomie spécifique au patient.

L'enjeu ne se limite plus à la simple conversion d'images en un modèle statique, mais à la génération d'une représentation dynamique et interactive de l'organe concerné. Cette approche transforme la planification peropératoire en un processus adaptatif en temps réel, facilitant la prise de décision et l'anticipation des étapes clés de l'intervention.

En urologie, plusieurs systèmes peropératoires permettent d'obtenir des rendus volumiques ou surfaciques du rein et de ses structures d'intérêt en RV [Antonelli 2019, Lasser 2012]. Ces technologies présentent une valeur clinique significative en aidant le chirurgien à visualiser précisément la tumeur, organiser mentalement les séquences de clampage et évaluer les risques de lésion des structures sous-jacentes. L'analyse tridimensionnelle apporte ainsi une précision accrue par rapport à l'interprétation traditionnelle des séquences d'imagerie en coupe.

Dans le contexte de la NPRA, plusieurs études ont démontré l'impact positif de l'utilisation de la RV sur les résultats périopératoires [Lasser 2012, Michiels 2021b, Azhar 2023, Moldovanu 2024]. Une analyse rétrospective menée sur une cohorte de 230 patients opérés avec assistance en RV (3D-IGRAPN) comparée à 415 patients opérés sans assistance en RV a mis en évidence des bénéfices significatifs de cette technologie [Michiels 2021b], notamment une réduction significative des complications majeures (3,8 % vs. 9,5 %, P = 0,04), une meilleure préservation de la fonction rénale avec une variation médiane du débit de filtration glomérulaire plus faible (-5,6 % vs. -10,5 %, P = 0,002) et un taux plus élevé d'atteinte du critère trifecta (55,7 % vs. 45,1 %, P = 0,005), traduisant une résection complète avec marges négatives, un temps d'ischémie réduit et une absence de complications majeures.

Toutefois, l'absence d'un alignement automatique du modèle avec la vue endoscopique représente une contrainte, imposant une interruption temporaire de l'acte chirurgical pour ajuster l'orientation de la reconstruction affichée en RV.

1.3.3 Réalité augmentée en chirurgie

L'intégration des données préopératoires dans le contexte peropératoire représente un défi majeur en chirurgie. En l'absence de systèmes de navigation, le chirurgien doit mentalement transposer la localisation des structures identifiées sur les images préopératoires vers une vue peropératoire endoscopique. Cette tâche est d'autant plus complexe que les modalités d'imagerie diffèrent et que des facteurs tels que les occlusions partielles, les variations de perspective et les déformations tissulaires induites par la respiration, le pneumopéritoine ou la traction instrumentale altèrent la perception anatomique.

La RA offre une solution en permettant l'alignement spatio-temporel des données préopératoires et peropératoires. Contrairement à la RV (cf. section 1.3.2.3), qui immerge l'utilisateur dans un environnement entièrement synthétique, la RA enrichit la vision du chirurgien en superposant des objets numériques à la scène opératoire réelle [Mendoza-Ramírez 2023]. Cette technologie s'intègre dans le continuum de Milgram, où coexistent informations virtuelles et monde réel [Milgram 1995]. Son application en CMI et en CO assistée par l'image renforce la précision du guidage peropératoire en rendant visibles des structures initialement imperceptibles.

Pour que la RA joue un rôle central dans l'assistance chirurgicale, trois étapes sont essentielles. Tout d'abord l'acquisition et la reconstruction des données, issues d'imageries préopératoires (TDM, IRM) (cf. sections 1.3.1.1 et 1.3.1.2) et peropératoires (échographie, fluorescence, endoscopie) (cf. sections 1.3.2.1 et 1.3.2.2), sont

déterminantes pour garantir la fiabilité du guidage. Ensuite, le recalage des référentiels d'imagerie est indispensable afin d'aligner avec précision les données préopératoires sur les images peropératoires. Cette étape est particulièrement complexe en raison des déformations anatomiques dynamiques induites par la respiration ou la manipulation instrumentale. Enfin, la visualisation optimisée des informations augmentées doit maximiser la perception du chirurgien en modulant les paramètres d'affichage tels que la transparence, la profondeur et l'ombrage, tout en minimisant l'encombrement visuel.

Sur le plan technique, différentes approches de fusion des données existent en RA chirurgicale. L'une repose sur l'utilisation d'une instrumentation externe, non présente conventionnellement et ayant un caractère invasif ou massif, tels que des marqueurs optiques et des capteurs électromagnétiques ou ultrasonores, pour suivre en temps réel la position des structures. Toutefois, cette méthode implique un équipement spécifique et peut être contraignante en CMI, où le champ opératoire est limité à l'angle de vue de la caméra endoscopique. Une alternative repose sur la vision par ordinateur, qui exploite l'analyse d'images et la reconnaissance de repères anatomiques pour ajuster en temps réel la superposition des données. Cette méthode, sans instrumentation supplémentaire, est facilement intégrable au bloc opératoire. L'approche adoptée dépend du niveau d'instrumentation disponible, variant d'un guidage purement visuel à des solutions intégrant du matériel d'imagerie comme celui des salles hybrides (cf. section 1.2.4.4).

Contrairement à certaines modalités d'imagerie peropératoire comme l'échographie, dont le champ de vision est réduit et l'interprétation exige une expertise élevée, la RA basée sur la vision permet une visualisation plus globale et intuitive des structures anatomiques. De plus, elle évite l'exposition aux radiations inhérentes aux salles hybrides, préservant ainsi la sécurité du patient et de l'équipe chirurgicale.

À la lumière des connaissances actuelles, l'intégration de la RA basée sur la vision par ordinateur en CMI est justifiée par sa capacité à fusionner des informations multi-sources pour une meilleure interprétation peropératoire sans avoir recours à une instrumentation supplémentaire. Elle contribue à réduire la charge cognitive du chirurgien en facilitant le repérage anatomique et favorise l'amélioration des résultats chirurgicaux grâce à une planification et un guidage optimisés.

L'application de la RA est particulièrement pertinente dans plusieurs spécialités chirurgicales qui nécessitent un guidage précis, gage de performance et de sécurité [Barcali 2022, Birlo 2022, Longo 2021]. En chirurgie hépatique, elle permet l'identification des structures intra-hépatiques (tumeurs, vaisseaux, segments de Couinaud) et la planification des trajectoires de résection en fonction des marges oncologiques [Giannone 2021, Labrunie 2023, Espinel 2024b]. En urologie, elle facilite la localisation tumorale et l'optimisation des abords vasculaires [Khaddad 2023, Teluob 2019, Amparore 2024b]. En chirurgie gynécologique, elle contribue au repérage des masses tumorales intrautérines [Collins 2020, Chauvet 2020]. Enfin, en chirurgie maxillo-faciale et otorhinolaryngologie, elle assiste la navigation peropératoire pour les reconstructions osseuses complexes et la chirurgie des sinus ou de la base du crâne [Kaplan 2023, Espinel 2024a, Ceccariglia 2022]. Quelques exemples

sont donnés en figure 1.37.

L'intérêt majeur de la RA réside dans sa capacité à améliorer la précision chirurgicale, à réduire le temps opératoire et à limiter les complications peropératoires. Toutefois, son implémentation repose encore sur la résolution de défis techniques et cliniques, notamment en matière de recalage peropératoire et d'ergonomie.

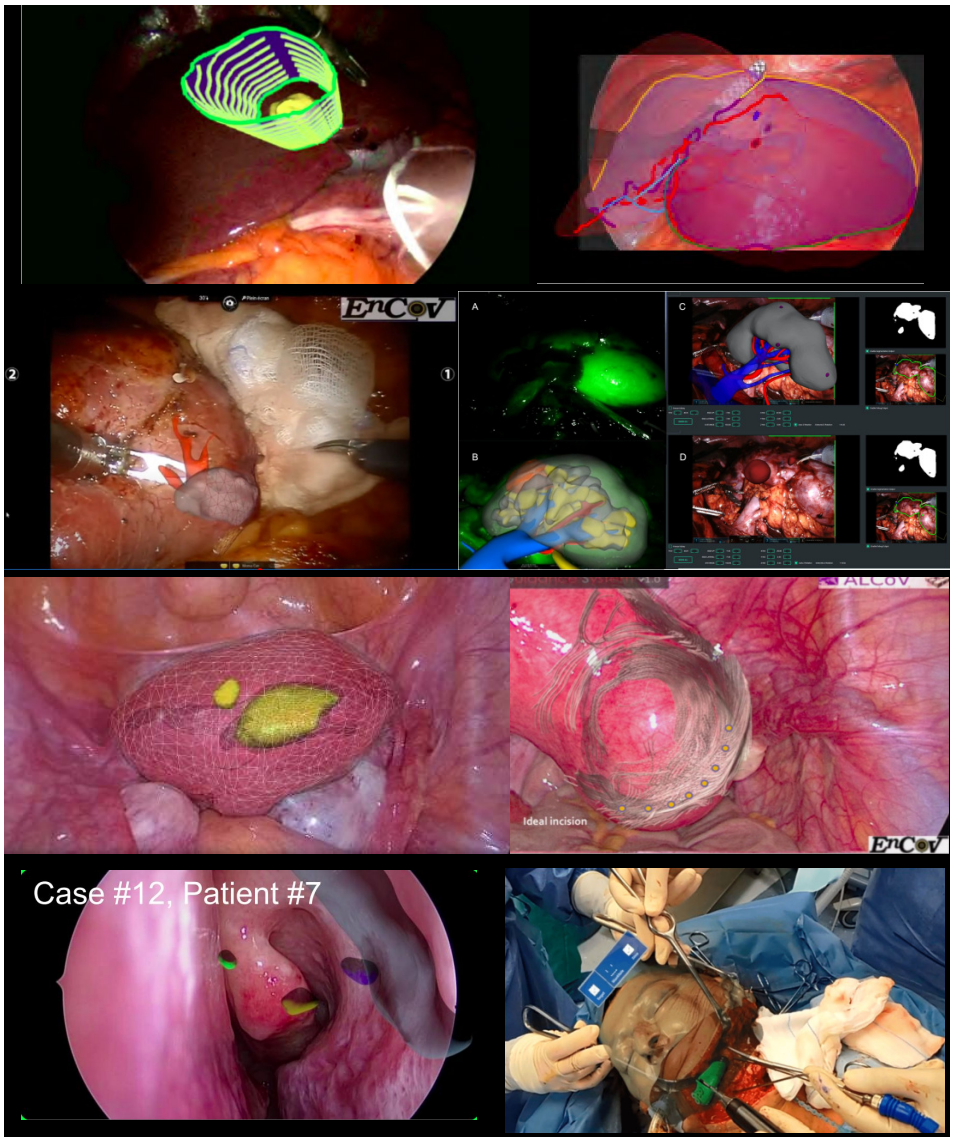


FIGURE 1.37 – Illustrations d'augmentations en réalité augmentée en : (de haut en bas) chirurgie hépatique, chirurgie urologique, chirurgie gynécologique et chirurgies maxillo-faciale et otorhinolaryngologie. Sources : [Espinel 2024b, Labrunie 2023, Teluob 2019, Amparore 2024b, Collins 2020, Chauvet 2020, Espinel 2024a, Ceccariglia 2022].

1.4 Contributions et structure du manuscrit

L'objectif de cette thèse est de concevoir et de déployer, en conditions réelles, un pipeline de RA visant à surmonter les limitations cliniques de la NPRA, notamment la localisation précise de la tumeur ainsi que des structures anatomiques sous-jacentes, vasculaires et urinaires. Au début de nos travaux, la méthodologie technique permettant l'implémentation de la RA en NPRA n'était pas encore définie. Nous disposions néanmoins d'une expertise préalable en développement de solutions pour la chirurgie de l'utérus et du foie. Toutefois, l'agencement des briques technologiques et scientifiques nécessaires restait indéterminé et les contraintes cliniques n'étaient que partiellement caractérisées. Seuls les résultats attendus ainsi que les données d'entrée étaient clairement établis. Dans ce contexte marqué par des incertitudes méthodologiques et cliniques, nous avons élaboré un système basé sur un jumeau numérique préopératoire et la vision artificielle afin d'apporter une réponse adaptée aux défis spécifiques posés par cette intervention chirurgicale.

Dans ce cadre, nous apportons trois contributions majeures :

Contribution 1 : Synchronisation des flux vidéo gauche-droit - Nous proposons une méthode permettant de synchroniser automatiquement les flux vidéo gauche-droite issus d'un endoscope stéréoscopique, en vue de garantir une cohérence temporelle optimale au sein d'un système de CAO. Une désynchronisation introduite par la chaîne d'acquisition et de traitement nuit aux calculs utilisant la stéréovision, limitant ainsi la précision des applications sous-jacentes. Après une analyse des contraintes techniques et réglementaires, nous avons opté pour une approche basée sur l'analyse d'image exploitant les reflets spéculaires des tissus. Nous avons développé un Score de Comparaison Gauche-Droite (SCGD) reposant sur un réseau de neurones, permettant de détecter, caractériser et apparier ces reflets. Notre méthode améliore significativement la précision de la stéréovision chirurgicale en temps réel, en garantissant une synchronisation plus fiable dans un système CAO ne recevant aucun signal de synchronisation.

Cette méthode a été publiée dans un article de journal [Chandelon 2024] et est détaillée au chapitre 3 de cette thèse.

Contribution 2 : Estimation automatique d'une pose globale par apprentissage profond - Nous proposons une approche de recalage automatique du jumeau numérique préopératoire sur l'image chirurgicale, spécifiquement adaptée à l'anatomie rénale, qui est dépourvue de repères visibles. Notre méthode repose sur un modèle générique de rein combiné à un réseau de neurones de bout en bout, entraîné sur un jeu de données dédié afin d'estimer directement l'alignement à partir d'images Rouge Vert Bleu (RVB) chirurgicales. L'évaluation sur sept cas cliniques démontre une forte concordance avec le recalage expert. Cette solution sans repère répond aux exigences cliniques en matière de précision et de rapidité, la rendant particulièrement adaptée à la Néphrectomie Partielle (NP).

Cette méthode a été publiée dans un article de journal [Chandelon 2025] et est présentée au chapitre 4 de cette thèse.

Contribution 3 : Suivi d'organe en temps réel par points clés utilisant une base d'images clés dynamique - Nous proposons un système de suivi dynamique en CMI intégrant une mise à jour dynamique des images clés afin de mieux gérer les variations peropératoires. Contrairement aux approches statiques, notre méthode sélectionne et met à jour les images clés en temps réel selon cinq critères optimisés, garantissant une couverture efficace de l'espace visible tout en évitant la redondance et la dérive. Cette approche renforce la précision et la robustesse du suivi tout en maintenant des performances adaptées aux contraintes du temps réel.

Initialement développée pour la chirurgie endoscopique monoculaire, cette méthode a été publiée dans un article de journal [Chandelon 2022] et a ensuite été étendue à la chirurgie endoscopique stéréoscopique dans un second article de journal [Chandelon 2023]. Elle est détaillée au chapitre 5 de cette thèse.

La portée potentielle de ces contributions dépasse le cadre spécifique de la NPRA. Elles ouvrent la voie à des avancées plus générales dans l'utilisation de la RA en CMI, avec des applications potentielles dans d'autres interventions nécessitant un guidage et un suivi précis des structures anatomiques.

Chapitre 2

Contexte : Néphrectomie partielle et méthodes de guidage par réalité augmentée

---

Dans ce second chapitre, nous définissons le cadre clinique et technique de nos travaux, centrés sur la chirurgie rénale guidée par RA. Nous commençons par une présentation du contexte clinique en abordant l'anatomie, la physiologie et les principales pathologies rénales ainsi que les techniques chirurgicales associées, avec un éclairage particulier sur la NP. Par la suite, nous exposons l'approche d'ingénierie logicielle adoptée dans nos travaux, fondée sur la méthodologie Agile, qui a permis un développement itératif et une évaluation continue des travaux de recherche au travers de preuves de concept testées en conditions réelles au bloc opératoire. Nous soulignons ensuite l'importance de l'approche centrée sur l'utilisateur, qui a constitué le fil conducteur de notre démarche scientifique et qui a guidé nos choix techniques afin de répondre aux besoins cliniques spécifiques de la NP. Nous poursuivons par une revue de la littérature des systèmes de guidage peropératoire en chirurgie rénale, en mettant en évidence les avancées en RA. Enfin, nous donnons des éléments techniques de compréhension sur les étapes d'acquisition du flux vidéo endoscopique, l'étalonnage géométrique de la caméra endoscopique et le recalage 3D-2D en chirurgie.

Sommaire

2.1 Contexte clinique de la chirurgie du rein 57
2.1.1 Anatomie et physiologie du rein 57
2.1.1.1 Emplacement anatomique 58
2.1.1.2 Morphométrie rénale 58
2.1.1.3 Fascia rénal et moyens de fixité 59
2.1.1.4 Vascularisation et innervation 61
2.1.1.5 Structure rénale 63
2.1.1.6 Organogénèse 65
2.1.1.7 Anomalies de développement 66
2.1.2 Pathologies et cancer du rein 67
2.1.2.1 Épidémiologie 68
2.1.2.2 Facteurs de risques 69
2.1.2.3 Dépistage 70
2.1.2.4 Diagnostic 70
2.1.3 Traitements du carcinome à cellules rénales 73
2.1.3.1 Traitement des petites tumeurs du rein 73
2.1.3.2 Traitement des tumeurs du rein localisées supérieures à quatre centimètres 75
2.1.3.3 Traitement d'un cancer du rein localement avancé . 77
2.1.4 Néphrectomie partielle 79
2.2 Guidage par réalité augmentée en chirurgie du rein 85
2.2.1 Généralités 85
2.2.2 Approche d'ingénierie logicielle 90
2.2.3 Approche utilisateur 92
2.2.4 État de l'art 94
2.2.5 Acquisition vidéo endoscopique 101
2.2.6 Étalonnage géométrique d'une caméra endoscopique 104

2.1 Contexte clinique de la chirurgie du rein

Dans cette section des éléments cliniques de contexte sont donnés afin d'appréhender la complexité de la situation anatomique du rein et de l'intervention chirurgicale.

2.1.1 Anatomie et physiologie du rein

Le système urinaire est responsable de la sécrétion et de l'excrétion de l'urine. Il est localisé dans les espaces rétro-péritonéal et sous-péritonéal, comme illustré en figure 2.1. Il est constitué de deux organes sécréteurs, les reins, chargés de la filtration glomérulaire et de la production de l'urine. Ceux-ci assurent un rôle physiologique essentiel à travers deux types d'actions : mécanique et hormonale.

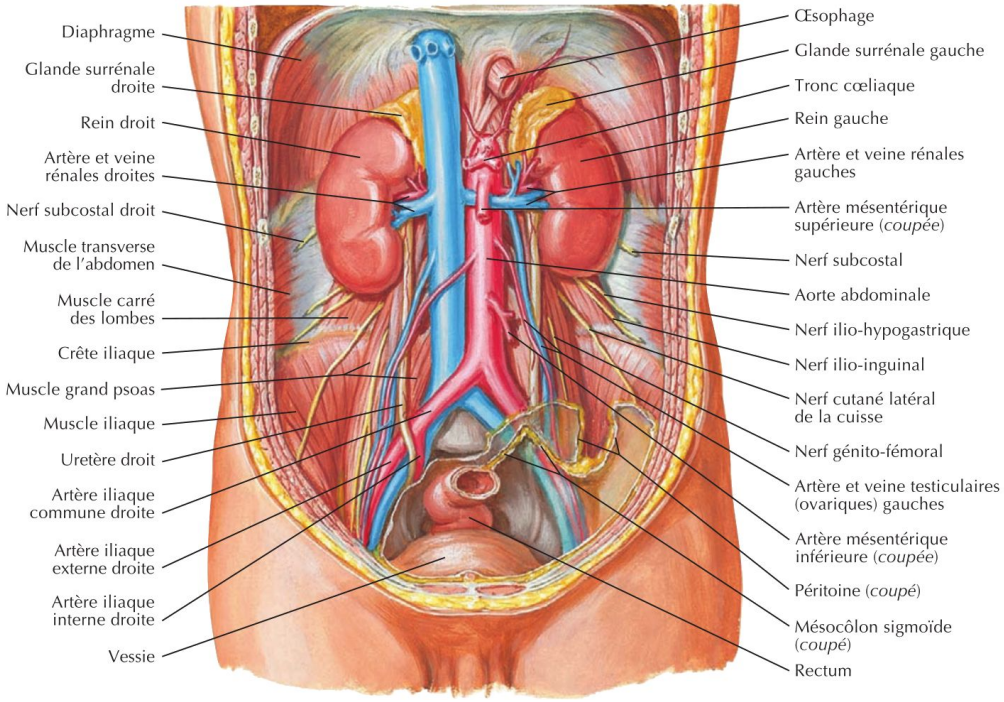


FIGURE 2.1 – Représentation schématique de la position des reins in situ en vue antérieure. Source : [Netter 2019].

L'action mécanique des reins repose sur la régulation de l'équilibre du milieu extracellulaire par la modulation des quantités d'eau et de sel éliminées dans les urines. Cette fonction permet de maintenir l'homéostasie hydrique et ionique de l'organisme. Par ailleurs, les reins interviennent dans l'élimination des déchets métaboliques par le biais du processus de filtration glomérulaire, suivi de la réabsorption et de la sécrétion tubulaire [Boudet 2017].

L'action hormonale des reins est tout aussi essentielle. Ils sécrètent l'Érythropoïétine (EPO), une hormone qui stimule la production des globules rouges par la moelle osseuse, jouant ainsi un rôle clé dans l'hématopoïèse. De plus, ils produisent la rénine, une enzyme impliquée dans la régulation de la pression artérielle via le système rénine-angiotensine-aldostérone. Ces mécanismes hormonaux participent activement au maintien de l'homéostasie cardiovasculaire et sanguine [Boudet 2017].

L'urine produite est ensuite acheminée par deux conduits excréteurs internes, les uretères, jusqu'à un organe de stockage, la vessie. L'évacuation de l'urine hors de l'organisme s'effectue par un conduit excréteur terminal, l'urètre [Kamina 2009].

2.1.1.1 Emplacement anatomique

Les reins sont localisés dans la partie supérieure de la fosse lombaire, au sein de l'espace rétro-péritonéal. Ils sont situés de part et d'autre de la colonne vertébrale, à la hauteur des vertèbres thoraciques T11 et T12 et des vertèbres lombaires L1 et L2. En raison de la présence du foie, le rein droit est situé légèrement plus bas que le rein gauche, il atteint le disque intervertébral L2-L3. Chaque rein est orienté obliquement en bas et latéralement. Son grand axe forme un angle d'environ 18° avec le plan sagittal médian. Dans le plan horizontal, cet axe présente une inclinaison postérieure comprise entre 40° et 60° par rapport au plan sagittal [Kamina 2009, Boudet 2017].

Chez l'adulte, la surface du rein est lisse, tandis qu'elle est irrégulière et polylobulée chez le nouveau-né. Morphologiquement, le rein présente une configuration en haricot, caractérisée par un bord latéral convexe et un bord médial concave. Ce dernier comporte, au niveau de son tiers moyen, le hile rénal, point d'entrée et de sortie des structures vasculaires, nerveuses et excrétrices. L'organe se caractérise par une teinte rouge brunâtre et une consistance ferme [Kamina 2009, Boudet 2017].

Sur le plan biomécanique, des mesures réalisées chez de jeunes volontaires présentant une fonction rénale normale ont montré des valeurs moyennes d'élasticité de $15,4 \pm 2,5$ kPa pour le cortex, contre $10,8 \pm 2,7$ kPa pour la médullaire [Grenier 2013]. Cette différence a également été observée lors d'expériences in-vivo sur des reins de porc [Gennisson 2012].

2.1.1.2 Morphométrie rénale

Le volume du rein varie entre 135 et 150 cm^3. Sa masse est en moyenne de 140 grammes chez l'homme et de 125 grammes chez la femme. Sa longueur moyenne est de 12 centimètres, avec des variations comprises entre 9 et 14,5 centimètres. Sa largeur est d'environ 6 centimètres, oscillant entre 4,5 et 7,5 centimètres, tandis que son épaisseur est de 3 centimètres, pouvant aller jusqu'à 4,5 centimètres. Le rein gauche est généralement plus long d'environ 1 centimètre que le rein droit [Kamina 2009, Boudet 2017].

Les dimensions rénales sont influencées par plusieurs paramètres, notamment l'âge, le sexe et la taille de l'individu. La longueur rénale est le paramètre le plus variable. Chez un individu mesurant 150 centimètres, elle est d'environ 12 centimètres, tandis qu'elle atteint 14,5 centimètres chez un sujet de 195 centimètres. Chez la femme, les dimensions du rein sont légèrement inférieures, avec une différence d'environ 0,5 centimètre par rapport à celles de l'homme. Après 60 ans, une réduction progressive des dimensions rénales est généralement observée [Kamina 2009, Boudet 2017].

2.1.1.3 Fascia rénal et moyens de fixité

Le rein et la glande surrénale sont contenus dans une loge rénale fermée, visible en figure 2.2, délimitée par le fascia périrénal résultant de la condensation périphérique de la capsule adipeuse du rein. Ce fascia comprend deux lames, une antérieure et une postérieure, qui fusionnent en haut avec le fascia diaphragmatique et en bas entre elles. Médialement, elles se prolongent dans les gaines vasculaires des vaisseaux rénaux, de l'aorte et de la veine cave inférieure. La lame postérieure adhère au fascia diaphragmatique et est séparée des muscles sous-jacents par le tissu adipeux pararénal [Kamina 2009, Boudet 2017].

La loge rénale englobe le rein, enchâssé dans une capsule adipeuse dont les trabéculations fibreuses assurent la liaison avec le fascia rénal. Le maintien en position du rein est principalement assuré par la pression abdominale. Le rein présente deux faces, postérieure et antérieure, illustrées en figure 2.3, deux bords, latéral et médial, et deux extrémités aussi appelées pôles, supérieur et inférieur [Kamina 2009, Boudet 2017].

La face postérieure, plane et rétro-péritonéale est en contact avec le diaphragme et la musculature lombaire. Sa protection osseuse est assurée par les côtes, avec un recouvrement par les onzième et douzième côtes à droite et uniquement par la douzième côte à gauche [Kamina 2009, Boudet 2017].

La face antérieure convexe présente des rapports anatomiques distincts selon le côté. À droite, elle est en contact avec l'angle colique droit, le mésocôlon ascendant, la portion descendante du duodénum et la face viscérale du foie. À gauche, la partie médiane s'articule avec l'extrémité du côlon transverse, tandis que la partie supérieure est en rapport avec la rate, la queue du pancréas et les vaisseaux spléniques [Kamina 2009, Boudet 2017].

Son bord latéral, également convexe, est en rapport avec le foie du côté droit et avec la rate du côté gauche [Kamina 2009, Boudet 2017].

Son bord médial concave au niveau du hile correspond à l'insertion du pédicule vasculo-nerveux. Il se divise en deux segments : un segment suprahilaire en contact avec la veine cave inférieure à droite et la glande surrénale à gauche et un segment infrahilaire, longée par l'uretère et positionnée en regard de la veine cave inférieure à droite et de l'aorte abdominale à gauche [Kamina 2009, Boudet 2017].

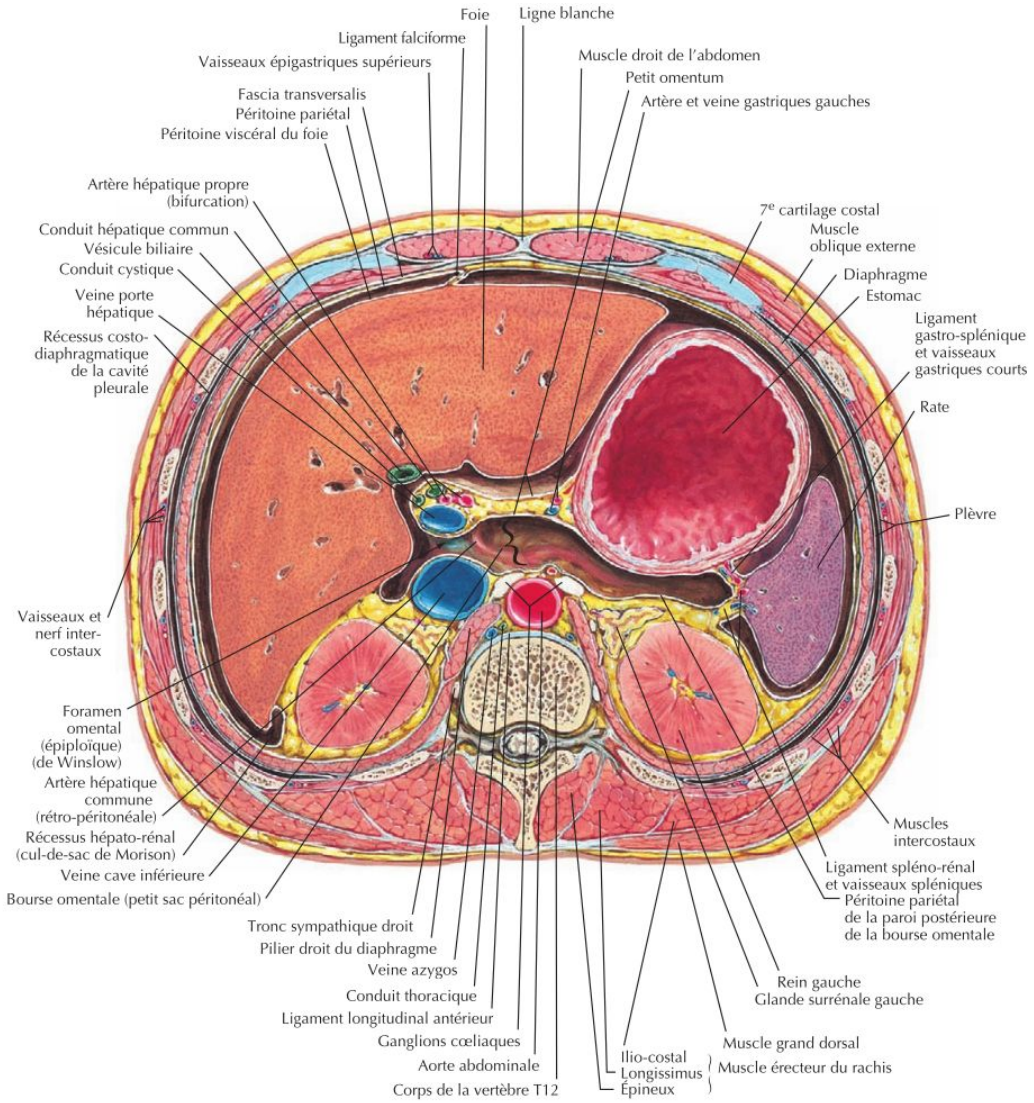


FIGURE 2.2 – Coupe transversale schématique de l'abdomen au milieu de T12. Source : [Netter 2019].

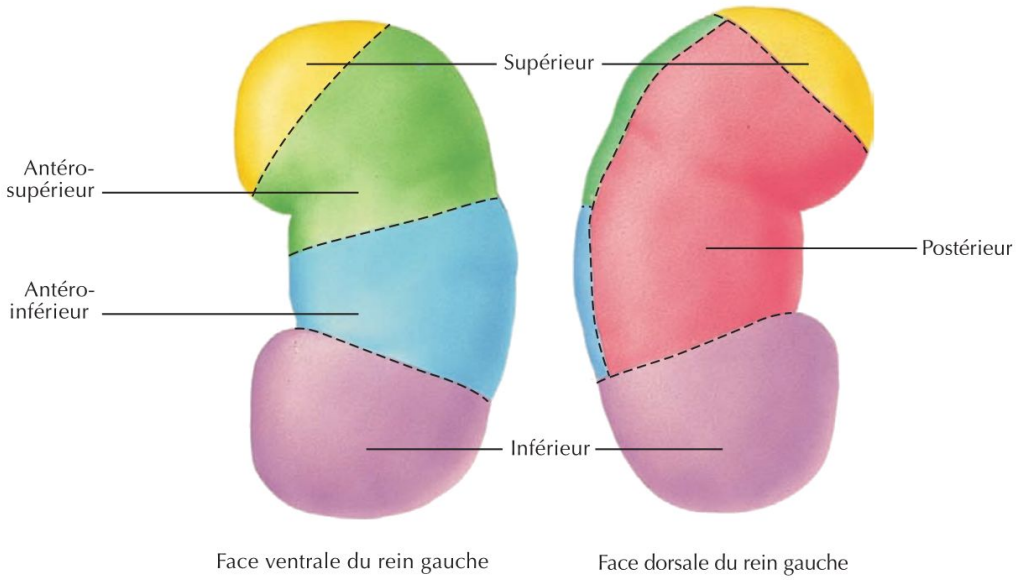


FIGURE 2.3 – Segments rénaux des faces ventrale et dorsale d'un rein gauche. Source : [Netter 2019].

Son extrémité supérieure est en relation avec la glande surrénale et le diaphragme, tandis que son extrémité inférieure, plus éloignée de la ligne médiane, se situe à distance des crêtes iliaques, à environ trois centimètres à droite et cinq centimètres à gauche [Kamina 2009, Boudet 2017].

Le hile rénal est une fente longitudinale d'environ quatre centimètres de longueur et un centimètre et demi d'épaisseur, situé à hauteur de la vertèbre lombaire L1. Cette dépression concave constitue l'entrée du sinus rénal et assure le passage des éléments du pédicule vasculo-nerveux. L'organisation de ce pédicule suit une stratification antéro-postérieure caractéristique, avec les veines rénales en position la plus antérieure, suivies des artères rénales et enfin le pelvis rénal situé en position la plus postérieure. Le hile est délimité par deux lèvres, antérieure et postérieure, qui encadrent l'entrée des structures vasculaires et excrétrices dans l'organe [Kamina 2009, Boudet 2017].

2.1.1.4 Vascularisation et innervation

Le rein est perfusé par l'artère rénale, généralement unique de chaque côté, qui naît du bord latéral de l'aorte, en dessous de l'artère mésentérique supérieure, au niveau des vertèbres lombaires L1 et L2 ainsi illustré sur la figure 2.1. L'artère rénale droite émerge légèrement plus haut que son homologue à gauche et suit un trajet oblique vers le hile rénal. Elle mesure environ 5 à 6 centimètres à droite et 3 à 4 centimètres à gauche pour un calibre de 4 à 6 millimètres [Kamina 2009, Boudet 2017].

L'irrigation rénale repose sur une distribution artérielle terminale segmentaire, divisant le rein en cinq territoires : supérieur, inférieur, antéro-supérieur, antéro-inférieur et postérieur. A proximité du hile, l'artère rénale se scinde en deux branches principales : une branche antérieure, donnant naissance aux artères des segments supérieurs, antérieurs et inférieurs et une branche postérieure, vascularisant le segment postérieur. Les artères segmentaires se ramifient en artères interlobaires, puis en artères arquées et interlobulaires, qui irriguent les néphrons, comme illustré en figure 2.4 [Kamina 2009, Boudet 2017].

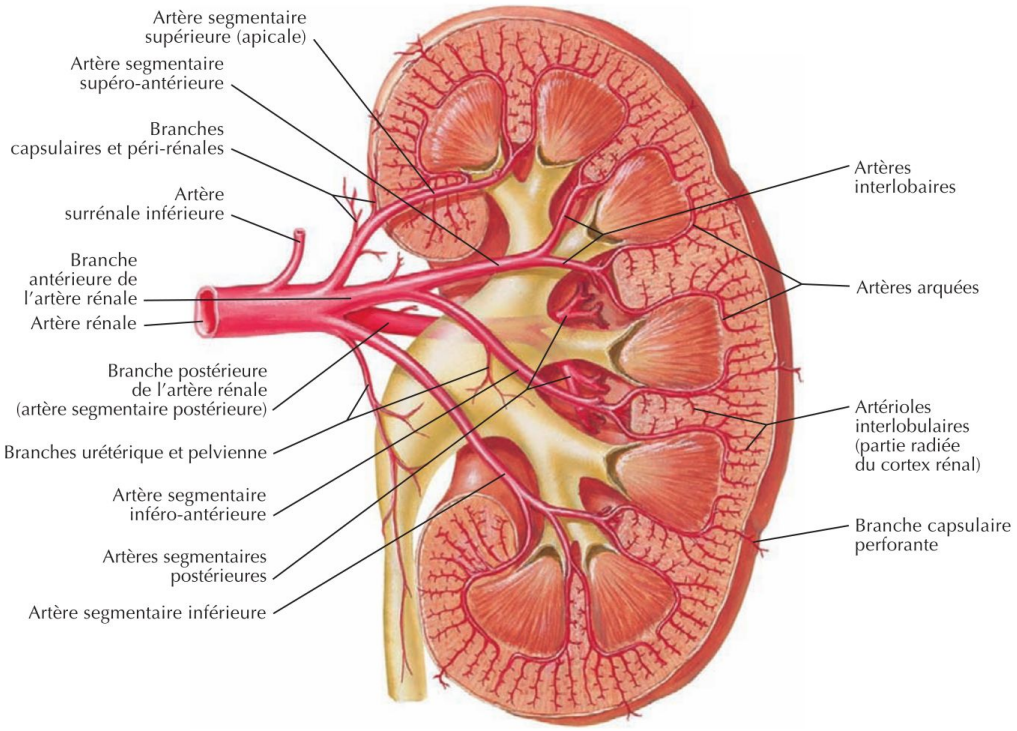


FIGURE 2.4 – Artères intrarénales rendues visibles via une coupe frontale du rein gauche : vue antérieure. Source : [Netter 2019].

Des artères polaires, présentes dans environ 15 % des cas, peuvent vasculariser les pôles rénaux sans passer par le hile. Par ailleurs, l'irrigation de la capsule adipeuse provient de l'artère rénale, des artères surrénales, testiculaires ou ovariennes et parfois de branches coliques, lombaires ou de l'aorte, formant un réseau exo-rénal [Kamina 2009, Boudet 2017].

Le drainage veineux est assuré par les veines rénales droite et gauche, qui collectent le sang des veines segmentaires au niveau du hile. Le drainage lymphatique s'organise autour de deux plexus : un intrarénal, entourant les tubules et convergeant vers les nœuds latéro-aortiques et latéro-caves et un capsulaire, qui peut s'anastomoser avec les lymphatiques coliques [Kamina 2009, Boudet 2017].

L'innervation du rein provient du plexus rénal, issu des ganglions cœliaques et aortico-rénaux. Ce réseau nerveux, dense à la face postérieure de l'artère rénale, contient des neurofibres sympathiques et parasympathiques (provenant du nerf vague), qui accompagnent les artères et les tubules jusqu'aux glomérules [Kamina 2009, Boudet 2017].

2.1.1.5 Structure rénale

Le rein est une structure anatomique complexe, constituée de cinq à treize lobes rénaux, organisée autour du parenchyme rénal et des voies excrétrices. Il est constitué de deux régions distinctes : le cortex et la médulla, dont l'organisation structurelle est illustrée en figure 2.5. L'ensemble de l'organe est enveloppé par une capsule fibreuse mince et extensible, composée de fibres élastiques et musculaires lisses. Cette capsule se prolonge à l'intérieur du sinus rénal, engainant les calices et le pelvis rénal. Elle contribue également au maintien de l'architecture rénale [Kamina 2009, Boudet 2017].

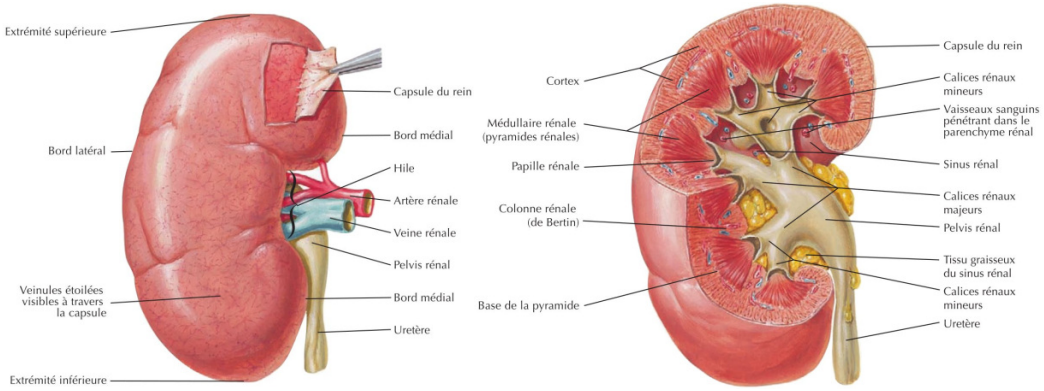


FIGURE 2.5 – Structure macroscopique du rein. (à gauche) Face antérieure du rein droit. (à droite) Rein droit coupé sur plusieurs plans, montrant le parenchyme et le pelvis rénal. Source : [Netter 2019].

Le cortex, zone périphérique brun-rouge et granuleuse, s'étend entre les pyramides rénales pour former les colonnes de Bertin. Il se divise en une zone externe, contenant les corpuscules rénaux et les tubules contournés et une zone interne abritant en plus les tubules collecteurs et les vaisseaux arqués. Son organisation histologique est représentée schématiquement en figure 2.6 [Kamina 2009, Boudet 2017].

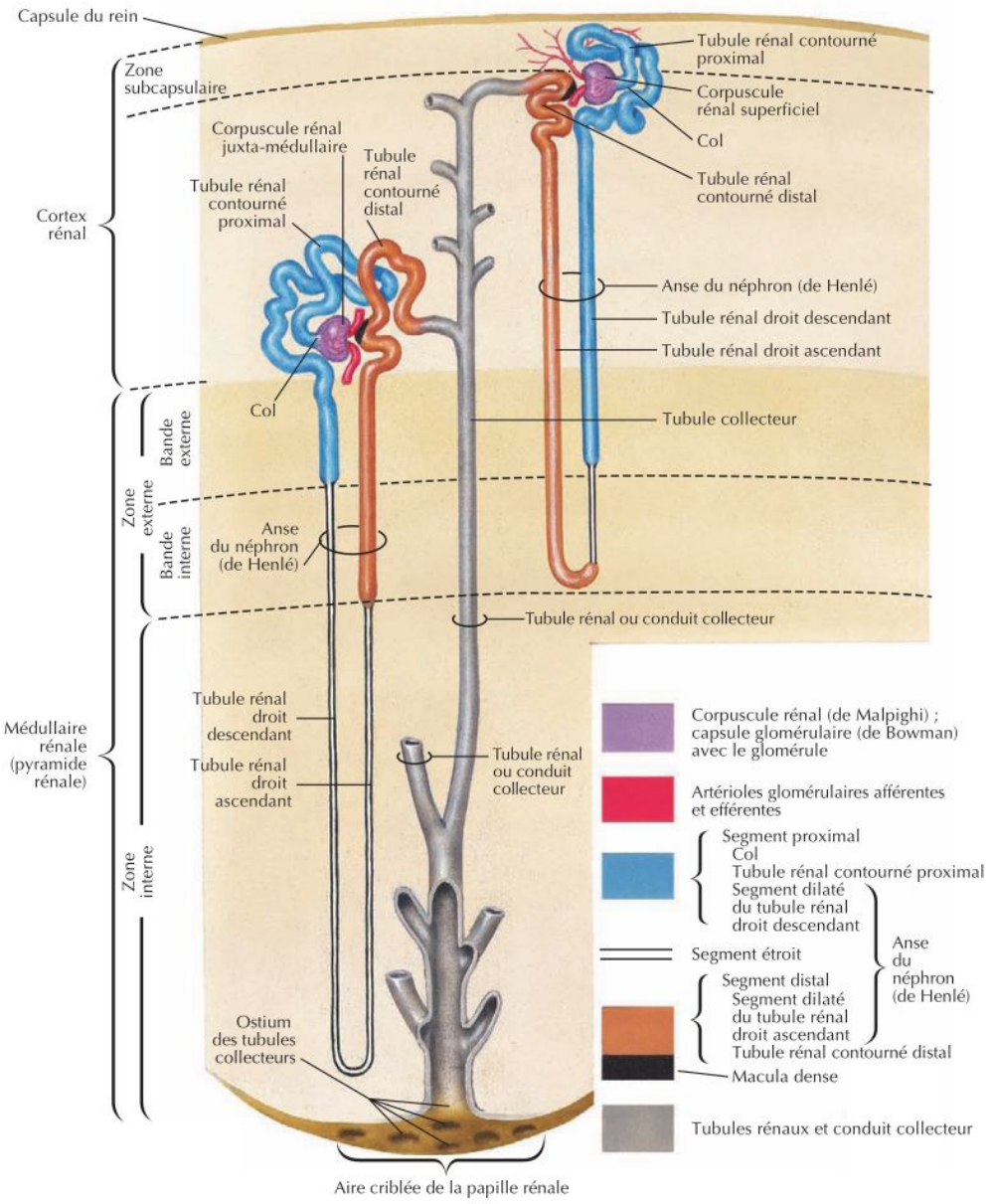


FIGURE 2.6 – Représentation schématique du néphron et tubule rénal collecteur. Source : [Netter 2019].

La médulla rénale se distingue par son aspect strié et sa teinte plus claire. Elle est constituée de cinq à onze pyramides rénales, également appelées pyramides de Malpighi. Chaque pyramide possède une base orientée vers le cortex et un sommet interne appelé papille rénale. Cette dernière débouche dans le sinus rénal, comme illustré en figures 2.5 et 2.6. Cette région profonde abrite les structures tubulaires essentielles à la fonction rénale, notamment les anses du néphron, les tubules collecteurs, les conduits papillaires et les vaisseaux droits. Les papilles rénales s'ouvrent dans les calices mineurs, structures en entonnoir qui collectent l'urine primaire. Elles convergent ensuite en trois calices majeurs (supérieur, moyen et inférieur), lesquels fusionnent à leur tour pour former le pelvis rénal, également désigné sous les termes de bassinet ou pyélon. L'ensemble calices et pelvis rénal forme les voies excrétrices rénales. Elles sont tapissées d'un épithélium transitionnel, favorisant leur adaptation aux variations de volume urinaire. Ces structures possèdent également une musculature responsable de contractions rythmiques facilitant le transport de l'urine. Les calices mineurs effectuent environ dix à douze contractions par minute, tandis que le pelvis rénal produit des ondes péristaltiques à une fréquence de trois à six par minute, propulsant l'urine vers l'uretère [Kamina 2009, Boudet 2017].

L'unité fonctionnelle du rein, le néphron, assure les processus de filtration glomérulaire, de réabsorption et de sécrétion tubulaire. Une représentation schématique est donnée en figure 2.6. Il comprend un corpuscule rénal, composé de la capsule glomérulaire et du glomérule capillaire où s'effectue l'ultrafiltration du plasma menant à la formation de l'urine primitive. Le tubule contourné proximal, d'environ 14 millimètres de long et 30 à 60 micromètres de calibre, assure la réabsorption active de l'eau et des solutés grâce à son épithélium prismatique ou cubique simple riche en microvillosités. Il se prolonge par l'anse du néphron, une structure en U aux segments ascendant et descendant tapissés d'un épithélium pavimenteux simple. Le tubule contourné distal, plus court (5 millimètres) et de calibre réduit (20 à 50 micromètres), rejoint ensuite un tubule collecteur. Ces derniers convergent pour former les conduits papillaires, qui s'abouchent aux papilles rénales et assurent l'évacuation de l'urine vers les calices [Kamina 2009, Boudet 2017].

2.1.1.6 Organogénèse

Le système uro-génital se développe à partir du mésoblaste intermédiaire selon un axe crânio-caudal, formant successivement trois structures : le pronéphros, le mésonéphros et le métanéphros, illustrées en figure 2.7 [Kamina 2009].

Le pronéphros, situé en région cervicale, apparaît brièvement puis régresse dès la quatrième semaine. Le mésonéphros, qui se développe ensuite en région thoracolombaire, contient des tubules reliés au conduit mésonéphrique, qui s'ouvrent dans le cloaque. La plupart de ces structures disparaissent, mais certains éléments persistent pour former les voies génitales masculines ou des vestiges embryonnaires chez la femme.

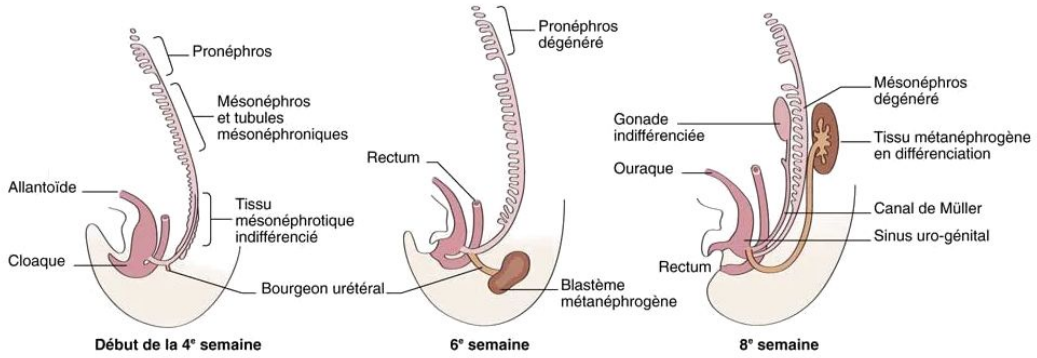


FIGURE 2.7 – Évolution du sinus uro-génital et schématisation de l'apparition du bourgeon urétéral, développement du blastème métanéphrogène. Source : [Cormier 2021].

Le métanéphros, qui donnera le rein définitif, apparaît en région caudale vers la cinquième semaine. Il résulte de l'interaction entre le diverticule métanéphrique, issu du conduit mésonéphrique et le blastème métanéphrogène, dérivé du mésoblaste. Le diverticule s'allonge vers le blastème, donnant naissance aux voies urinaires : uretère, pelvis rénal, calices et conduits collecteurs.

Le blastème métanéphrogène se fragmente en coiffes qui se transforment en vésicules, puis en tubules. L'extrémité des tubules s'invagine pour former la capsule glomérulaire. Ce processus aboutit à la formation du néphron et de la capsule rénale, assurant le développement du rein fonctionnel.

2.1.1.7 Anomalies de développement

Les anomalies du développement rénal peuvent affecter la formation, la position ou la fusion des reins, donnant lieu à diverses malformations [Kamina 2009].

Parmi elles, l'agénésie rénale se caractérise par l'absence de formation d'un ou des deux reins. Lorsqu'elle est bilatérale, cette anomalie est incompatible avec la vie, entraînant une issue fatale peu après la naissance. En revanche, l'agénésie unilatérale est souvent associée à une absence partielle ou totale des organes génitaux du même côté, bien qu'elle puisse être compatible avec une vie normale [Kamina 2009].

Le rein polykystique, quant à lui, résulte d'un défaut de développement des conduits collecteurs, provoquant la formation de multiples kystes qui altèrent la fonction rénale [Kamina 2009].

Les reins ectopiques, déplacés hors de leur position normale, peuvent parfois être surnuméraires. Ils se situent généralement en région iliaque ou pelvienne [Kamina 2009].

Les anomalies de fusion concernent des reins qui s'unissent partiellement ou totalement. Le rein en fer à cheval résulte de la fusion des pôles inférieurs ou supérieurs des deux reins. Une fusion complète aboutit à la formation d'un rein discoïde, un organe unique. Enfin, les reins doubles, situés du même côté, présentent une fusion partielle ou totale des extrémités en contact [Kamina 2009].

2.1.2 Pathologies et cancer du rein

Les pathologies rénales regroupent un ensemble d'affections allant des néphropathies bénignes aux Maladies Rénales Chroniques (MRC) sévères, susceptibles d'évoluer vers une Insuffisance Rénale Chronique Terminale (IRCT). La néphropathie désigne toute atteinte rénale, qu'elle soit d'origine inflammatoire (néphrite), dégénérative (néphrose) ou métabolique (néphropathie diabétique, néphropathie hypertensive). La MRC se définit par des altérations fonctionnelles et/ou structurelles rénales persistantes depuis plus de trois mois et constitue un enjeu majeur de santé publique en raison de son impact sur la morbi-mortalité cardiovasculaire [Haute Autorité de Santé 2023].

En France, la prévalence de la MRC est estimée entre 7 % et 10 % de la population adulte, soit environ 5,7 millions de personnes. Parmi ces patients, en 2020, 91 875 étaient en IRCT, nécessitant une prise en charge par traitement de suppléance, soit la dialyse, soit la transplantation rénale [Haute Autorité de Santé 2023]. Selon le rapport REIN 2022 de l'agence de la biomédecine, le nombre total de patients sous traitement de suppléance continue d'augmenter, avec une prévalence de l'IRCT traitée estimée à environ 1 pour 1 000 habitants [Agence de la biomédecine 2022].

L'incidence de l'IRCT a cependant légèrement diminué depuis 2017 (-0,6 % par an). Cette évolution s'explique en partie par l'amélioration des stratégies de prise en charge, notamment un meilleur dépistage des néphropathies à risque (hypertension, diabète, néphropathies d'origine vasculaire) et un accès accru à la transplantation rénale. Des inégalités persistent encore dans l'accès à la greffe, avec des disparités selon l'âge, les comorbidités et le lieu de résidence des patients [Agence de la biomédecine 2022].

Le traitement de suppléance repose principalement sur deux approches : la dialyse (hémodialyse ou dialyse péritonéale) et la transplantation rénale. En 2022, 55 % des patients étaient en dialyse et 45 % porteurs d'un greffon fonctionnel. La transplantation préemptive, réalisée avant l'instauration de la dialyse, reste minoritaire (environ 4 % des greffes), malgré son bénéfice démontré sur la survie et la qualité de vie des patients [Agence de la biomédecine 2022].

L'amélioration du dépistage précoce des néphropathies et la mise en place de parcours de soins optimisés, impliquant une coordination pluridisciplinaire entre médecins généralistes, néphrologues et autres professionnels de santé, constituent des axes prioritaires pour ralentir la progression de la MRC et améliorer la prise en charge des patients [Haute Autorité de Santé 2023].

Toutefois, au-delà des pathologies rénales chroniques, une autre problématique majeure affecte les reins : le cancer du rein. Le reste de cette partie lui est dédié,

car il constitue la cible clinique de nos travaux de recherche.

2.1.2.1 Épidémiologie

En 2022, l'incidence mondiale du cancer s'élevait à 20 millions de nouveaux cas, accompagnés de 9,7 millions de décès. Parmi les cancers les plus fréquemment diagnostiqués figurent ceux du poumon (12,4 %), du sein (11,5 %) et du côlon (9,6 %) [Ferlay 2024].

En France, le cancer constitue la principale cause de mortalité chez l'homme et la deuxième chez la femme, avec 433 136 nouveaux cas recensés en 2023, dont 57 % chez les hommes et 43 % chez les femmes [Lapôtre-Ledoux 2023].

Le cancer du rein est le quatorzième cancer le plus fréquent à travers le monde, avec une estimation de 434 840 nouveaux cas en 2022 [Ferlay 2024]. Son incidence varie selon les régions et les populations, avec des taux plus élevés chez les hommes, les personnes âgées de plus de 75 ans et dans les pays occidentaux, notamment en Europe et en Amérique du Nord [Capitanio 2019]. Le nombre de cancers du rein a augmenté dans la plupart des pays au cours des dernières décennies et en particulier chez les moins de 50 ans [Huang 2022]. Les causes font l'objet de débat, mais pourraient être expliquées par une utilisation croissante des examens d'imagerie associée à des découvertes fortuites ainsi que la prévalence grandissante de facteurs de risque tels que l'obésité et le vieillissement de la population [Wong 2020].

En France, 17 141 nouveaux cas de cancers du rein ont été estimés en 2023, soit une augmentation d'incidence de 199 % entre 1990 et 2023. Cette augmentation est particulièrement marquée chez les individus de moins de 50 ans et chez les hommes [Lapôtre-Ledoux 2023]. La mortalité reste cependant stable avec 5 589 décès en 2018, soit un taux de mortalité de 5 décès pour 100 000 hommes et 1,5 pour 100 000 femmes [Defossez 2019].

Sur le plan histologique, 90 % des tumeurs rénales de l'adulte correspondent à des Carcinomes à Cellules Rénales (CCR), dominés par le sous-type à cellules claires (70 %), suivi des papillaires (10 à 15 %) et chromophobes (5 %). Les autres formes restent rares, chacune représentant moins de 1 % des cas [Bukavina 2022]. Des coupes histologiques des sous-types les plus courants sont illustrés en figure 2.8.

Chez l'enfant, la tumeur de Wilms constitue le cancer du rein le plus fréquent, il apparaît habituellement chez l'enfant de moins de 5 ans, mais elle peut être parfois diagnostiquée chez un enfant plus âgé et exceptionnellement à l'âge adulte. Ce néphroblastome représente environ 5 % des cancers pédiatriques chez les moins de 15 ans, avec une incidence estimée entre 5 % et 14 % de l'ensemble des cancers infantiles, soit environ un cas pour 10 000 naissances et une centaine de cas annuels en France [Spreafico 2021]. Des formes bilatérales et synchrones sont observées dans environ 5 % des cas. Sur le plan histopathologique, cette tumeur embryonnaire du rein associe des cellules blastématiques, stromales et épithéliales.

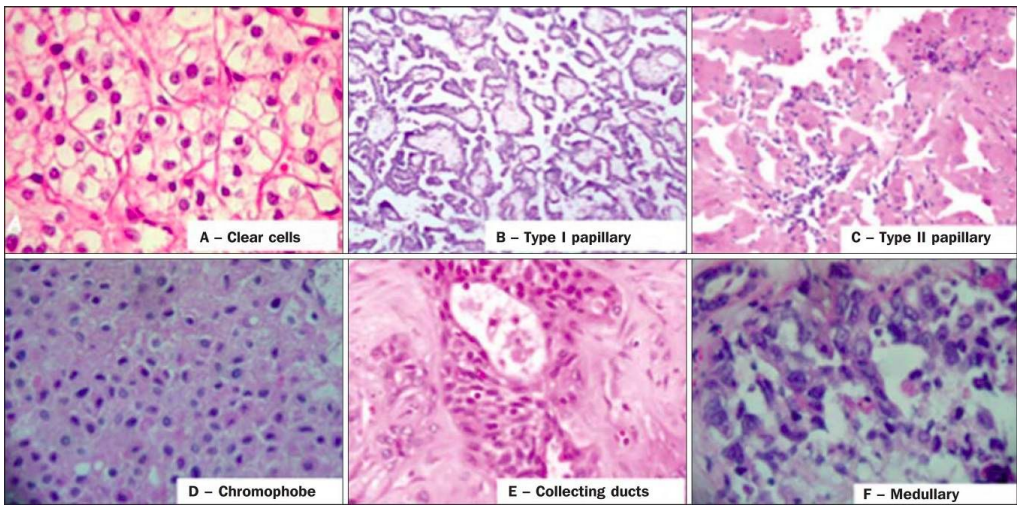


FIGURE 2.8 – Coupes histologiques des sous-types de Carcinomes à Cellules Rénales (CCR) les plus courants. A : CCR à cellules claires / B : CCR papillaire de type 1 / C : CCR papillaire de type 2 / D : CCR chromophobe / E : CCR du canal collecteur / F : CCR médullaire. Source : [Muglia 2015].

2.1.2.2 Facteurs de risques

Les facteurs de risque non modifiables du cancer du rein incluent l'âge, avec un pic d'incidence après 75 ans [Znaor 2015, Defossez 2019], le sexe caractérisé par un rapport d'incidence homme/femme de 1,5 [Scelo 2018, Scelo 2018, Defossez 2019] ainsi que les formes familiales qui représentent 3 % à 5 % des cas [Haas 2014].

Les facteurs de risque modifiables du cancer du rein sont majoritairement liés au mode de vie et constituent des cibles potentielles de prévention primaire. Le syndrome métabolique, l'obésité, l'hypertension artérielle et le tabagisme figurent parmi les principales influences reconnues [Macleod 2013, Sanfilippo 2014, Defossez 2019, Capitanio 2019, Al-Bayati 2019, Liu 2019b, Huang 2022, Du 2022, Graff 2022, Alcala 2022, Deng 2023]. L'insuffisance rénale chronique accroît également le risque de développer un cancer du rein, avec des atteintes fréquemment bilatérales et/ou multifocales en cas d'IRCT [Hora 2008, Macleod 2013, Christensson 2013, Lowrance 2014, Pan 2019, Åkerlund 2021, Bukavina 2022]. Le rôle du diabète comme facteur de risque indépendant du cancer du rein est cependant controversé [Bukavina 2022, Al-Bayati 2019, Macleod 2013].

Des comportements protecteurs tels qu'une activité physique régulière et une alimentation riche en légumes, fruits ou vitamine C pourraient contribuer à réduire le risque de cancer du rein [Behrens 2013, Jia 2015, Moore 2016, Zhang 2017, Tahbaz 2018, Bukavina 2022, Liao 2022]. Toutefois, les mécanismes physiopathologiques impliqués restent insuffisamment caractérisés et aucune étude de haut niveau de preuve n'a permis de confirmer ces hypothèses de manière concluante.

2.1.2.3 Dépistage

La faible prévalence du cancer du rein limite l'efficacité d'un dépistage systématisé. À ce jour, aucun biomarqueur sanguin ou urinaire ni protocole d'imagerie n'a mis en évidence un rapport coût-bénéfice favorable [Rossi 2018, Campi 2021, Campi 2023, Diana 2023]. Néanmoins, l'intégration du dépistage du cancer du rein à celui du poumon suscite un intérêt croissant et fait actuellement l'objet d'investigations [Usher-Smith 2022, Harrison 2023].

Une surveillance ciblée des individus à haut risque apparaît plus pertinente. Dans le cadre de la maladie génétique de von Hippel-Lindau, un dépistage par IRM abdominale biennale est recommandé dès l'âge de 15 ans [Daniels 2023]. D'autres formes héréditaires de cancer du rein pourraient également bénéficier de stratégies de dépistage adaptées [Forde 2020].

Chez les patients atteints d'insuffisance rénale terminale, dont le risque de CCR est multiplié par 10, une surveillance régulière pourrait représenter une approche valable [Åkerlund 2021]. Le comité de transplantation de l'association française d'urologie préconise un suivi radiologique des reins natifs chez les patients greffés [Goujon 2021].

2.1.2.4 Diagnostic

La majorité des tumeurs rénales sont découvertes de manière fortuite lors d'examens d'imagerie, principalement par échographie ou TDM abdominale, prescrits pour d'autres indications médicales [Jayson 1998]. Le diagnostic et la caractérisation de ces masses reposent sur une approche multimodale combinant échographie-Doppler, TDM et IRM, réalisées avec ou sans injection de produit de contraste.

L'échographie-Doppler constitue souvent le premier examen révélant fortuitement une masse rénale. Elle permet d'évaluer la nature kystique ou solide de la lésion et de détecter des caractéristiques évocatrices, comme une hyperéchogénicité marquée, suggérant un AngioMyoLipome (AML). Le Doppler aide à estimer la vascularisation tumorale et à rechercher une extension veineuse. Toutefois, son utilité reste limitée pour l'évaluation préopératoire des tumeurs rénales.

La TDM représente l'examen de référence pour l'évaluation des tumeurs rénales, sous réserve de l'absence de contre-indication aux Produits de Contraste Iodés (PCI). Les PCI sont associés à des réactions d'hypersensibilité dans environ 3 % des cas, avec un risque de réaction anaphylactique sévère de 0,04 % et un potentiel de néphrotoxicité dont la prévalence varie entre 2 % et 6 % selon la voie d'administration [Thomsen 1998, Van der Molen 2018a, Van der Molen 2018b]. Chez les patients à risque, une alternative diagnostique doit être envisagée, soit par IRM ou échographie. Chez les patients dialysés, une TDM avec injection est généralement préférable à une IRM si le bénéfice diagnostique est équivalent.

L'acquisition des images en TDM (cf. section 1.3.1.1) suit plusieurs phases : une acquisition sans injection, indispensable pour quantifier la prise de contraste ; une phase de néphrographie corticale 30 à 40 secondes post-injection ; une phase de

néphrographie tubulaire 80 à 100 secondes post-injection et une phase excrétoire tardive selon l'indication [Kim 2009, Guzzo 2009, Bosniak 2012]. Un exemple est donné en figure 2.9. La TDM bi-énergie et spectrale permet une réduction de la dose de PCI (inférieure à 100 ml) tout en augmentant la sensibilité au rehaussement des tumeurs hypovasculaires [Çamlıdağ 2019, Sadoughi 2019, Dai 2018] et à la caractérisation des masses kystiques complexes [Mileto 2015, Stehlé 2023, Reizine 2024]. Elle limite les artefacts liés aux masses rénales centrales et métalliques. Dans le cadre de l'étude des masses solides, la TDM spectrale pourrait permettre de développer de nouveaux biomarqueurs afin de mieux différencier les tumeurs de haut grade [Zhang 2024b, Liang 2022].

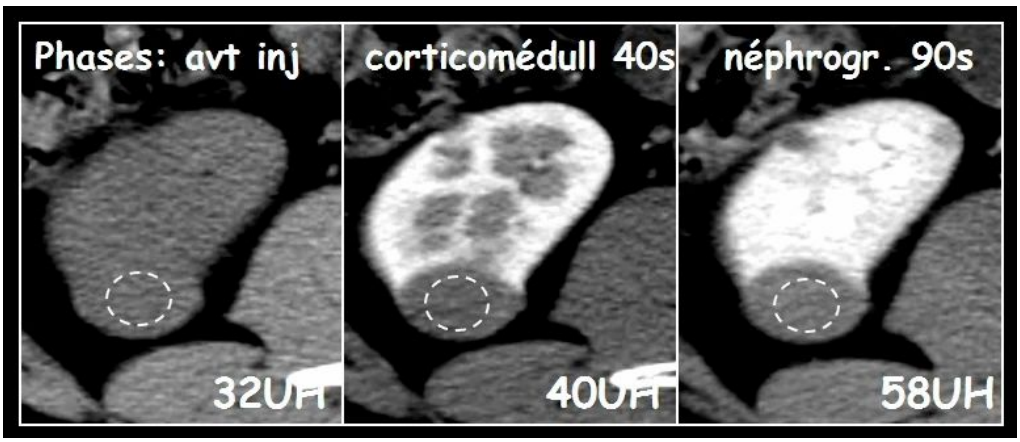


FIGURE 2.9 – Exemple d'une acquisition d'images en tomodensitométrie pour le diagnostic d'une tumeur rénale papillaire (de gauche à droite) avant injection, phase de néphrographie corticale et phase de néphrographie tubulaire. Source : Pr. J.M. Correas.

L'IRM ne se limite plus à une modalité alternative à la TDM en cas de contre-indication aux PCI. Elle est particulièrement indiquée pour la caractérisation des masses kystiques et solides atypiques demeurées indéterminées au scanner, le bilan d'extension vasculaire et le suivi des formes héréditaires [Maurer 2017, Schieda 2017]. La séquence multiparamétrique repose sur des acquisitions en pondération T2 (avec et sans saturation de graisse), T1 en phase et opposition de phase (pour la détection de la graisse intratumorale), en diffusion à B élevé (800 à 1000), ainsi qu'une séquence dynamique après injection de gadolinium [Schieda 2015, Cornelis 2017, De Leon 2019, Lopes Vendrami 2017, Ludwig 2020].

L'IRM (cf. section 1.3.1.1) peut apporter des informations complémentaires à la TDM pour l'exploration préopératoire des tumeurs rénales avec thrombus cave. Elle permet une meilleure évaluation de l'envahissement de la paroi veineuse pouvant nécessiter une reconstruction chirurgicale ou de la limite supérieure du thrombus [Janus 1991, Kabala 1991, Hallscheidt 2005, Mueller-Lisse 2010].

L'évaluation en TDM et en IRM permet également d'orienter le diagnostic histologique présomptif en fonction du caractère hypervasculaire ou homogène du rehaussement tumoral. Pour la prédiction du risque de malignité des lésions kystiques, la classification de Bosniak révisée, donnée en annexe D, est utilisée [Silverman 2019]. Ces examens sont essentiels pour l'évaluation préopératoire en définissant la taille tumorale, sa localisation, ses rapports avec les structures avoisinantes (voies excrétrices, vaisseaux, graisse périrénale) et en recherchant une tumeur synchrone controlatérale. De multiples scores morphométriques ont été décrits avec pour objectif d'aider le clinicien dans la planification opératoire [Rosevear 2012]. Les plus anciens et utilisés sont les scores RENAL ou PADUA, détaillés en annexes E et F [Kutikov 2009, Ficarra 2009, Pinto 2024]. Sur les dix dernières années, seize autres classifications ont été décrites [Veccia 2020].

La biopsie rénale, réalisée sous guidage échographique ou scanographique est indiquée lorsque la confirmation histologique est nécessaire pour orienter la stratégie thérapeutique [Yano 2018, Amaral 2021, Patel 2021, Jasinski 2022, Jiang 2022a, Ferrari 2021, Okhunov 2021, Lobo 2020, Couture 2021, Patel 2020a]. Le Comité de Cancérologie de l'Association Française d'Urologie (CCAFU) recommande son utilisation avant un traitement médical, un traitement focal (radiofréquence, curiethérapie ou radiothérapie), une néphrectomie élargie pour tumeur localisée jugée inaccessible à une chirurgie partielle ou en cas d'incertitude diagnostique. Les biopsies sont réalisées avec des aiguilles coaxiales 16-18G afin de limiter le risque de dissémination tumorale. La sensibilité et la spécificité de la biopsie rénale dépassent 95 %, avec une concordance biopsie/pièce opératoire atteignant 90 % [Marconi 2016, Pierorazio 2016]. Les complications restent rares (2 à 5 %) et sont principalement de grade faible selon la classification de Clavien-Dindo [Posielski 2019, Amaral 2021]. L'hématome post-biopsique constitue l'effet indésirable le plus fréquent et ne nécessite généralement pas de prise en charge spécifique [Robert 2023]. Les complications majeures nécessitant une intervention concernent 0,4 % à 2 % des cas [Marconi 2016, Posielski 2019, Garbens 2020].

L'imagerie fonctionnelle par TEP-scanner au fluorodésoxyglucose (TEP 18-FDG) n'a pas démontré d'intérêt dans la caractérisation des tumeurs rénales, mais son utilité est étudiée dans le bilan d'extension initial et la détection des récidives. Le développement des nouvelles modalités d'imagerie associées à l'IA, notamment en radiomique, pourrait améliorer la caractérisation des sous-types tumoraux, affiner la stratification pronostique et prédire la réponse aux traitements [Mühlbauer 2021, Rallis 2021].

En complément de l'imagerie et de la biopsie, le bilan biologique standard d'une tumeur rénale localisée doit comprendre le dosage de la créatinine sérique avec calcul du Débit de Filtration Glomérulaire (DFG) selon l'équation du Chronic Kidney Disease Epidemiology Collaboration 2009 (CKD EPI) car plus performante [Levey 2015] et la numération formule sanguine. Le bilan est complété par d'autres analyses en cas de cancer métastatique. L'objectif de ce bilan est d'évaluer la fonction rénale, de déterminer les facteurs pronostiques de l'International Metastatic Renal Cell Carcinoma Database (IMDC) [Heng 2009], de dépister un éventuel syn-

drome paranéoplasique : anémie, polyglobulie, hypercalcémie, cholestase, de monitorer et dépister les contre-indications à l'immunothérapie [Heng 2009, Sufrin 1989].

La classification histologique du CCR suit la dernière révision de l'Organisation Mondiale de la Santé (OMS) 2022, regroupant les tumeurs malignes en six catégories : CCR, carcinomes papillaires, tumeurs oncocytaires et chromophobes, carcinomes des tubes collecteurs, tumeurs avec altérations moléculaires et nouvelles entités [Moch 2022]. Les facteurs histopronostiques incluent le sous-type histologique, le grade nucléolaire de la société internationale d'uro-pathologie (ISUP en anglais), la présence de nécrose extensive, le stade pTNM décrit en annexe G, la présence d'emboles vasculaires et surtout la composante sarcomatoïde ou rhabdoïde [Amin 2017, Moch 2022, Delahunt 2013, Tan 2013].

Le réseau national Cancers rares du rein de l'adulte (Carare), labellisé par l'Institut National du Cancer (INCa), propose une relecture histopathologique par un groupe d'experts en tumeurs rénales et organise depuis 2022 des Réunions de Concertation Pluridisciplinaire (RCP) pour la prise en charge des cas complexes, tant sur le plan chirurgical que oncologique.

Le diagnostic du cancer du rein repose sur une combinaison d'imagerie multimodale et d'analyse histopathologique, visant à préciser la nature tumorale, le stade évolutif et le profil pronostique afin d'optimiser la prise en charge thérapeutique.

2.1.3 Traitements du carcinome à cellules rénales

La prise en charge thérapeutique du CCR dépend étroitement des caractéristiques morphologiques de la tumeur, en particulier de sa taille. La liste des traitements présentée ci-dessous n'est pas exhaustive mais résume les principales recommandations établies par l'Association Française d'Urologie (AFU) et l'European Association of Urology (EAU).

2.1.3.1 Traitement des petites tumeurs du rein

Les alternatives thérapeutiques disponibles pour la prise en charge des petites tumeurs rénales, définies par des masses tumorales inférieures ou égales à quatre centimètres de diamètre, reposent sur plusieurs approches. La NP constitue la stratégie chirurgicale de référence, offrant un contrôle oncologique optimal tout en préservant la fonction rénale et en minimisant les complications périopératoires. Cette intervention est détaillée en section 2.1.4.

La NP permet d'obtenir un contrôle oncologique équivalent à celui de la Néphrectomie Totale (NT) [Talenfeld 2018, Kunath 2017, Van Poppel 2011]. La survie sans récidive à cinq ans après une NP pour une tumeur de stade T1a (cf. annexe G) excède 95 % [Talenfeld 2018, Jamil 2020]. La présence d'une marge chirurgicale positive est associée à un risque accru de récidive locale, sans pour autant impacter significativement la survie spécifique [Shah 2016, Bensalah 2010, Bernhard 2010]. La totalisation immédiate par néphrectomie n'est généralement pas requise [Bensalah 2010, Wood 2018], bien que dans le cas de tumeurs de haut grade

et/ou de marges positives extensives, une discussion en RCP puisse recommander cette approche. Le traitement des récidives locales après NP reste peu étudié mais plusieurs options thérapeutiques sont envisageables, incluant la néphrectomie élargie, une reprise de NP ou une ablation percutanée en fonction du délai de récidive, du type histologique et du grade tumoral [Wood 2018, Kriegmair 2018]. L'étude Re-Part, conduite par le CCAFU et publiée en 2022, souligne l'intérêt des Techniques Ablatives (TA) percutanées pour le traitement des récidives locales post-NP, en mettant en avant une meilleure préservation de la fonction rénale et une morbidité réduite par rapport au traitement chirurgical de rattrapage [Brassier 2022].

La NT ne constitue pas une stratégie de première intention dans la gestion des petites tumeurs rénales. Elle peut néanmoins être indiquée en présence d'un rein non fonctionnel, d'une suspicion de tumeur classée T3a (cf. annexe G) ou d'une complexité chirurgicale importante. En cas de tumeur jugée particulièrement complexe, une évaluation par un centre expert est recommandée avant toute indication de NT.

Les différentes voies d'abord en NT affichent des résultats oncologiques comparables [Hemal 2007, Auffenberg 2020]. La NT cœlioscopique exploite les bénéfices de la chirurgie mini-invasive, section 1.2.4, tels qu'une hospitalisation plus courte, une réduction de la consommation d'antalgiques et une diminution du saignement peropératoire [Ebbing 2015]. Les approches transpéritonéale et rétropéritonéale présentent des performances similaires en termes de contrôle oncologique et de qualité de vie [Desai 2005]. Peu d'études ont comparé la chirurgie robot-assistée et la cœlioscopie standard, bien que les résultats oncologiques paraissent équivalents [Hemal 2009].

Les TA constituent une alternative thérapeutique pour le traitement des petites tumeurs rénales. Une biopsie diagnostique préalable est recommandée avant leur mise en œuvre. Parmi ces techniques, il est possible de citer la radiofréquence, les micro-ondes, la cryothérapie, ainsi que des approches plus récentes telles que l'électroporation irréversible ou la radiothérapie stéréotaxique.

Le TA peut être proposé lorsque la destruction tumorale complète est réalisable. Les résultats oncologiques et fonctionnels à long terme sont satisfaisants [Talenfeld 2018, Andrews 2019, Uhlig 2019, Larcher 2015]. Cependant, en l'absence d'essais randomisés comparant TA et NP, les données disponibles sont issues d'études rétrospectives sujettes à divers biais. Le contrôle oncologique local du TA semble inférieur à celui de la NP [Talenfeld 2018, Uhlig 2019, Andrews 2019, Thompson 2015, Abu-Ghanem 2020], bien que certaines études suggèrent une meilleure préservation de la fonction rénale [Pantelidou 2016, Patel 2017]. De plus, les taux de complications globales et majeures apparaissent moindre avec le TA comparativement à la NP [Yano 2018, Ahmed 2014, Krokidis 2017].

La radiothérapie stéréotaxique ablative représente une option thérapeutique non invasive particulièrement intéressante chez les patients inopérables ou à haut risque chirurgical, atteints de tumeurs T1a et T1b (cf. annexe G) [Ali 2022, Correa 2019, Bentahila 2024]. Toutefois, des études prospectives avec un suivi prolongé sont nécessaires pour consolider son indication en situation élective.

La Surveillance Active (SA) consiste en une surveillance clinique et radiologique régulière des patients atteints d'une petite tumeur rénale. Une intervention est envisagée en cas de progression tumorale ou d'apparition de symptômes. Cette approche se distingue de l'abstention thérapeutique, qui est réservée aux patients à espérance de vie limitée, dans un objectif purement palliatif. Les études rapportent une croissance tumorale majoritairement linéaire, avec une augmentation de volume oscillant entre 0,09 et 0,37 centimètre par an [Finelli 2020, Alam 2019, Abou Youssif 2007, Crispen 2009, Jewett 2011, Mir 2018, Petros 2019, Uzosike 2018, Klatte 2021]. Aucune différence significative de croissance n'a été observée entre les tumeurs bénignes et malignes [Jewett 2011].

En matière de qualité de vie, la SA est associée à un état de santé physique moindre, ce qui s'explique par l'inclusion majoritaire de patients âgés et comorbides, ainsi qu'à un niveau d'anxiété accru par rapport aux patients ayant bénéficié d'une prise en charge thérapeutique immédiate [Parker 2013, Goldberg 2020, Pierorazio 2013].

Le choix thérapeutique dépend principalement de l'espérance de vie et de la taille tumorale. Chez les patients à espérance de vie limitée, une surveillance clinique seule peut être privilégiée. Pour les tumeurs de moins de deux centimètres, la SA avec un traitement différé si nécessaire constitue une approche recommandée, notamment chez les patients âgés et polymorbides. Au-delà de trois centimètres, la NP demeure le traitement de référence.

2.1.3.2 Traitement des tumeurs du rein localisées supérieures à quatre centimètres

Les tumeurs rénales localisées supérieures à quatre centimètres nécessitent une prise en charge optimisée afin de garantir un contrôle oncologique efficace tout en préservant la fonction rénale lorsque cela est possible.

Pour les tumeurs T1b (cf. annexe G), la NP est non inférieure à la NT en termes de survie spécifique [Kunath 2017, Lee 2014, Mir 2017, Borghesi 2013, Capitanio 2015, Scosyrev 2014a, MacLennan 2012]. Une étude prospective unique a démontré une efficacité comparable entre la NP et la NT pour des tumeurs de moins de cinq centimètres [Van Poppel 2011].

Concernant les tumeurs T2 (cf. annexe G), les données disponibles sont principalement issues d'études rétrospectives, suggérant un avantage en survie pour la NP, malgré des pertes sanguines plus importantes et des complications accrues [Mir 2017, Janssen 2018, Klett 2021, Huang 2021]. Une étude multicentrique récente a mis en évidence un bénéfice en termes de survie spécifique et globale en faveur de la NP par rapport à la NT [Janssen 2018].

Pour les tumeurs T3a (cf. annexe G) biopsiées, les résultats des études rétrospectives sont discordants. Deux études suggèrent une équivalence oncologique entre NP et NT [Patel 2020b, Boulenger De Hauteclocque 2023], tandis qu'une troisième rapporte une survie sans récidive plus courte après NP [Shah 2017]. Une méta-analyse récente incluant 1278 patients en NP et 2113 en NT n'a pas démontré de différence significative de survie spécifique [Liu 2021a].

L'approche en chirurgie robot-assistée semble réduire la morbidité périopératoire de la NP pour les tumeurs supérieures à quatre centimètres. Elle est associée à une diminution des pertes sanguines, des complications et de la durée d'hospitalisation par rapport à la chirurgie ouverte [Wang 2017, Garisto 2018], ainsi qu'à une réduction du taux de conversion, de l'ischémie chaude et de l'altération de la fonction rénale par rapport à la cœlioscopie standard [Choi 2015, Hinata 2021]. La morbidité périopératoire demeure néanmoins comparable à celle de la NT dans les centres experts [Bradshaw 2020, Schiavina 2020, Wang 2017, Garisto 2018].

Les revues systématiques et une étude prospective non randomisée ont montré des résultats oncologiques équivalents entre NT cœlioscopique et NT ouverte, y compris pour les tumeurs localement avancées [Steinberg 2004, Hemal 2007, Ebbing 2015]. La NT cœlioscopique se distingue par une morbidité réduite et une hospitalisation plus courte [Hemal 2007, Gratzke 2009, MacLennan 2012].

Une méta-analyse récente incluant 1832 patients a démontré la non-infériorité de la NT robot-assistée par rapport à la cœlioscopie standard en termes de résultats périopératoires et de complications [Li 2020a]. Toutefois, son coût reste significativement plus élevé [Jeong 2017].

Une étude multicentrique observationnelle n'a pas révélé de différence significative entre NT transpéritonéale et rétropéritonéale en termes de complications postopératoires, marges chirurgicales positives et résultats fonctionnels. Cependant, la voie transpéritonéale est associée à une durée opératoire plus courte et à un moindre taux de complications peropératoires [Porpiglia 2021].

La localisation tumorale au pôle supérieur n'est pas un facteur prédictif d'envahissement de la surrénale. L'absence de différence en termes de survie globale à dix ans [Lane 2009] justifie l'indication restrictive de la surrénalectomie, limitée aux cas d'envahissement tumoral direct ou de lésion suspecte à l'imagerie préopératoire [Lane 2009, Bekema 2013].

L'intérêt du curage ganglionnaire associé à la NT demeure controversé [Bekema 2013, Herrlinger 1991, Capitanio 2014, John 2019]. Une large étude multicentrique rétrospective (2722 patients) n'a pas montré d'amélioration de la survie après lymphadénectomie [Gershman 2018], confirmée par un essai randomisé de l'EORTC pour les patients N0 (cf. annexe G) [Blom 2009]. Cependant, une méta-analyse incluant 1983 patients suggère un bénéfice en survie spécifique pour les patients à haut risque sans impact significatif sur la morbidité [Capitanio 2014].

L'embolisation préopératoire de l'artère rénale ne confère aucun bénéfice en termes de résultats oncologiques [May 2009]. Elle peut être envisagée en cas de symptômes (hématurie, douleurs) chez des patients non éligibles à la chirurgie [Maxwell 2007].

Les TA pour les tumeurs de plus de quatre centimètres font l'objet d'études rétrospectives aux résultats hétérogènes, avec un suivi limité [Cazalas 2023, Chang 2015, Shapiro 2020, Hasegawa 2018, Yanagisawa 2022, Atwell 2007, Schmit 2010]. Une analyse récente de 5763 patients issus de la base SEER a rapporté un risque de mortalité spécifique multiplié par 2,5 pour les tumeurs T1b (cf. annexe G) traitées par TA en comparaison à la NP [Pecoraro 2019].

La radiothérapie stéréotaxique ablative est une option envisageable chez les patients inopérables nécessitant un contrôle tumoral local. Une étude rétrospective de l'International Radiosurgery Oncology Consortium for Kidney (IROCK) rapporte une survie sans progression et une survie spécifique à deux ans de 81 % et 96,1 %, respectivement, sans altération significative de la fonction rénale et avec une toxicité limitée [Siva 2020].

Si elle est techniquement réalisable, la NP reste le traitement recommandé pour les tumeurs T1b (cf. annexe G). La NT est préconisée pour les tumeurs T1b et T2 (cf. annexe G) lorsque la NP n'est pas envisageable ou en cas de rein non fonctionnel.

2.1.3.3 Traitement d'un cancer du rein localement avancé

La néphrectomie élargie constitue le traitement de référence dans une visée curative. L'objectif principal, indépendamment de la technique utilisée, est d'obtenir des marges chirurgicales négatives [Zini 2008]. La voie d'abord standard pour les tumeurs localement avancées reste la CO. Toutefois, des séries récentes rapportent la faisabilité d'une approche cœlioscopique ou robot-assistée, avec des résultats périopératoires favorables chez des patients sélectionnés, sous réserve de garantir une exérèse en marges saines [Bragayrac 2016, Verhoest 2016, Becher 2020].

Un curage ganglionnaire est indiqué en présence d'adénopathies cliniquement palpables ou détectées en imagerie préopératoire [Bhindi 2018, Correa 2021]. L'exérèse chirurgicale de ces adénopathies pourrait avoir un intérêt à la fois pronostique et thérapeutique dans une situation oligométastatique [Srivastava 2020, Zaid 2017]. En l'absence d'atteinte ganglionnaire identifiée, le bénéfice oncologique du curage ganglionnaire systématique reste indéterminé. En cas d'atteinte ganglionnaire avérée, la survie à cinq ans après néphrectomie varie entre 10 et 45 % [Correa 2021, Gershman 2017].

L'administration d'un traitement médical en néo-adjuvant entraîne une réponse tumorale dans 9 à 28 % des cas avec les thérapies ciblées anti-angiogéniques et jusqu'à 35 % de réduction volumique avec l'immunothérapie [Abel 2011, Karam 2014, Lebacle 2019, Zhu 2024, Gorin 2022, Albiges 2022, Carlo 2022]. En présence d'un thrombus cave tumoral, un traitement systémique préopératoire permet une réduction de la hauteur du thrombus allant jusqu'à 35 % [Gu 2024, Stewart 2015]. Toutefois, la stratégie néo-adjuvante ne peut être envisagée que dans le cadre d'essais cliniques ou après discussion en RCP, en raison du faible niveau de preuve actuel et de l'absence de bénéfice oncologique démontré [Westerman 2020, Leow 2023].

L'envahissement des organes adjacents est difficile à prédire [Tanaka 2021]. Certains patients classés T4 (cf. annexe G) sont en réalité sur-stadifiés et peuvent bénéficier d'une résection chirurgicale. La survie à ce stade demeure limitée (inférieure à 10 % à cinq ans) et l'obtention de marges chirurgicales négatives constitue un facteur pronostique essentiel [Borregales 2016, Capitanio 2009, Margulis 2007].

La chirurgie d'exérèse reste le traitement de référence des tumeurs rénales avec thrombus, quel que soit le niveau d'extension du thrombus. Cette intervention est réalisable sous réserve que la tumeur et le thrombus soient techniquement résécables avec des marges saines et en l'absence de métastases avérées (M0, cf. annexe G) [Bigot 2022, Reese 2013]. Toutefois, elle est associée à une mortalité à 90 jours pouvant atteindre 20 %, avec un taux de complications proportionnel au niveau du thrombus [Agochukwu 2014].

Le bilan préopératoire doit être adapté à la complexité attendue de la chirurgie, notamment en fonction du type de clampage de la Veine Cave Inférieure (VCI) dépend de l'extension tumorale, comme illustrée en figure 2.10. Pour les thrombus de niveau I, un clampage partiel ou total mais bref de la VCI juxta-rénale est suffisant. En revanche, les thrombus de niveau III ou IV nécessitent un clampage du hile hépatique et des veines sus-hépatiques, voire une chirurgie à cœur battant avec circulation extracorporelle pour certains thrombus de niveau IV.

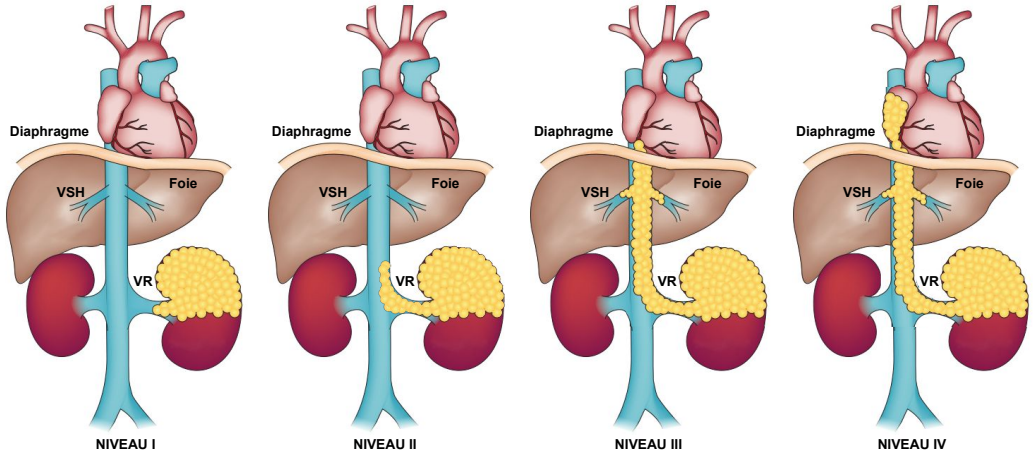


FIGURE 2.10 – (de gauche à droite) Thrombus de niveau I situé dans Veine Rénale (VR). Thrombus de niveau II situé dans la Veine Cave Inférieure (VCI) infra-hépatique. Thrombus de niveau III remontant jusqu'au hile hépatique et aux Veines Sus-Hépatiques (VSH). Thrombus de niveau IV avec extension dans l'oreillette droite. Source : générée par l'auteur à partir de [Ludwig 2017].

Le choix de la technique chirurgicale est dicté en premier lieu par l'extension supérieure du thrombus et en second lieu par le risque d'envahissement pariétal de la VCI. La voie ouverte est la recommandation standard pour la chirurgie du thrombus cave. L'approche mini-invasive robot-assistée a émergé ces dernières an-

nées, démontrant une diminution du risque transfusionnel et une réduction des complications globales, sans impact significatif sur l'incidence des complications majeures [Garg 2022, Shi 2020, Abaza 2016, Gill 2015].

2.1.4 Néphrectomie partielle

Comme mentionné précédemment en section 2.1.3.1, la NP constitue le traitement chirurgical de référence des tumeurs rénales classées T1a (cf. annexe G). Son indication peut être étendue aux tumeurs T1b et T2 (cf. annexe G) lorsque l'intervention est techniquement réalisable (cf. section 2.1.3.2).

En France, la NP est nettement prédominante par rapport à la NT. Elle représente 85,6 % des actes chirurgicaux de néphrectomie, soit plus de 5 300 interventions en 2023 (voir tableau 2.1). Parmi celles-ci, plus de 95,8 % sont réalisées en CMI suivant les données de la Classification Commune des Actes Médicaux (CCAM) résumées en annexe H. Comme illustré en figure 2.11, la répartition géographique des actes de NP est toutefois hétérogène sur le territoire métropolitain. Les établissements hospitaliers situés à Paris et dans le département de la Gironde concentrent le plus grand nombre d'actes, avec respectivement 8,6 % et 7,7 % des NP nationales.

	CO	CMI	Total
NP	223	5077	5300 (85,6 %)
NT	402	490	892 (14,4 %)
Total	625 (10,1 %)	5567 (89,9 %)	6192

TABLE 2.1 – Répartition des actes de Néphrectomie Partielle (NP) et Totale (NT) réalisés en France en 2023 selon le type de chirurgie pratiquée : Ouverte (CO) ou Mini-Invasive (CMI). Source : générée par l'auteur à partir de la base de données Open CCAM 2023.

D'un point de vue clinique, la NP favorise une préservation optimale du parenchyme rénal sain, limitant ainsi la détérioration post-chirurgicale de la fonction rénale [Scosyrev 2014b, Volpe 2015, Lane 2011, Zhang 2016]. Cette préservation fonctionnelle réduit le risque d'événements cardiovasculaires et pourrait améliorer la survie globale en comparaison à la NT [Capitanio 2015, Miller 2008, Huang 2006, Huang 2020]. De plus, la conservation néphronique est associée à une meilleure qualité de vie et à un bénéfice psychologique significatif [Gratzke 2009]. En revanche, les techniques d'énucléation tumorale minimisant la marge de tissu sain résiduel ne semblent pas augmenter le taux de récidive locale. Le bénéfice en termes de préservation rénale reste incertain encore aujourd'hui [Minervini 2011, Cao 2017, Bertolo 2023].

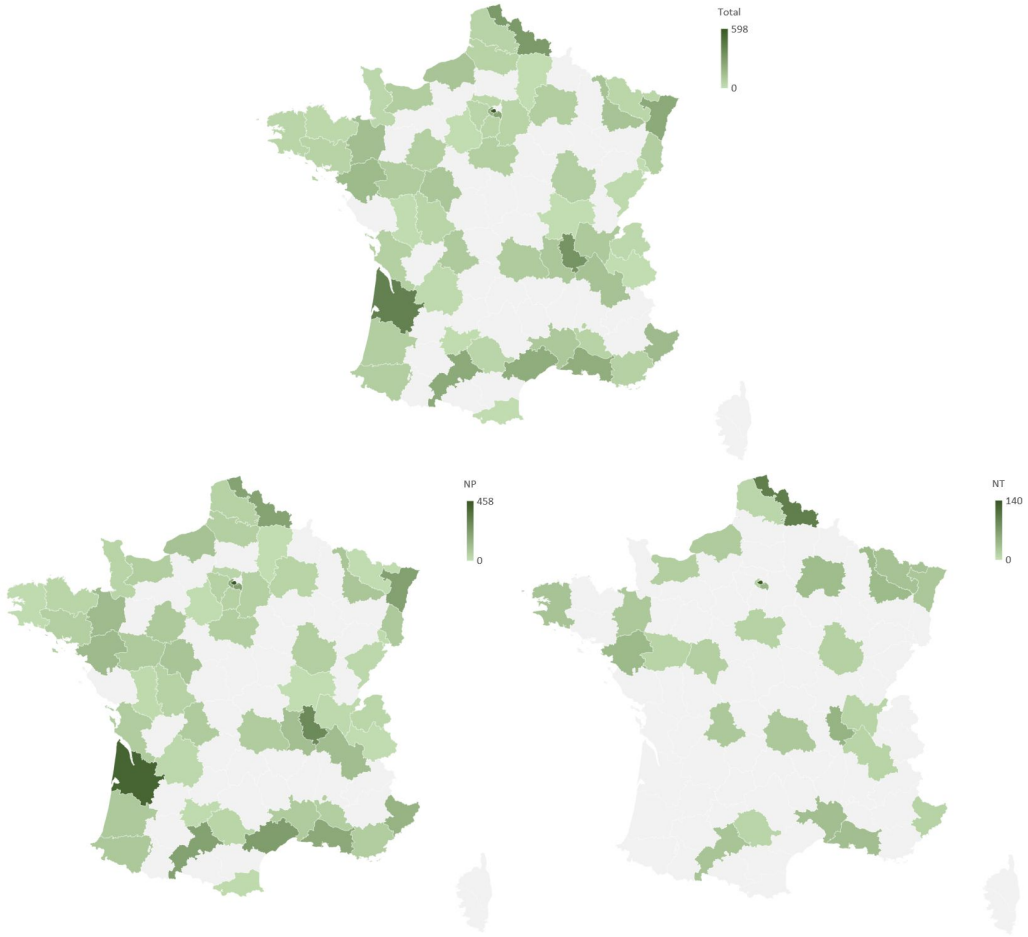


FIGURE 2.11 – Cartographie de la répartition en France métropolitaine des actes de néphrectomie. En haut, toutes les néphrectomies. A gauche, les Néphrectomies Partielles (NP). A droite, les Néphrectomies Totale (NT). Source : générée par l'auteur à partir de la base de données Open CCAM 2023.

Toutefois, la NP présente une morbidité significative, estimée à environ 20 % selon la littérature. Les deux complications les plus redoutées sont les saignements secondaires dues à un faux-anévrysme et les fistules urinaires [Potretzke 2016]. Les complications graves peuvent atteindre un taux de 11 %. Leur incidence dépend de plusieurs facteurs, incluant la complexité tumorale, l'âge et les comorbidités du patient, l'expertise chirurgicale, le volume d'activité du centre et la voie d'abord choisie [Simhan 2011, Khene 2020, Khene 2022]. L'étude prospective française RoPAN, publiée en 2023, précise que près de 10 % des patients ayant bénéficié d'une NPRA développent des complications tardives au cours de la première année, principalement d'ordre vasculaire, infectieux ou pariétal. Elles apparaissent majoritairement dans les trois mois suivant l'intervention [Margue 2023].

La survie spécifique et la survie globale après NP ne diffèrent pas selon la voie chirurgicale employée [Gill 2007, Chang 2018, Bravi 2021]. Comparée à la voie ouverte, la NP cœlioscopique présente des avantages en matière de réduction des pertes sanguines et de durée d'hospitalisation (cf. section 1.2.4.2), malgré un allongement significatif des durées opératoire et d'ischémie. Les taux de complications postopératoires et la fonction rénale à long terme restent toutefois similaires [Gill 2007, Bravi 2021, Guglielmetti 2022].

La NPRA, par rapport à la voie ouverte, améliore significativement les résultats en réduisant les pertes sanguines, les durées d'hospitalisation, d'ischémie et le taux global de complications postopératoires. Le taux de marges positives ainsi que la fonction rénale résiduelle sont cependant équivalents entre ces deux approches [Masson-Lecomte 2013, Ni 2022, Peyronnet 2016, Bravi 2021]. Cette réduction de la morbidité permet désormais, pour certains patients sélectionnés, une prise en charge ambulatoire [Bernhard 2022, Bernhard 2023].

En comparaison avec la voie cœlioscopique standard, la NPRA offre des bénéfices clairs concernant la durée d'ischémie ainsi qu'un moindre taux de conversion en CO ou en NT, particulièrement dans le cas des tumeurs complexes ou hilaires [Choi 2015, Chang 2018, Bravi 2021, Long 2012].

Une étude prospective multicentrique observationnelle n'a pas démontré de différence significative entre la NP transpéritonéale et rétropéritonéale en termes de complications postopératoires, de marges chirurgicales positives ou de résultats fonctionnels évalués par le DFG. Toutefois, la voie transpéritonéale s'est avérée supérieure par une durée opératoire plus courte et une incidence moindre de complications peropératoires [Porpiglia 2021].

L'impact de la durée d'ischémie sur la fonction rénale postopératoire reste débattu, en particulier chez les patients souffrant d'une insuffisance rénale préexistante [Huang 2020]. Bien que les données disponibles soient contradictoires, il est recommandé de limiter le temps de clampage du pédicule rénal (cf. section 2.1.1.3) à moins de 25 minutes [Volpe 2015, Mir 2015, Zhang 2016]. Les techniques sans clampage ou avec clampage sélectif réduisent certes la dégradation précoce de la fonction rénale mais n'ont pas prouvé de bénéfice à long terme par rapport au clampage total inférieur à 25 minutes [Rais-Bahrami 2012, Greco 2019, Dong 2018, Cignoli 2023, Fong 2024, Mellouki 2021].

En France, une étude prospective réalisée en 2010 au sein d'institutions hospitalières publiques, universitaires et privées a mis en évidence un effet centre pour la NP, soulignant une morbidité et un taux de marges positives accrus dans les établissements à faible activité chirurgicale [Couapel 2014]. Une analyse du CCAFU a également démontré que les centres à haut volume présentaient de meilleurs résultats en matière de NPRA [Peyronnet 2018].

Sur le plan technique, la NPRA est aujourd'hui très largement pratiquée en France en raison de ses bénéfices cliniques démontrés et l'accessibilité croissante aux systèmes de télé-chirurgie. Durant l'intervention, le patient est installé en décubitus latéral avec une flexion maximale de la table opératoire afin d'accroître la distance entre la dernière côte flottante et la crête iliaque [Hoznek 2007]. La préparation cutanée du patient et la création du pneumopéritoine suivent les étapes décrites précédemment en section 1.2.4.2. Cinq à six trocarts sont ensuite positionnés selon la configuration illustrée en figure 2.12, afin d'assurer l'introduction optimale des instruments chirurgicaux et de la caméra endoscopique, comme illustré en figure 2.13. En NPRA, la distance de travail entre l'optique et les tissus est plus faible qu'en cœlioscopique. Cette proximité offre au chirurgien une vision magnifiée, améliorant la précision des gestes de dissection tumorale mais limitant le champ visuel à des vues partielles de l'organe.

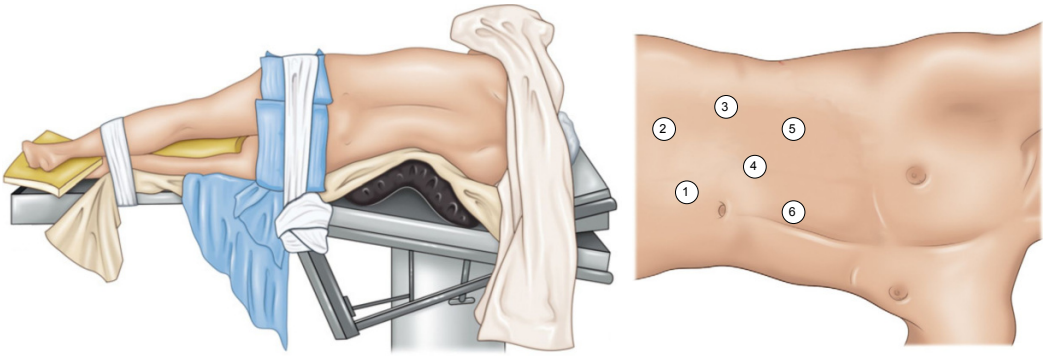


FIGURE 2.12 – (à gauche) Positionnement typique du patient en décubitus latéral pour une Néphrectomie Partielle (NP) cœlioscopique. (à droite) Positionnement typique des trocarts pour une Néphrectomie Partielle Robot-Assistée (NPRA) par voie transpéritonéale. Source : générée par l'auteur à partir de [Hoznek 2007].

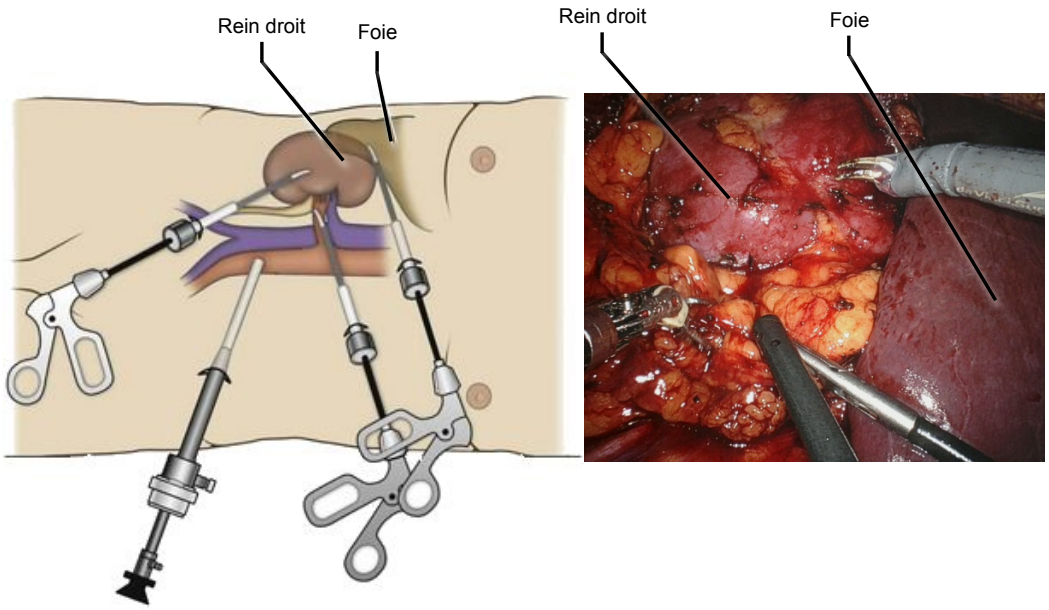


FIGURE 2.13 – (à gauche) Représentation schématique du champ de vue peropératoire lors d'une Néphrectomie Partielle Robot-Assistée (NPRA). (à droite) Image endoscopique lors d'une NPRA après exposition du rein. Source : générée par l'auteur à partir de [Bishoff 2007].

La NP comprend généralement entre quatre et seize phases opératoires selon les descriptions de la littérature. Cinq étapes chirurgicales clés ont été retenues dans ce manuscrit [Farinha 2023, De Backer 2023], correspondant à la figure 2.14 :

1. L'exposition du rein : mobilisation chirurgicale du rein par ouverture et dissection des attaches péritonéales ou rétropéritonéales, permettant un accès anatomique clair à la tumeur et au pédicule rénal, si besoin.
2. La dissection vasculaire : isolement et contrôle des pédicules vasculaires rénaux (artère et veine principales et leurs branches segmentaires éventuelles). Cette étape conditionne la qualité du clampage vasculaire ultérieur. Cette étape n'est pas réalisée systématiquement.
3. L'identification et la délimitation de la tumeur : repérage précis de la tumeur, assisté si nécessaire par échographie peropératoire et marquage des marges chirurgicales avec l'électrocoagulation monopolaire ou bipolaire.
4. La résection tumorale : ablation de la tumeur par incision du parenchyme rénal selon les marges prédéfinies, en minimisant la perte de tissu rénal sain pour préserver la fonction rénale postopératoire.
5. La reconstruction du lit tumoral : fermeture chirurgicale du parenchyme rénal résiduel par suture profonde à visée hémostatique et urinostatique. L'application de produits hémostatiques (colles biologiques) complète généralement cette étape afin de réduire les complications.

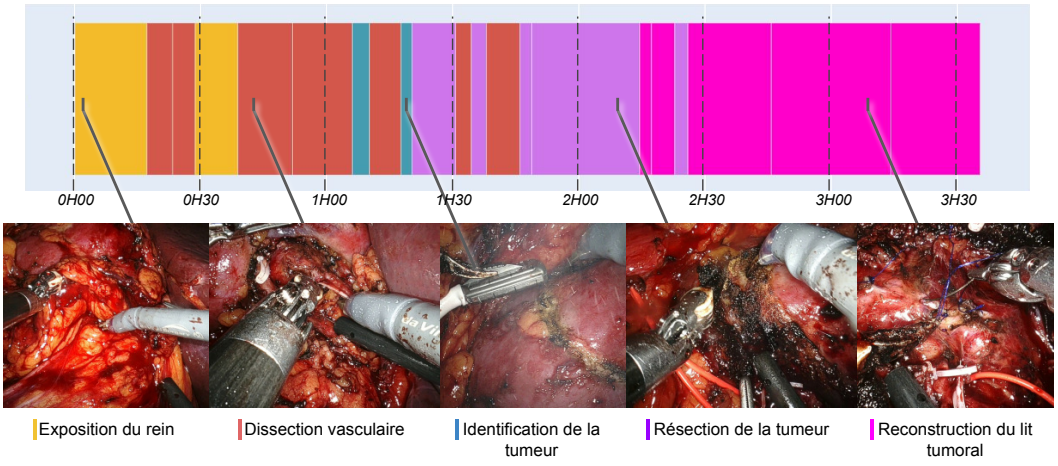


FIGURE 2.14 – Représentation schématique sous la forme d'un chronogramme des différentes phases chirurgicales lors d'une Néphrectomie Partielle Robot-Assistée (NPRA). Source : auteur.

Chacune des phases chirurgicales peut être caractérisée précisément par les structures anatomiques exposées ainsi que par l'instrumentation chirurgicale spécifique employée. Ces éléments sont détaillés en annexe I.

2.2 Guidage par réalité augmentée en chirurgie du rein

Dans le contexte de la chirurgie du rein et plus particulièrement de la NP guidée par l'image, nous nous proposons d'explorer scientifiquement et techniquement le guidage par RA, de la levée des verrous scientifiques à la validation expérimentale. In fine, nous cherchons à enrichir la perception visuelle du chirurgien en superposant directement sur les images endoscopiques, les structures internes et adjacentes du rein afin d'améliorer la précision et la sécurité du geste chirurgical.

2.2.1 Généralités

Nos travaux portent spécifiquement sur l'augmentation peropératoire des images endoscopiques avec des informations tridimensionnelles préopératoires issues du jumeau numérique du patient. Ces modèles anatomiques personnalisés permettent de visualiser en peropératoire les structures d'intérêts, notamment les lésions tumorales à réséquer ainsi que les éléments vasculaires et urinaires à préserver. Les images augmentées peuvent également intégrer des données dérivées du jumeau numérique, comme les zones de perfusion artérielle illustrées en figure 2.15 [Amparore 2023b, Amparore 2023a] ou un plan de résection chirurgical.

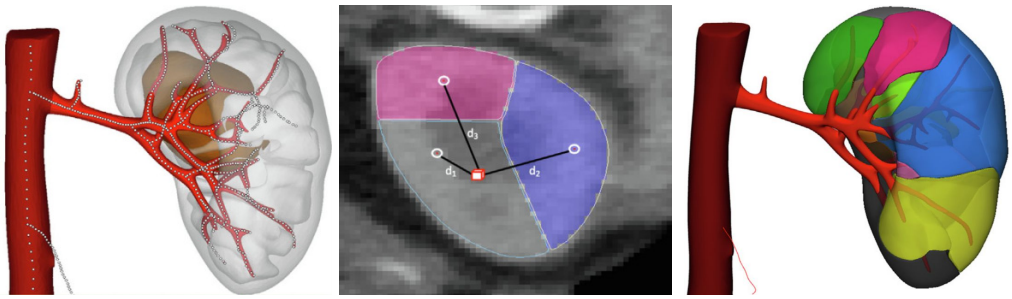


FIGURE 2.15 – Calcule des zones de perfusion artérielle du rein à partir du jumeau numérique du patient via un diagramme de Voronoï. (à gauche) Détection des lignes centrales des branches artérielles et discrétisation de celles-ci pour une utilisation en tant que graines. (au centre) Regroupement des voxels du parenchyme rénal en fonction de leur distance euclidienne aux graines. (à droite) Visualisation des zones de perfusion. Source : générée par l'auteur à partir de [Amparore 2023a].

La reconstruction tridimensionnelle de ces modèles est réalisée par le chirurgien et/ou le radiologue à partir des données d'imagerie en coupe, TDM ou IRM (cf. section 1.3.1.1), au cours de la phase de planification chirurgicale préopératoire, comme explicité en section 1.3.1. Cette étape aboutit à la génération de maillages surfaciques de chacune des structures à augmenter.

Le jour de l'intervention, avant l'entrée du patient en salle d'opération, le système de RA est mis en place. Il se compose au minimum d'une unité informatique de haute performance, équipée d'un processeur graphique dédié ou Graphics Processing Unit

(GPU) en Anglais, capable d'effectuer en quasi-temps réel à la fois du traitement d'images complexes et le rendu graphique requis pour l'augmentation visuelle des images endoscopiques. Le système de RA est connecté à la colonne de cœlioscopie ou au système robotique de télé-chirurgie par une liaison filaire assurant la captation du flux vidéo provenant de la caméra endoscopique, comme explicité en section 2.2.5. La génération des augmentations visuelles s'appuie essentiellement sur des méthodes de vision par ordinateur, qui nécessitent typiquement un étalonnage géométrique de la caméra utilisée, selon les méthodes décrites dans la littérature scientifique présentée en section 2.2.6.

Une fois le patient installé (cf. section 2.1.4), anesthésié et le pneumopéritoine établi par insufflation de dioxyde de carbone (CO_2) (cf. section 1.2.4.2), les instruments chirurgicaux sont introduits dans la cavité abdominale et le rein libéré de sa graisse péri rénale (cf. section 2.1.4). A ce stade de l'intervention, le jumeau numérique du patient est aligné sur les images endoscopiques par diverses approches décrites en section 2.2.4. Ce recalage initial 3D-2D constitue une étape technique importante qui permet la mise en correspondance spatiale des entités virtuelles avec la scène chirurgicale réelle, comme illustré en figure 2.16. Lorsque ce recalage initial est jugé cliniquement conforme, les informations de positionnement du jumeau numérique peuvent être exploitées afin de réaliser un suivi dynamique de l'organe (contribution présentée au chapitre 5) et enrichir en temps réel le flux vidéo endoscopique avec des augmentations visuelles pertinentes.

Pour cela, l'étape de rendu graphique revêt une importance clinique majeure, car il doit permettre au chirurgien d'interpréter avec précision l'image augmentée en termes de localisation et d'interactions entre les structures adjacentes, tout en simulant une perception d'un volume malgré l'affichage d'une image plane. Cette représentation repose sur un pipeline graphique structuré et optimisé pour une intégration peropératoire en temps réel.

Les modèles anatomiques sont représentés sous forme de maillages tridimensionnels associés à des propriétés optiques définies (couleur, opacité), préchargés en mémoire GPU. Ces maillages subissent une transformation géométrique depuis leur espace de référence (espace modèle) vers l'espace de la caméra virtuelle, en exploitant les résultats du recalage initial ou du suivi dynamique. Cette transformation repose sur l'application successive des matrices de modèle, de vue et de projection.

Après la phase de rastérisation, un modèle d'éclairage simplifié, souvent basé sur l'approche de Blinn-Phong [Blinn 1977], est appliqué. Une source lumineuse virtuelle de type projecteur ("spot") colocalisée avec la caméra virtuelle est utilisée afin de simuler de manière réaliste les conditions d'éclairage intra-abdominales.

Enfin, le pipeline graphique, implémentable via des interfaces standardisées telles qu'OpenGL, intègre des étapes terminales essentielles comme le "depth peeling" utilisé pour un rendu stratifié des objets transparents superposés ou la composition pour assurer la fusion de l'image endoscopique en arrière-plan avec les modèles anatomiques en surimpression au premier plan. L'image résultante de l'ensemble des ces étapes techniques, de l'acquisition du flux vidéo au rendu graphique, est affichée par un dispositif prévu à cet effet.

Le principal avantage de la RA est donc de réaliser la fusion multimodale des données préopératoires sur la source vidéo peropératoire, sans impact important sur la charge mentale du chirurgien habituellement éprouvée par l'agrégation des informations et les efforts de mentalisation nécessaires pour cartographier et aligner spatialement les données d'imagerie en coupe (cf. section 1.3.3).

Un autre avantage de cette solution comparativement aux autres outils de guidage peropératoire (cf. section 1.3.2) est l'absence de dispositifs tiers dans le champ opératoire comme ceux nécessaires aux systèmes d'échographie peropératoire ou de navigation chirurgicale (sonde ou dispositif de localisation). Le chirurgien dispose sur une unique vue de l'ensemble des informations nécessaire à la réalisation du geste. Le bénéfice clinique potentiel pour le patient est alors important avec un temps total d'intervention réduit et une amélioration de la qualité des résections, car le chirurgien peut opérer en utilisant directement les images augmentées. Il est d'ailleurs intéressant de noter que 87 % des urologues apprécient le rôle de la RA pour le guidage peropératoire [Puliatti 2022]. Pour ces raisons, un système de RA doit garantir un recalage précis à tout moment lors de son utilisation.

De manière concrète, le guidage par RA en chirurgie rénale présente un potentiel impact clinique significatif, notamment lors des phases chirurgicales considérées comme complexes ou risquées. Quatre phases sont classiquement concernées : la dissection vasculaire, la délimitation de la tumeur, la résection tumorale et la reconstruction du lit tumoral (cf. section 2.1.4). Plus précisément quatre étapes sont ciblées pour la génération d'augmentations en RA : la dissection du hile rénal, le clampage vasculaire, la navigation du plan d'excision et la réparation du système collecteur et des lésions vasculaires [Joeres 2019]. Les défis et les besoins informationnels associés à ces phases sont synthétisés, de manière non exhaustive, dans le tableau 2.2.

Bien que plusieurs solutions de RA aient été proposées dans la littérature scientifique (cf. section 2.2.4) pour la chirurgie du rein et plus précisément pour la NP, elles partagent toutes certaines similitudes. Nous proposons en figure 2.17 un schéma technique illustrant le processus de réalisation de la NP assistée par RA, qui résulte de l'intégration et de l'enchaînement structuré des différentes étapes introduites précédemment.

Afin que ce processus technique de génération d'augmentations pour guider le geste chirurgical soit utilisable en clinique par le chirurgien, il requiert un effort très important d'ingénierie. Cet effort est d'autant plus important que la contrainte temps réelle est indispensable pour une utilisation peropératoire.

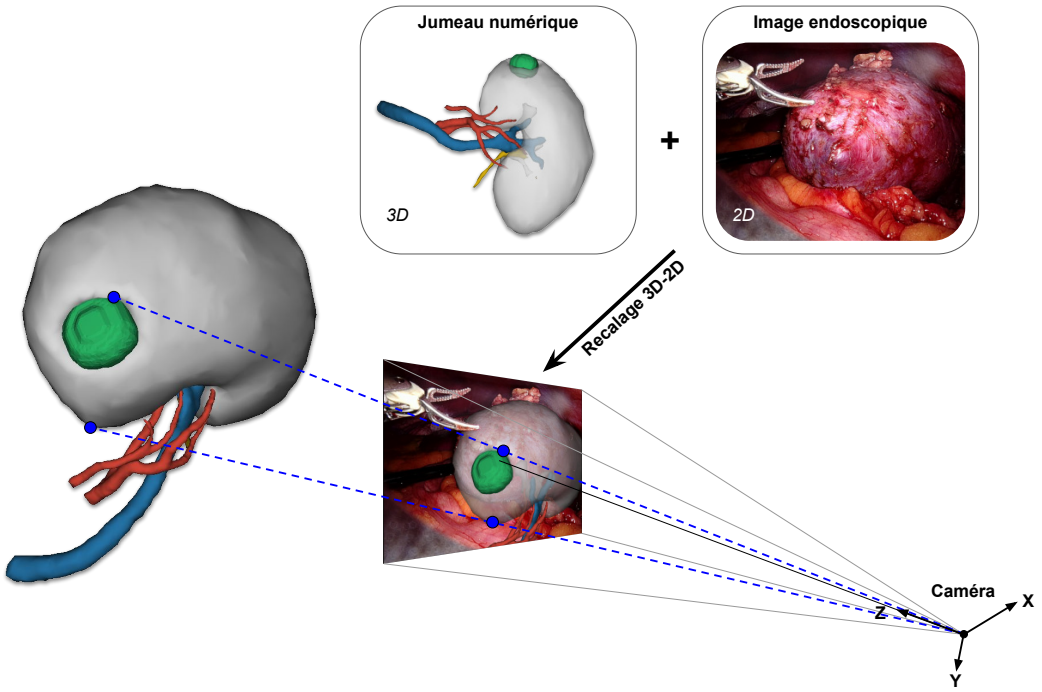


FIGURE 2.16 – Recalage 3D-2D du jumeau numérique du patient sur une image endoscopique de chirurgie du rein. Le modèle préopératoire doit être ajusté de manière à ce que sa projection sur le plan de l'image soit cohérente avec les informations de la vue caméra. Source : auteur.

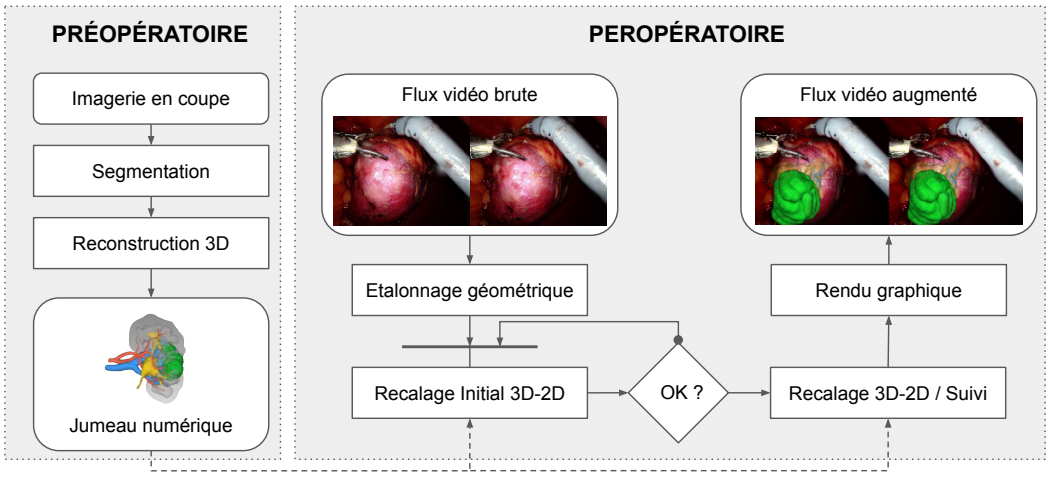


FIGURE 2.17 – Processus technique général pour la Néphrectomie Partielle (NP) assistée par Réalité Augmentée (RA). Source : auteur.

Phases chirurgicales	Défis / Challenges	Informations nécessaires
Dissection vasculaire	Le clampage est-il nécessaire et dans l'affirmative, quels sont les vaisseaux qui doivent être clampés ?	Informations sur la taille et la position de la tumeur, ainsi que sur les vaisseaux qui l'alimentent.
	Dissection hilaire en cas d'anatomie très variable d'un patient à l'autre.	Informations sur l'uretère, les principaux vaisseaux extrarénaux, l'arbre vasculaire rénal et la mise en évidence du pôle inférieur.
	Identifier, localiser et disséquer toutes les branches vasculaires pertinentes.	Disponibilité peropératoire des données d'imagerie préopératoire. Mise en évidence des vaisseaux occlus. Informations sur la proximité de l'instrument par rapport aux artères principales.
	Toutes les branches vasculaires concernées ont-elles été clampées ?	Informations sur la perfusion segmentaire. Confirmation que les clamps sont complètement fermés.
Délimitation de la tumeur	Localiser et naviguer jusqu'à la tumeur.	Disponibilité peropératoire des données d'imagerie préopératoire. Informations sur la position de la tumeur.
	Trouver le plan de résection idéal.	Disponibilité peropératoire des données d'imagerie préopératoire, parties endophytiques de la tumeur, profondeur de la tumeur, relation spatiale entre les outils et la tumeur, plan d'excision préopératoire.
Résection tumorale	Identifier, localiser et protéger les structures à risque (vaisseaux, système collecteur).	Information ou mise en évidence du parenchyme, des principaux vaisseaux alimentant la tumeur, du système collecteur. Disponibilité peropératoire des données d'imagerie préopératoire. Mise en évidence des principaux vaisseaux occlus.
	Préserver la perfusion du tissu rénal restant.	Informations sur la perfusion segmentaire.
	Détecter et gérer les lésions des structures à risque (vaisseaux, système collecteur).	Informations sur les lésions du système collecteur. Informations sur le tissu tumoral dans le plan de résection.
	Le clampage rétroactif est-il nécessaire ?	Informations sur la perfusion segmentaire.
	La résection a-t-elle été un succès sur le plan oncologique ?	Informations sur le type de tissu dans le lit de résection.
Reconstruction du lit tumoral	Appliquer un positionnement, une force et une distance corrects pour les sutures.	Informations sur les artères et les tissus qui peuvent se trouver sur le trajet de l'aiguille.
	Identifier, localiser et gérer les lésions du système collecteur.	Informations sur les lésions du système collecteur.
	Identifier, localiser et prendre en charge les lésions des vaisseaux principaux.	Informations sur les vaisseaux traversant la zone de résection.
	Prévenir et gérer les problèmes de visibilité dus à des saignements abondants.	Informations sur les principaux vaisseaux sanguins en peropératoire. Informations sur la source du saignement.
	Distinguer les vaisseaux qui nécessitent une suture individuelle de ceux qui n'en nécessitent pas.	Informations sur les artères. Quantification et visualisation de la gravité des saignements.

TABLE 2.2 – Défis et besoins d'information pour les phases chirurgicales les plus complexes et les plus difficiles lors de la Néphrectomie Partielle (NP) par voie mini-invasive. Source : généré par l'auteur à partir de [Joeres 2019].

2.2.2 Approche d'ingénierie logicielle

La mise en œuvre expérimentale des travaux de recherche en vision par ordinateur et RA nécessite un investissement temporel significatif en ingénierie logicielle afin de garantir la testabilité, la robustesse et la pertinence clinique des solutions développées pour une utilisation au bloc opératoire.

La méthodologie adoptée dans le cadre de ces travaux s'articule autour de deux niveaux de développement distincts mais complémentaires : la preuve de concept et le prototype fonctionnel.

La preuve de concept constitue une première étape dont l'objectif est de démontrer la faisabilité technique ainsi que l'intérêt scientifique et applicatif d'une idée innovante. Cette phase exploratoire est essentiellement réalisée en Python, un langage de programmation retenu pour sa flexibilité, sa simplicité d'utilisation, ainsi que pour son vaste écosystème de bibliothèques scientifiques qui accélèrent la mise en œuvre expérimentale des méthodes proposées. À ce stade, l'accent est mis sur la rapidité d'implémentation et la validation préliminaire des méthodes proposées, afin de permettre une valorisation scientifique et/ou industrielle par une publication ou un dépôt de brevet.

Le prototype fonctionnel représente une seconde phase, plus avancée, dans laquelle la solution est transférée et développée afin de permettre une évaluation approfondie en conditions précliniques ou directement en contexte clinique réel. Ce niveau de développement requiert un effort très important d'ingénierie logicielle. Il implique notamment une implémentation rigoureuse en langages C++ et QML, ainsi qu'une programmation optimisée exploitant les ressources Central Processing Unit (CPU) et GPU. Cet effort d'optimisation est indispensable pour assurer les performances compatibles avec une utilisation quasi temps réel au bloc opératoire.

Plus généralement, le travail d'ingénierie logicielle dépasse souvent le cadre purement scientifique, intégrant des tâches spécifiques de développement telles que la gestion avancée des flux vidéo, le rendu graphique ou encore l'interface utilisateur. Bien que ces éléments relèvent davantage du domaine du développement logiciel que de la recherche, ils sont essentiels à l'élaboration de solutions de RA robustes et testables.

Dans une optique de transfert efficace des contributions entre les équipes de recherche et de développement, ces travaux ont conduit à une restructuration des bibliothèques logicielles. L'intégration d'interfaces C++ et Python permet une interopérabilité accrue entre les prototypes et les solutions industrialisables, facilitant ainsi la maintenabilité et le transfert des algorithmes issues de la recherche à des applications cliniques opérationnelles.

Sur le plan méthodologique, l'approche Agile s'avère particulièrement adaptée aux exigences de ce projet. En effet, elle offre une grande flexibilité en favorisant une adaptation rapide aux évolutions des besoins clinico-techniques et facilite une collaboration étroite et continue entre cliniciens, chercheurs et développeurs. Contrairement au modèle traditionnel en V, caractérisé par une séquence linéaire rigide et une distinction nette entre les phases de vérification et validation (V&V), l'approche

Agile permet d'intégrer ces processus de manière itérative et incrémentale tout au long du cycle de recherche et développement.

Les contraintes spécifiques liées à l'intégration matérielle et logicielle dans l'environnement opératoire cible sont également prises en compte dès les premières phases du projet, garantissant ainsi une intégration fluide et cohérente lors du déploiement des prototypes en conditions réelles.

Ainsi, entre septembre 2022 et février 2025, huit versions successives du prototype de RA destinée à la NP ont été développées, d'abord de manière autonome, puis en étroite collaboration avec un ingénieur de recherche. Les principales fonctionnalités des différentes versions sont répertoriées à titre informatif en annexe J, tandis que deux exemples d'interfaces du prototype sont illustrés en figure 2.18.

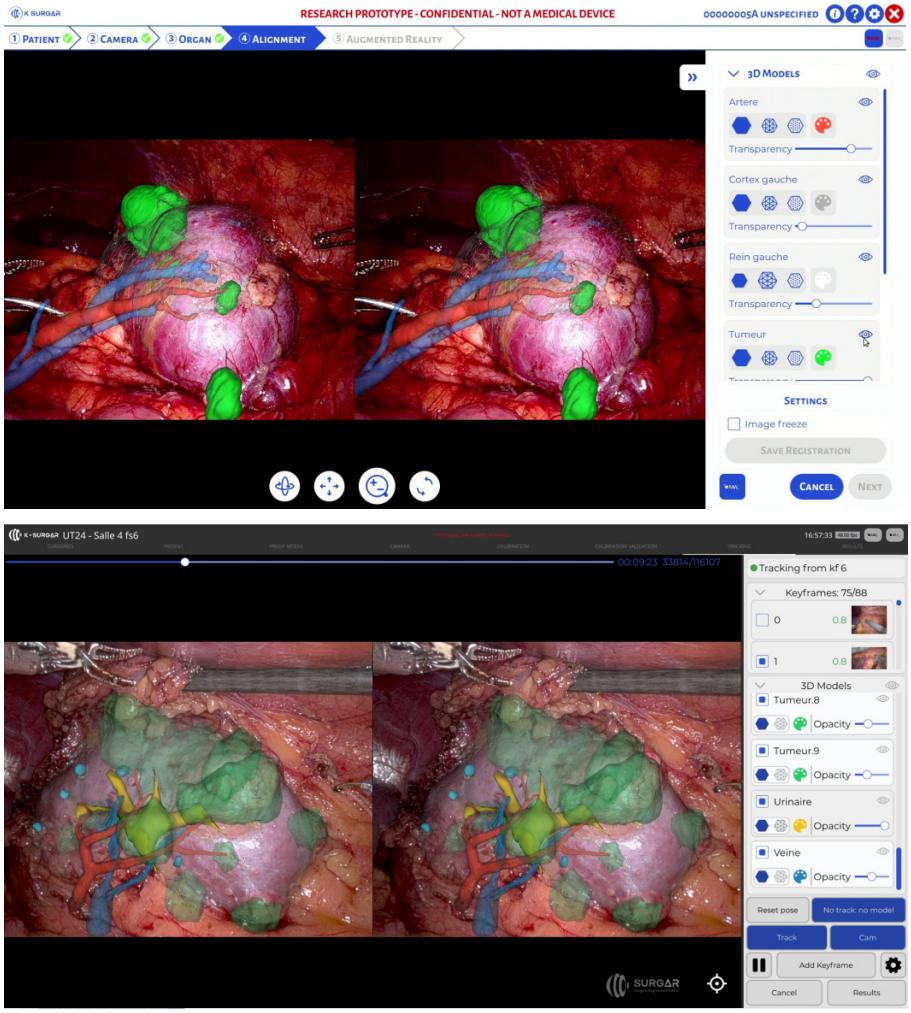


FIGURE 2.18 – Exemples d'interfaces du prototype de Réalité Augmentée (RA) destiné à la Néphrectomie Partielle (NP). (en haut) Version 0.1.0 d'octobre 2022. (en bas) Version 0.3.5 de février 2025. Source : auteur.

2.2.3 Approche utilisateur

Dans le cadre de la recherche et du développement d'un dispositif médical innovant dédié à la NP, les travaux de recherche ont intégré dès leur conception les contraintes spécifiques liées au domaine applicatif, conformément à une approche de conception pilotée par le domaine ou Domain-Driven Design (DDD) en Anglais. Cette méthodologie permet une intégration itérative et adaptative des besoins cliniques dans le développement technologique, assurant ainsi une pertinence optimale pour les utilisateurs finaux.

L'approche adoptée a permis une confrontation régulière des méthodes issues de la recherche avec les conditions cliniques réelles. Entre octobre 2022 et février 2025, huit sessions d'essais de faisabilité ont été conduites au bloc opératoire d'urologie du CHU de Bordeaux. Ces essais ont permis de recueillir des données pseudonymisées provenant de 31 interventions chirurgicales, afin de constituer une base de données pour la recherche et le développement. Cette dernière intègre notamment la modélisation numérique du rein sous forme de jumeaux numériques, visibles en annexe K, l'enregistrement du flux vidéo stéréo-endoscopique ainsi que l'évaluation de l'expérience et de la charge cognitive des chirurgiens monitorés via des instruments validés comme les échelles NASA-TLX [Hart 1988], Surg-TLX [Wilson 2011], Health-ITUES [Brown III 2013] ou l'UEQ (User-Experience Questionnaire, ISO 62366-2). L'exploitation rétrospective de ces cas, hors du bloc opératoire, permet de rejouer des scénarios chirurgicaux et de contribuer au développement de nouvelles méthodes pour l'assistance ou la planification du geste chirurgical. Les caractéristiques démographiques et cliniques de l'échantillon de patients opérés pour une tumeur rénale lors de ces essais sont présentées dans le tableau 2.3. L'inclusion des patients s'est faite de manière successive sans critère de présélection. La distribution des caractéristiques est conforme aux tendances épidémiologiques, notamment en ce qui concerne la prévalence masculine et la répartition des âges. Les cas analysés présentent majoritairement une complexité chirurgicale intermédiaire à élevée, justifiant l'intérêt clinique potentiel d'un système de chirurgie assistée par ordinateur.

Ces essais sur données s'inscrivent dans le cadre du consortium de RHU "Digital Urology 3D (DU3D)" (ANR-21-RHUS-0015), qui vise à transformer la chirurgie du cancer du rein en exploitant les technologies numériques et tridimensionnelles. Coordonné par le CHU de Bordeaux, ce programme repose sur une étroite collaboration entre des partenaires académiques et industriels afin de favoriser l'innovation en santé. Lancé le 15 avril 2022 pour une durée de cinq ans, il bénéficie d'un financement total de 18,3 millions d'euros, dont 6,1 millions d'euros de subventions publiques. Sa finalisation est prévue pour le 14 avril 2027.

Le projet DU3D réunit cinq partenaires industriels (Fujifilm, Rescol, Sophia Genetics et SURGAR) et trois partenaires académiques (le CHU de Grenoble, le CNRS via les équipes du CRPP et du LOF et l'Université de Bordeaux via Cohabit, ISPED et CREDIM). Le projet se structure autour de trois axes principaux :

- WP1 : Chirurgie guidée par l'image – De la réalité virtuelle à la réalité augmentée ;

- WP2 : Impact de la modélisation virtuelle et de l'impression 3D dans l'éducation des patients ;
- WP3 : Modèles réalistes du rein imprimés en 3D pour l'entraînement chirurgicale.

	Patients n = 31
Genre	
Homme	23 (74,19 %)
Femme	8 (25,81 %)
Age (ans)	
< 40	5 (16,13 %)
40-44	3 (9,68 %)
45-49	3 (9,68 %)
50-54	0 (0,00 %)
55-59	3 (9,68 %)
60-64	6 (19,35 %)
65-69	4 (12,90 %)
70-74	3 (9,68 %)
75-79	2 (6,45 %)
≥ 80	2 (6,45 %)
Intervention chirurgicale	
Néphrectomie partielle robot-assistée	28 (90,32 %)
Thrombectomie de la veine cave robot-assistée	3 (9,68 %)
Tumeur multifocale	
Oui	3 (9,68 %)
Taille moyenne de la tumeur (cm)	4,33 ± 2,04
Latéralité	
Gauche	13 (41,94 %)
Droite	18 (58,06 %)
Localisation tumorale	
Supérieure	8 (25,81 %)
Equatoriale	16 (51,61 %)
Inférieure	7 (22,58 %)
Face du rein	
Postérieure	12 (38,71 %)
Antérieure	17 (54,84 %)
Autre	2 (6,45 %)
Bord du rein	
Externe	17 (54,84 %)
Interne	14 (45,16 %)
Localisation hilaire	
Oui	11 (35,48 %)
Composante endophytique ou exophytique	
Entièrement endophytique	8 (25,81 %)
< 50\,\% exophytique	13 (41,94 %)
≥ 50\,\% exophytique	10 (32,25 %)
Proximité avec le système collecteur (mm)	
≥ 7	6 (19,35 %)
4-7	1 (3,23 %)
≤ 4	24 (77,42 %)
Complexité chirurgicale suivant le score RENAL	
Basse	4 (12,91 %)
Intermédiaire	16 (51,61 %)
Haute	11 (35,48 %)

TABLE 2.3 – Caractéristiques démographiques et cliniques de l'échantillon de patients opérés pour une tumeur du rein lors des essais de faisabilité. Source : auteur.

2.2.4 État de l'art

Dans cette section, nous nous intéresserons aux méthodes de l'état de l'art en mesure de générer des augmentations peropératoires en RA lors de la chirurgie du rein. La littérature fait parfois référence à une visualisation dite holographique [Zeng 2021] qui peut être communiquée au chirurgien par un affichage bidimensionnel (2D) ou tridimensionnel (3D) sur un dispositif comme la console du télémanipulateur utilisée en chirurgie robotique, voir section 1.2.4.3.

Le guidage chirurgical assisté par RA repose fondamentalement sur une étape de recalage, qui consiste à aligner le jumeau numérique avec l'image endoscopique peropératoire. La précision du recalage conditionne directement la fiabilité du guidage et par conséquent l'exactitude du geste chirurgical [Andrews 2020, Bertolo 2020]. En chirurgie rénale, cette tâche est particulièrement complexe en raison de l'absence de repères anatomiques communs entre l'imagerie préopératoire et l'image peropératoire. Ces éléments sont souvent désignés sous le nom de primitives de recalage. De surcroît, la manipulation peropératoire des structures anatomiques entraîne des déformations et rend l'alignement d'autant plus difficile [Magaribuchi 2024].

Les méthodes de recalage décrites dans la littérature sont généralement réparties en trois grandes catégories [Khaddad 2023, Wang 2024a, Roberts 2022, Manfredi 2022, Magaribuchi 2024, Puliatti 2022, Rassweiler 2014, Bertolo 2020, Detmer 2017, Hughes-Hallett 2014a, Makiyama 2023] : le recalage manuel aussi appelé approche interactive, le recalage basé sur des marqueurs artificiels aussi appelés fiduciaires et le recalage fondé sur la reconstruction de surface.

Le recalage manuel demeure aujourd'hui la méthode de référence dans la majorité des études, notamment en raison de sa simplicité de mise en œuvre. Il s'appuie sur l'expertise du chirurgien, sa capacité cognitive à interpréter l'anatomie en trois dimensions et à aligner visuellement le jumeau numérique avec l'image peropératoire. Cette approche autorise des ajustements en temps réel selon les six DDL – trois translations et trois rotations – et peut être effectuée à tout moment de l'intervention [Makiyama 2023]. Néanmoins, l'absence d'automatisation impose de répéter l'ajustement à chaque mouvement de la caméra ou de l'organe, ce qui rend le processus fastidieux, potentiellement imprécis et dépendant de l'expérience de l'opérateur [Makiyama 2023]. Des erreurs de recalage ou Target Registration Error (TRE) en Anglais de l'ordre de $7,0 \pm 2,0$ millimètres ont ainsi été rapportées [Kobayashi 2020].

La faisabilité technique du guidage par RA basé sur un recalage manuel a été largement démontrée [Ukimura 2008, Teber 2009, Nakamura 2010, Zhang 2022a, Zhang 2021a, Porpiglia 2018a, Zeng 2021, Li 2022b, Teluob 2019], y compris à travers de larges séries cliniques présentant des résultats prometteurs tant sur le plan peropératoire que postopératoire [Chen 2014, Schiavina 2021a, Amparore 2022b]. Par exemple, une étude incluant 65 patients a montré une adoption accrue du clampage sélectif (58,4 %) et de l'énucléation pure (63,1 %), accompagnée de faibles taux de marges positives (2,5 %) et de complications majeures (1,5 %) [Piramide 2022]. Deux études cas-témoins [Porpiglia 2018b, Porpiglia 2020], utilisant la RA avec reca-

lage manuel pour des tumeurs rénales complexes, ont démontré une réduction significative du taux de clampage artériel total (45,8 % contre 69,7 %, p = 0,03) et de l'ouverture des voies urinaires (10,4 % contre 46,5 %, p = 0,0003), sans impact significatif sur les complications post-opératoires. Des bénéfices sur la préservation fonctionnelle rénale ont également été observés, bien que non significatifs pour le DFG estimé à trois mois (-14,26 % contre -19,46 %, p-value = 0,11), mais significatifs pour le débit plasmatique rénal effectif (-12,3 contre -18,14, p-value = 0,01) [Porpiglia 2020]. Cette approche manuelle est également utilisée dans des interventions plus complexes, telles que les chirurgies de thrombus cave [Amparore 2024a].

L'ergonomie de l'interface d'interaction joue un rôle clé dans l'efficacité du recalage manuel. Elle doit être en mesure de numériser l'action de recalage de l'opérateur afin de la transposer dans la scène virtuelle. Divers dispositifs ont été testés comme la souris standard qui permet des déplacements physiques dans le plan et des interactions via des boutons ; la souris 3D de type SpaceMouse (3Dconnexion) à 6 DDL qui permet des mouvements tridimensionnels sans mouvement physique de la souris [Schiavina 2021a] ; des systèmes de type tablettes tactiles [Hughes-Hallett 2014b] ; des interfaces sans contact par caméra RGB-D ou infrarouge qui permettent le suivi des mains de l'opérateur [Ruppert 2012]. Enfin, en chirurgie cœlioscopique, la mobilisation du jumeau numérique peut également être asservie au suivi du mouvement d'un outil ou de la caméra endoscopique. Ces solutions spécifiques offrent une précision variable, certaines entraînant des erreurs de rotation deux fois supérieures à celles de méthodes classiques par appariement de points [Joeres 2021a].

Afin de simplifier le recalage manuel au prix d'une légère perte de précision, certains auteurs limitent l'alignement initial à trois DDL, en fixant manuellement un point commun dans l'image et le modèle, puis en ajustant par rotations. Cette approche présente une erreur moyenne d'alignement de 4,76 millimètres contre 3,16 millimètresmm pour un recalage complet à 6 DDL [Pratt 2012].

En complément des transformations rigides, des déformations paramétriques linéaires ou non linéaires [Sederberg 1986, Barr 1984] peuvent être appliquées au jumeau numérique pour simuler la plasticité des tissus. Bien que pertinentes, elles alourdissent le processus de recalage, souvent chronophage, et sont généralement déléguées à un assistant pour ne pas interrompre trop longtemps la chirurgie [Porpiglia 2020].

L'automatisation du recalage a fait l'objet de nombreuses recherches, notamment à travers l'utilisation de marqueurs artificiels radio-opaques implantés en peropératoire. Ces derniers sont localisés par des systèmes d'imagerie ionisante comme le Cone-Beam Computed Tomography (CBCT) en Anglais et permettent de déterminer la position relative des marqueurs et de l'organe cible. Typiquement au nombre de cinq, ces marqueurs visibles à la fois dans l'imagerie et par la caméra endoscopique, servent de primitives de recalage [Teber 2009, Baumhauer 2008]. Au cours de l'intervention, des acquisitions en mode scopie (4 images par seconde) peuvent être réalisées afin de compenser dynamiquement les déformations anatomiques [Simpfendörfer 2016].

Parallèlement, des systèmes proches de la navigation chirurgicale (cf. section 1.2.4.4) s'appuient sur des capteurs optiques ou électromagnétiques pour assurer le suivi relatif de l'instrumentation chirurgicale par rapport à la position de la caméra endoscopique. Le recalage initial repose généralement sur l'utilisation de marqueurs artificiels comme décrit précédemment, qui permettent de synchroniser la scène peropératoire avec le jumeau numérique. Les instruments sont ensuite localisés par des capteurs externes selon une approche dite inside-out, dans laquelle des marqueurs extracorporels sont détectés pour trianguler la position des instruments intra-abdominaux. Des méthodes telles que POSIT ou ORTHIT sont couramment utilisées [Wild 2016]. Les erreurs de recalage TRE associées à ces techniques sont faibles, atteignant 0,78 millimètre pour la navigation optique et 0,93 millimètre pour la navigation électromagnétique [Jackson 2020]. En contexte clinique, l'erreur globale de visualisation ou Target Visualisation Error (TVE) en Anglais sur l'image augmentée a été estimée à 2,7 millimètres pour des néphrectomies assistées par RA, avec un allongement médian du temps opératoire de 15,5 minutes.

Une alternative innovante aux marqueurs classiques consiste en l'utilisation de grilles de points implantables à la surface du parenchyme rénal. Ce dispositif sert de référentiel de localisation, dont la position par rapport à la tumeur est déterminée par échographie. Le déplacement relatif de la sonde échographique est ensuite analysé visuellement dans ce référentiel. Les instruments, quant à eux, sont localisés à partir des données cinématiques du télémanipulateur robotique, afin de mesurer la distance entre la pointe de l'outil et le centre de la tumeur [Singla 2017, Edgcumbe 2018].

Cependant, l'implantation de tels dispositifs peut entraîner des saignements, justifiant le développement de marqueurs non invasifs, appliqués par simple collage [Magaribuchi 2024, Wild 2016]. Par ailleurs, l'emploi de marqueurs fluorescents améliore la robustesse du système face à des perturbations visuelles telles que le sang ou la fumée.

Une approche particulièrement attractive, car elle ne requiert ni instrumentation additionnelle ni intervention manuelle conséquente, repose sur le recalage basé sur la reconstruction de surface. Il a été estimé qu'une cartographie d'environ 28 % de la surface rénale est suffisante pour un recalage précis [Benincasa 2008], à l'aide d'algorithmes de mise en correspondance de nuages de points. Deux stratégies de reconstruction tridimensionnelle peropératoire sont identifiées : le sondage par instrument et la stéréovision.

Dans le cas du sondage, la navigation optique permet de suivre l'instrument utilisé comme stylet afin d'échantillonner la surface rénale [Herrell 2009, Joeres 2021b, Altamar 2011]. Les coordonnées tridimensionnelles des points sondés sont obtenues soit par triangulation des données optiques, soit via les modèles cinématiques du télémanipulateur robotique, bien que cette dernière solution reste complexe à intégrer en pratique clinique [Makiyama 2023, Kavoussi 2021, Ferguson 2020]. Les points ainsi collectés permettent de reconstituer la surface anatomique, ensuite alignée avec le jumeau numérique.

Des expériences sur phantoms en hydrogel ont permis d'évaluer l'erreur quadratique moyenne de recalage de l'organe complet, estimée à 2,53 millimètres pour une tumeur endophytique et 4,88 millimètres pour une tumeur exophytique. Les erreurs localisées dans la zone de résection de la tumeur sont respectivement de 0,75 millimètre et 2,15 millimètres.

Dans les contextes dynamiques, des recalages itératifs peuvent être requis pour corriger les erreurs d'alignement. Une méthode basée sur la Perspective-n-Point (PnP) peut être employée lorsque des points de référence ont été identifiés au préalable à l'aide du stylet de détection.

La stéréovision constitue une alternative intéressante à l'échantillonnage manuel [Zhang 2019b, Stoyanov 2010, Zampokas 2022, Hamarneh 2014, Su 2009, Amir-Khalili 2013]. Elle repose sur l'analyse de cartes de disparité obtenues par des algorithmes de type Semi-Global Block Matching (SGBM) en Anglais [Hirschmuller 2007], nécessitant un traitement GPU pour atteindre des performances de l'ordre de vingt images par seconde [Zampokas 2022]. Néanmoins, plusieurs défis persistent en chirurgie rénale comme la variabilité de la morphologie rénale, la présence de graisse péri-rénale, les vues partielles, les saignements et la fumée.

Pour limiter l'impact négatif des structures adjacentes, des techniques de segmentation automatique fondées sur des réseaux de neurones profonds tels que Mask-R-CNN [He 2017] sont utilisées. Une stratégie d'agrégation de plusieurs vues adjacentes peut également permettre de créer un nuage de points unifié à partir de différents patchs de surface et donc d'éviter les minimas locaux [Zhang 2019b]. Pour cela, des points clés sont extraits via des méthodes comme Speeded-Up Robust Features (SURF) en Anglais [Bay 2006], mis en correspondance via une stratégie RANdom SAmple Consensus (RANSAC) en Anglais [Fischler 1981]. Les nuages de points sont alignés, seules les parties sans recouvrement sont conservées. Une étape d'ajustement de faisceaux est réalisée pour atténuer les erreurs. Le nuage de point résultant est sous-échantillonné pour réduire le bruit.

Indifféremment de la stratégie de reconstruction tridimensionnelle peropératoire, le recalage entre le jumeau numérique et le nuage de point reconstruit peut alors être réalisé suivant une approche en deux étapes (coarse-to-fine), comme illustrée sur la figure 2.19. La première étape comprend souvent une Analyse en Composante Principale (ACP) [Wold 1987] suivie d'un alignement par Iterative Closest Point (ICP) en Anglais [Zhang 2023b, Zhang 2019b]. La seconde étape, consiste à prendre en compte la déformation via Coherent Point Drift (CPD) en Anglais [Myronenko 2010] qui formule le problème d'alignement entre les deux nuages de points comme un problème de classification non supervisée modélisé par un modèle de mélange gaussien ou Gaussian Mixture Model (GMM) en Anglais [Reynolds 2009] avec une contrainte de cohérence de mouvement. Des essais sur phantoms déformables ont révélé une erreur moyenne de $1,28 \pm 0,68$ millimètre [Zhang 2019b]. Une version améliorée, incluant une estimation volumique de la déformation par interpolation via Gaussian Radial Basis Function (GRBF) en Anglais [Orr 1996] et une sélection partielle via Projection-Based Partial Selection (PBPS) en Anglais du nuage de points, atteint

un TRE de $1,69 \pm 0,31$ millimètre avec des temps de traitement compris entre 1 et 24 secondes [Zhang 2020b, Vagvolgyi 2008].

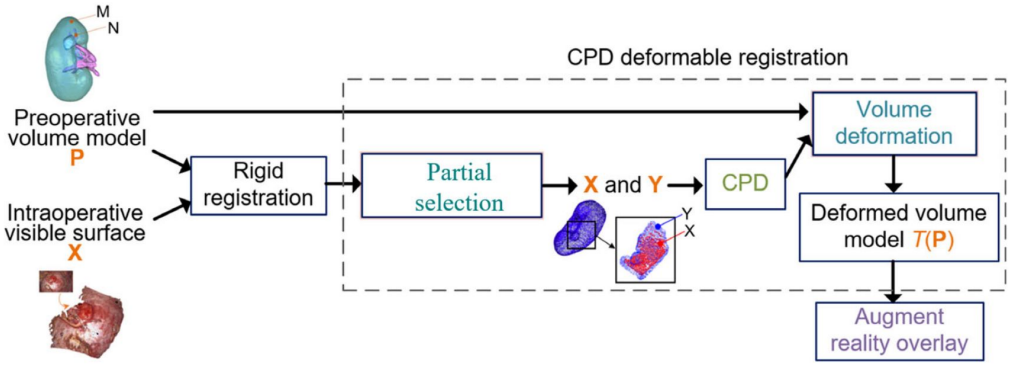


FIGURE 2.19 – Représentation schématique d'un pipeline de recalage déformable par Coherent Point Drift (CPD) pour la Néphrectomie Partielle (NP) guidée par Réalité Augmentée (RA). Source : [Zhang 2020b].

Plus récemment, des approches de recalage automatique reposant sur l'apprentissage profond ont émergé. Une première méthode utilise un encodeur PointNet++ [Qi 2017] combiné à un décodeur MultiLayer Perceptron (MLP) en Anglais [Rosenblatt 1958] pour prédire une carte d'occupation à partir d'un nuage de points de la surface rénale. Le modèle est entraîné par simulation à partir du jumeau numérique déformé du patient [Henrich 2024]. Une autre méthode, illustrée en figure 2.20, propose une estimation de pose à six DDL à partir de la segmentation de primitives dans l'image endoscopique, comme la surface rénale. Le masque ainsi obtenu permet, selon les auteurs, d'estimer la translation tridimensionnelle et l'échelle du jumeau numérique dans l'image. Il en suit une estimation de la rotation tridimensionnelle via un réseau de neurones entraîné sur des données synthétiques. Toutefois, dans 50 % des cas, l'erreur de rotation dépasse 10° sur les axes X/Y et 5° sur Z, nécessitant une correction manuelle [Pecoraro 2024, Padovan 2022]. Les performances restent limitées par la nécessité d'une segmentation quasi-complète, rarement atteinte en pratique [Piazzolla 2025].

Une étude sur 105 patients rapporte un taux de succès de 95 % pour un recalage basé sur deux repères veineux suivi d'un ajustement de l'échelle et de la rotation via un modèle d'apprentissage, bien que les détails méthodologiques soient absents [Xiaolei 2025].

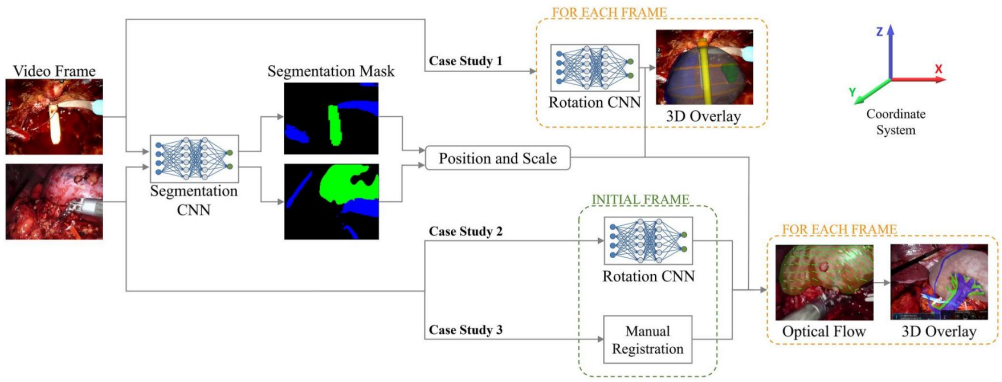


FIGURE 2.20 – Représentation schématique d'un pipeline de recalage rigide basé sur la segmentation sémantique et l'inférence d'une rotation tridimensionnelle. Source : [Padovan 2022].

Afin de simplifier l'interprétation visuelle de la scène chirurgicale, une autre stratégie exploite l'injection de vert d'indocyanine ou ICG en Anglais et la capture en fluorescence proche infrarouge pour identifier la surface rénale (cf. section 1.3.2.2), comme illustré en figure 2.21. Celle-ci est segmentée via un seuillage adaptatif ou par un réseau de neurones avec un encodeur ResNet-50 [He 2016], puis une ellipse est ajustée sur le masque de segmentation afin d'estimer la pose du rein [Amparore 2024b]. Bien que le recalage prenne en moyenne sept secondes en mode ICG et onze secondes en mode RVB, une intervention humaine demeure nécessaire pour ajuster la rotation du jumeau numérique. Des résultats chirurgicaux satisfaisants ont été rapportés pour dix néphrectomies partielles robot-assistées [Amparore 2022a]. Les déplacements extrêmes du rein restent néanmoins problématiques [Pecoraro 2024].

Les recalages initiaux décrits précédemment sont réalisés sur image fixe ou ensemble d'images. Ils imposent un renouvellement de cette même opération en cas de mouvement de l'organe ou de la caméra. Pour pallier à cette limitation, des techniques de suivi ont été proposées dans la littérature, principalement basées sur le flux optique [Pecoraro 2024] ou sur le Simultaneous Localisation And Mapping (SLAM) [Mur-Artal 2015], avec des résultats encourageants [Jia 2021]. Ces méthodes nécessitent généralement la segmentation continue de la surface rénale, souvent assurée par un modèle de type SiamMask [Wang 2019]. Les points de référence (keypoints en Anglais) sont extraits et décrits à l'aide d'algorithmes comme Features From Accelerated Segment Test (FAST) [Rosten 2006] et Oriented FAST and Rotated BRIEF (ORB) [Rublee 2011], appariés via Brute Force Matching (BFM) puis filtrés par des techniques comme Grid-based Motion Statistics (GMS) [Bian 2017]. La stéréovision peut aussi être intégrée pour la mise en correspondance des points de référence. L'estimation de pose est faite entre les points appariés et le nuage de

points épars reconstruit, un raffinage par ajustement de faisceaux peut également être effectué.

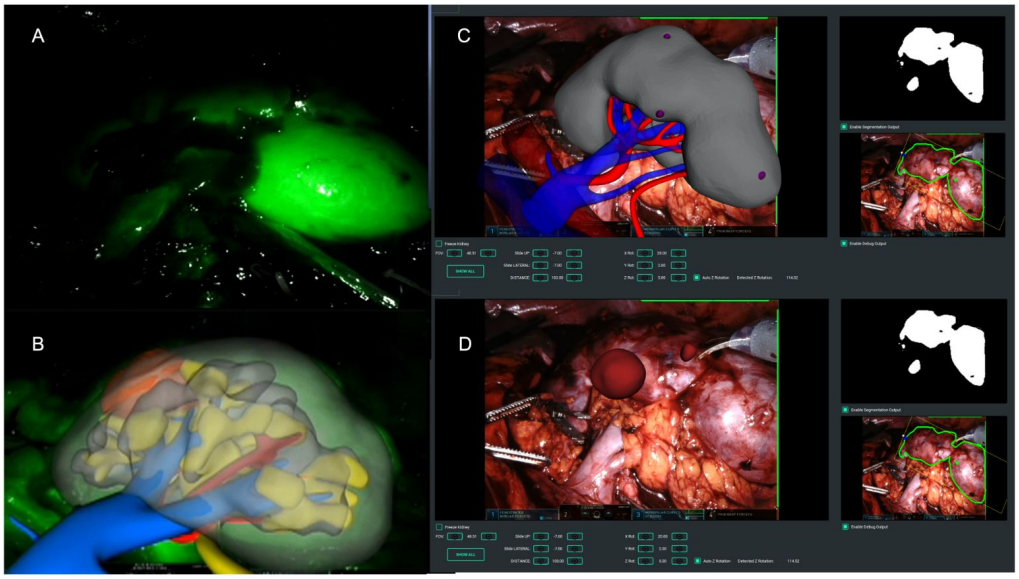


FIGURE 2.21 – Solution de Réalité Augmentée (RA) pour la Néphrectomie Partielle (NP) capable d'utiliser la surface rénale comme primitive de recalage automatique. Source : [Amparore 2024b].

Un enjeu majeur demeure l'évaluation in vivo de ces systèmes. En effet, les conditions peropératoires sont bien plus complexes que celles rencontrées ex-vivo et les métriques cliniques courantes ne permettent pas de quantifier précisément l'erreur de recalage [Porpiglia 2020, Hughes-Hallett 2014b, Chen 2014]. Le protocole "Bystander" a été développé à cette fin. Il repose sur une "double intervention chirurgicale" où un premier opérateur réalise la procédure de manière indépendante, pendant qu'un second utilise exclusivement le système de RA pour le recalage et la mesure de points d'intérêts (centre tumoral, site de clampage). La comparaison des localisations respectives avec et sans utilisation du système de guidage pendant une même intervention, permet alors de calculer l'erreur de localisation à la cible ou Target Localisation Error (TLE) en Anglais [Kavoussi 2021, Nimmagadda 2022]. Ce protocole présente néanmoins des limitations liées à la capacité de l'être humain a pointer un point précis, à la déformation de l'organe, à la difficulté d'estimer la position d'un point cible intraparenchymateux [Cannon 2023].

2.2.5 Acquisition vidéo endoscopique

Dans un système de CAO qui intègre un ou plusieurs flux vidéo parmi ses données d'entrée, le module d'acquisition constitue un composant à la fois matériel et logiciel fondamental. Ses performances influencent l'ensemble de la chaîne de traitement, depuis la capture de l'image jusqu'à son affichage ou son traitement algorithmique.

Les dispositifs vidéo couramment utilisés en chirurgie, tels que les colonnes de cœlioscopie, les microscopes opératoires ou encore les systèmes robotisés de télémanipulation, disposent généralement de sorties vidéo standardisées. Ces connecteurs permettent de relier des systèmes tiers, typiquement en vue de redistribuer le flux caméra vers un dispositif d'affichage, qu'il s'agisse d'un écran mural ou d'un moniteur monté sur bras articulé.

La transmission du signal vidéo entre la source et l'affichage repose majoritairement sur des liaisons filaires, utilisant soit des conducteurs en cuivre, soit plus rarement des fibres optiques. Selon le mode de transmission, le signal peut transiter via une liaison série (monoconducteur) ou parallèle (multiconducteurs) et se présenter sous forme analogique ou numérique. Un signal analogique consiste en une onde électrique continue dont la tension code directement les informations chromatiques, tandis qu'un signal numérique repose sur des données binaires encodées et souvent compressées.

L'usage des signaux analogiques vidéo est aujourd'hui marginal, en raison de la transition vers des chaînes de traitement entièrement numériques et de la vulnérabilité intrinsèque de ces signaux aux interférences électromagnétiques. Le format S-Video en est un exemple représentatif. À l'inverse, les signaux numériques se sont imposés comme la norme, leur production est plus simple, leur robustesse accrue et leur fidélité de reproduction nettement supérieure.

Différentes interfaces sont utilisées pour le transport des flux vidéo numériques, en fonction du contexte d'utilisation, qu'ils soient domestiques ou professionnels. Le Serial Digital Interface (SDI) en Anglais, est notamment déployé dans les environnements professionnels comme les plateaux de télévision et permet une transmission sans compression de contenus en très haute définition, avec une quantification sur 8, 10 ou 12 bits. Ce signal est véhiculé sur un ou plusieurs câbles coaxiaux à impédance caractéristique de 75 ohms, terminés par des connecteurs Bayonet Neill–Concelman (BNC), assurant un verrouillage sécurisé par quart de tour. Pour les usages grand public, des standards comme High-Definition Multimedia Interface (HDMI) en Anglais, Digital Visual Interface (DVI) en Anglais ou encore DisplayPort (DP) sont largement répandus. Chacun de ces protocoles présente des caractéristiques propres en matière de bande passante, de robustesse face aux pertes de signal et de compatibilité matérielle. Le tableau 2.4 reprend les spécifications essentielles.

Certaines de ces interfaces intègrent également des métadonnées complémentaires au signal vidéo principal. L'Extended Display Identification Data (EDID) en Anglais joue par exemple un rôle déterminant dans la détection automatique des résolutions et paramètres vidéo entre la source et l'afficheur. Le High-bandwidth Digital Content Protection (HDCP) en Anglais quant à lui, assure la protection contre la copie numérique directe en restreignant l'usage des contenus transmis.

Caractéristique	SDI	HDMI	DVI	DP
Type de signal	Numérique série unidirectionnel	Numérique TMDS (Transition Minimized Differential Signaling)	Numérique TMDS (DVI-D), analogique (DVI-A)	Numérique via micro-paquets (Main Link)
Normes principales	SMPTE 259M, 292M, 424M, 2081, 2082, 2082	HDMI 1.0 – 2.1b (norme CEA-861)	DVI 1.0 (1999, par le DDWG)	VESA DP 1.0 – 2.1a
Débit binaire théorique	12G-SDI : 11,88 Gbit/s	HDMI 2.1 : 48 Gbit/s	DVI-D Dual Link : 7,92 Gbit/s	DP 2.0 : 77,4 Gbit/s
Résolution maximale	Jusqu'à 4K@60Hz (12G), 8K@60Hz avec 24G/Quad-link	Jusqu'à 10K@120Hz (HDMI 2.1)	3840×2400@30Hz (Dual Link)	Jusqu'à 16K@60Hz (DP 2.0)
Audio transporté	Oui	Oui	Non	Oui
Longueur maximale	Jusqu'à 100 m (3G), ~70 m (6G), ~30 m (12G)	10 m (HDMI 2.0), ~2–3 m (HDMI 2.1)	~5 m pour Dual Link	~1–2 m (DP 2.1)
Connectique	BNC 75Ω verrouillable	HDMI Type A (Standard), Type C (Mini), Type D (Micro)	DVI-D, DVI-A, DVI-I (Single/Dual Link)	DP Standard, Mini DP
Sécurité	Très élevée (verrouillage BNC)	Faible (pas de verrouillage natif)	Variable	Variable
Support HDCP	Non	Oui	Oui	Oui
Support de la compression	Non	Oui	Non	Oui
Chaînage d'écrans (MST)	Non	Non	Non	Oui

Table 2.4 – Principales caractéristiques des interfaces vidéo numérique SDI, HDMI, DVI et DP. Source : auteur.

Pour permettre leur transmission, les flux vidéo numériques sont soumis à des processus d'encodage à la source, puis de décodage au niveau de l'affichage. Les paramètres d'encodage, tels que le débit binaire (bitrate en Anglais), le mode de rafraîchissement (Variable Frame Rate (VFR) ou Constant Frame Rate (CFR) en Anglais), ainsi que les éventuelles techniques de compression, conditionnent directement la qualité visuelle perçue et la latence du signal, deux éléments critiques dans des contextes comme l'endoscopie. Dans certains cas, notamment en chirurgie cœlioscopique ou robotique, des systèmes d'imagerie stéréoscopique sont déployés (cf. sections 1.2.4.1 et 1.2.4.3), impliquant la transmission simultanée de paires d'images (gauche et droite). Cette transmission peut s'effectuer soit via une seule liaison selon

des techniques de multiplexage comme le Frame Packing, le Side-by-Side ou le Top-and-Bottom – comme dans le cas des colonnes de cœlioscopie du fabricant Storz – soit à travers deux canaux indépendants, comme c'est le cas pour le télémanipulateur chirurgical Da Vinci (Intuitive Surgical). La chaîne d'acquisition matérielle que nous utilisons en chirurgie robotique est représentée schématiquement en figure 2.22.

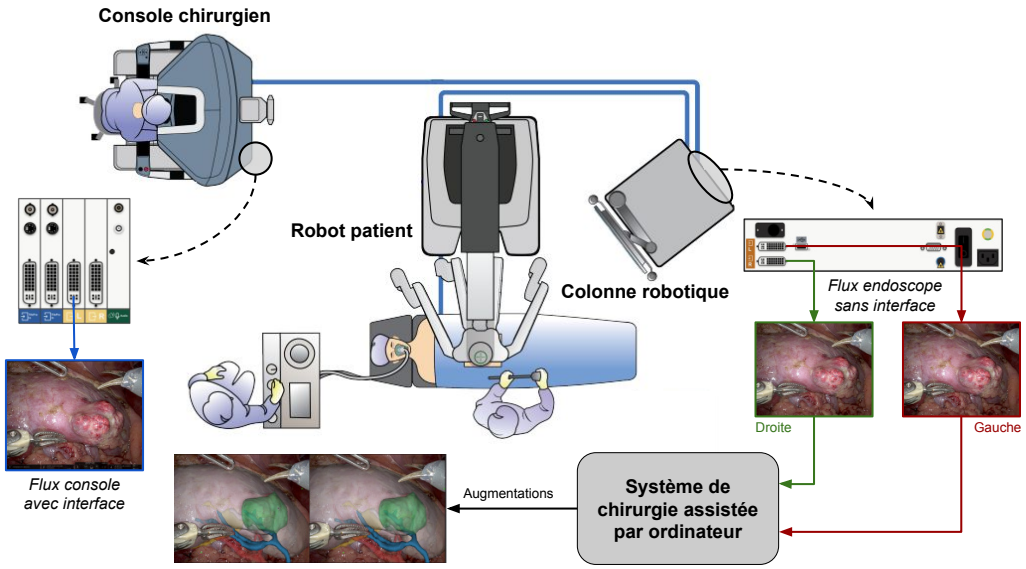


FIGURE 2.22 – Représentation schématique de la chaîne d'acquisition vidéo stéréoscopique utilisable en chirurgie robotique par un système de chirurgie assistée par ordinateur. Générée par l'auteur à partir de [Surgical 2025b].

Dans les systèmes de CAO intégrant des technologies de RA, l'acquisition du flux vidéo endoscopique repose généralement sur l'utilisation d'une carte d'acquisition vidéo, souvent connectée à une carte mère via le bus Peripheral Component Interconnect (PCIn) en Anglais. Le choix de cette carte est déterminé par plusieurs critères techniques : le facteur de forme, le nombre de canaux pris en charge, la connectique d'entrée, la résolution supportée, la fréquence d'échantillonnage, la présence d'un moteur de décodage matériel ou logiciel ainsi que la capacité à gérer la synchronisation intercanaux. Ces cartes exploitent généralement un traitement embarqué via Field-Programmable Gate Array (FPGA) en Anglais, couplé aux ressources de l'hôte grâce à un Software Development Kit (SDK) en Anglais. Le flux capturé devient ainsi accessible aux applications logicielles du système hôte, qui peuvent s'appuyer sur des bibliothèques spécialisées telles que FFmpeg, GStreamer ou DeepStream (Nvidia) pour le traitement vidéo.

Dans le contexte des dispositifs électromédicaux, le respect des normes de sécurité électrique (ISO 60601-1) et électromagnétique (ISO 60601-1-2) constitue un impératif. Ces normes visent à prévenir à la fois les fuites de courant vers le patient et les perturbations électromagnétiques pouvant affecter l'environnement opératoire ou être générées par celui-ci. Concernant les signaux vidéo, une isolation galvanique est généralement mise en œuvre entre la source et les dispositifs de capture ou d'affichage, garantissant une isolation électrique efficace entre les circuits.

2.2.6 Étalonnage géométrique d'une caméra endoscopique

Sur le plan physique, une caméra permet de transformer une scène tridimensionnelle du monde réel en une représentation bidimensionnelle projetée sur la rétine d'un capteur d'image. Ce processus repose sur la collecte des rayons de lumière émis ou réfléchis par les objets de la scène, acheminés à travers un système optique composé de lentilles. En vision par ordinateur, une caméra est fréquemment modélisée selon le principe de la projection perspective centrale, désigné sous le terme de modèle de caméra sténopé. Ce modèle géométrique simplifié fait abstraction des phénomènes complexes associés à la propagation de la lumière dans des systèmes aux multiples lentilles. Il décrit exclusivement les relations projectives entre les points tridimensionnels de la scène et leurs correspondances bidimensionnelles sur le plan image. Malgré ces approximations, ce modèle a fait ses preuves dans de nombreux domaines et convient généralement à un grand nombre d'applications.

Dans le modèle sténopé, la scène est observée depuis un centre de projection également désigné sous le nom de centre optique, noté $O = (0, 0, 0)^T$, lequel constitue l'origine du repère caméra. Tous les rayons de lumière convergent vers ce point unique, traversent un orifice ponctuel et viennent former l'image sur une rétine imaginaire placée à une distance focale f en avant de ce point. Cette disposition, bien que conceptuelle, adopte une configuration géométrique symétrique à celle du capteur physique, qui dans une caméra réelle est situé en arrière du système optique. L'axe Z de la caméra est l'axe optique, il traverse le centre du système optique et définit l'orientation principale du repère caméra dans la direction de l'espace observé, comme illustré en figure 2.23. Un point tridimensionnel de la scène $\overrightarrow{P_c} = (X_c, Y_c, Z_c)^T$ exprimé dans le repère caméra, se projette au point d'intersection $\overrightarrow{p} = (u, v, f)^T$ entre le plan image et une droite passant par le point P et le centre de projection O selon :

$$\overrightarrow{p} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{f_x}{Z_c} X_c + c_x \\ \frac{f_y}{Z_c} Y_c + c_y \end{bmatrix} \tag{2.1}$$

où f_x et f_y désignent les distances focales exprimées en pixels et (c_x, c_y) correspondent aux coordonnées du point principal, c'est-à-dire l'intersection entre l'axe optique et le plan image [Hartley 2003, Yeong 2021].

Cette formulation est souvent exprimée en coordonnées homogènes pour intégrer les transformations projectives dans un formalisme matriciel :

$$\underbrace{\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}}_{\text{Coordonnées image}} = \underbrace{\alpha}_{\text{Facteur d'échelle}} \underbrace{\begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{K,\text{Matrice intrinsèque}} \underbrace{\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}}_{\text{Coordonnées caméra}} \tag{2.2}$$

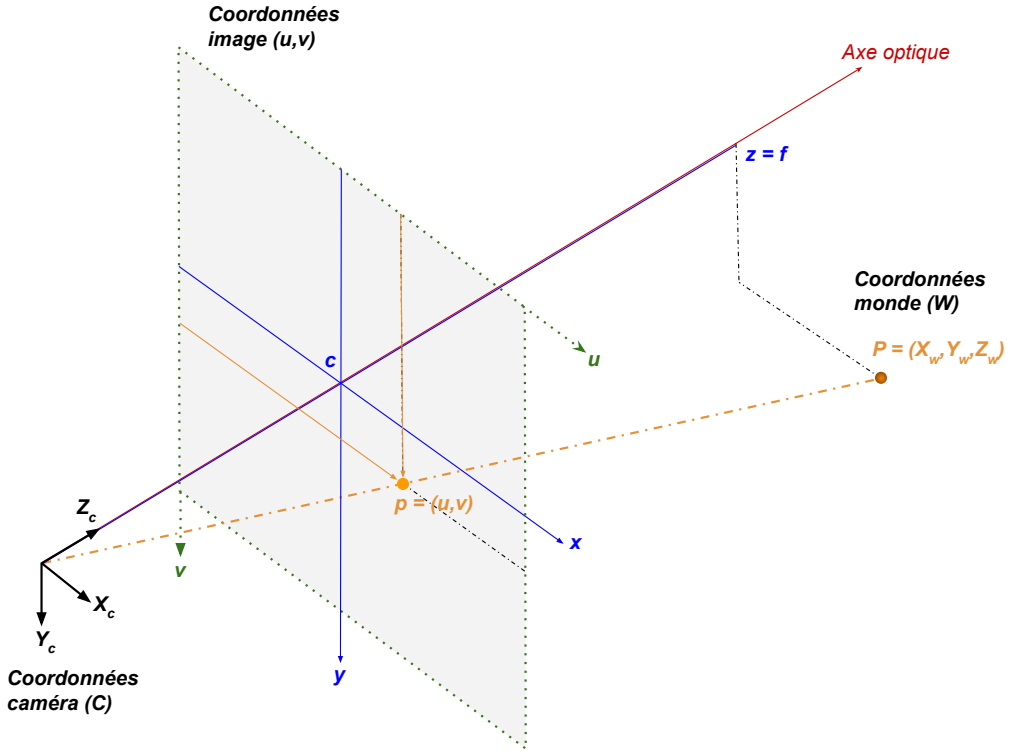


FIGURE 2.23 – Représentation schématique du modèle de caméra sténopé. Un point tridimensionnel $\overrightarrow{P}$ dans le repère monde se projette sur le plan image virtuel en un point bidimensionnel $\overrightarrow{p}$. Source : auteur.

Le passage du monde réel à l'image requiert une connaissance des paramètres intrinsèques de la caméra, regroupés dans la matrice intrinsèque K, mais aussi des paramètres extrinsèques qui définissent la transformation rigide, aussi appelée pose de la caméra, entre le repère monde et le repère caméra. Ces derniers s'expriment sous forme d'une matrice $[R|t]$, où $R \in SO(3)$ est une matrice de rotation et $t \in \mathbb{R}^3$ est un vecteur de translation. L'ensemble de ces éléments permet de relier un point $\overrightarrow{P}_w$ exprimé dans le repère monde à sa projection $\overrightarrow{p}$ sur le plan image via la relation :

$$\overrightarrow{p} = \alpha K \, [R|t] \, \overrightarrow{P}_w \tag{2.3}$$

où $\alpha \in \mathbb{R}$ est un facteur d'homogénéité, inhérent à la représentation en coordonnées projectives.

Le modèle sténopé suppose une projection idéale dans laquelle la lumière suit des trajectoires rectilignes à travers un orifice ponctuel, sans subir d'altération. Toutefois, cette approximation s'écarte des phénomènes observés en pratique, où les systèmes optiques réels introduisent diverses formes d'aberrations géométriques. Parmi celles-ci, les distorsions radiales et tangentielles constituent les plus fréquentes [Ramírez-Hernández 2020]. La distorsion radiale se manifeste par une déviation systématique des rayons de lumière à mesure qu'ils s'éloignent de l'axe optique, entraînant une courbure visible des lignes droites dans l'image. Elle peut être modélisée par un développement polynomial en fonction de la distance radiale r au point principal (c_x, c_y), où $r^2 = m_u^2 + m_v^2$ avec $m_u = u - c_x$ et $m_v = v - c_y$. La position d'un point idéal m affecté par la distorsion radiale peut alors s'exprimer par :

$$m'' = m \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{2.4}$$

où k_1, k_2 et k_3 sont les coefficients du modèle de distorsion radiale.

La distorsion tangentielle résulte généralement d'un léger désalignement entre les axes optiques des différentes lentilles. Elle peut se modéliser pour deux coefficients de distorsion tangentielle p_1 et p_2 pour un point idéal m par :

$$m''_u = m_u + 2p_1 m_u m_v + p_2(r^2 + 2m_u^2) \tag{2.5}$$

$$m''_v = m_v + 2p_2 m_u m_v + p_1(r^2 + 2m_v^2) \tag{2.6}$$

L'étalonnage géométrique d'une caméra consiste à estimer conjointement l'ensemble des paramètres intrinsèques (focale, point principal, coefficients de distorsion) et extrinsèques (pose de la caméra) à partir de l'observation d'objets dont la géométrie est connue dans l'espace tridimensionnel avec précision.

Plusieurs approches d'étalonnage géométrique existent dans la littérature [Cui 2023]. Parmi les méthodes les plus courantes figurent celles basées sur une référence externe, souvent sous la forme d'une mire bidimensionnelle ou tridimensionnelle. Ces méthodes nécessitent une image unique (Single Image Calibration (SIC) en Anglais) [Barreto 2009] ou une séquence d'images. A côté des techniques classiques, des méthodes dites d'auto-étalonnage sont parfois mises en œuvre.

Elles ne nécessitent pas de référence externe et exploitent les mouvements de la caméra dans une scène statique inconnue et la mise en correspondance de points/lignes détectés entre les vues pour réaliser l'étalonnage [Peng 2010, Stoyanov 2005a].

La méthode d'étalonnage géométrique la plus répandue a été proposée par Zhang et al. [Zhang 2002] et reste une référence dans les bibliothèques de vision par ordinateur, comme OpenCV. Cette méthode repose sur l'observation de plusieurs vues d'une mire plane rigide, typiquement un damier ou un motif ChArUco, comportant des points caractéristiques facilement détectables par des algorithmes de traitement d'images, comme illustré en figure 2.24. Via des correspondances entre les coordonnées tridimensionnelles connues des points de la mire et leurs positions bidimensionnelles détectées dans les images, les paramètres du modèle sont optimisés afin de minimiser l'erreur de reprojection, c'est-à-dire la somme totale des carrés des distances entre les points détectés et les points projetés.

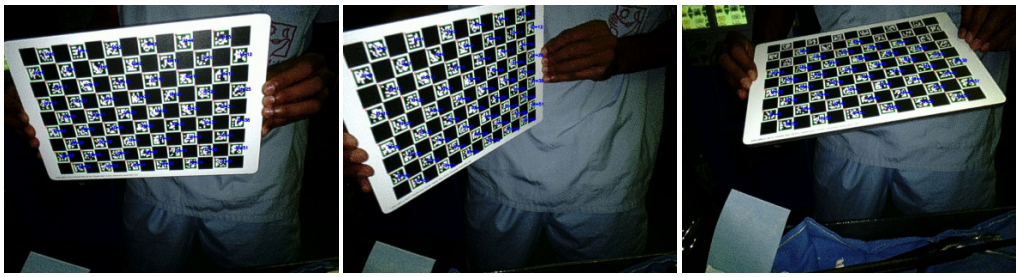


FIGURE 2.24 – Points caractéristiques détectés dans différentes vues de la mire d'étalonnage ChAruCo 9x13 DICT 7X7, taille de marqueurs 16 millimètres. Source : auteur.

Lorsqu'une caméra est équipée d'un dispositif d'autofocus, qui est un mécanisme d'ajustement automatique de la distance focale visant à assurer une netteté optimale de l'image indépendamment de la distance objet-capteur, les méthodes classiques d'étalonnage se révèlent insuffisantes. En effet, les travaux de Ricolfe-Viala et al. [Ricolfe-Viala 2021] ont montré que la variation automatique de la focale induite par l'autofocus perturbe le modèle sténopé. L'expérimentation menée avec une caméra montée sur un bras robotisé révèle que les paramètres extrinsèques estimés sont systématiquement biaisés, la position de la caméra est évaluée en moyenne cent millimètres trop proche de la cible, malgré une erreur de reprojection très faible. Ce biais provient du couplage fort entre la focale, la position de la caméra et le point principal. Les auteurs introduisent alors la notion de distance hyperfocale, définie comme la distance minimale à partir de laquelle la focale reste constante malgré un changement de distance caméra-objet. Au-delà de cette distance, l'image reste nette sans que l'autofocus n'ajuste la focale. En exploitant cette propriété, ils proposent une méthode d'étalonnage en deux étapes : (1) estimer expérimentalement les facteurs d'échelle liés à la focale à différentes distances afin d'identifier la zone de

stabilité focale ; (2) fixer ces valeurs pour effectuer l'optimisation non linéaire uniquement sur les paramètres extrinsèques, le point principal et la distorsion. Cette approche permet de limiter le couplage entre les paramètres intrinsèques et extrinsèques et d'améliorer significativement la précision de l'étalonnage pour les systèmes à focale variable.

Les systèmes stéréoscopiques, qui reposent sur l'acquisition simultanée d'une vue gauche et d'une vue droite, s'appuient fondamentalement sur la géométrie épipolaire. Elle définit les contraintes projectives reliant les points homologues entre deux images. Un point tridimensionnel $\overrightarrow{P}$ de la scène se projette en deux points $\overrightarrow{p_G}$ et $\overrightarrow{p_D}$ situés respectivement dans les plans image des caméras gauche et droite, comme illustré en figure 2.25. La position de $\overrightarrow{p_G}$ contraint la localisation possible de $\overrightarrow{p_D}$ à une droite dans l'image droite, appelée droite épipolaire, notée d_D.

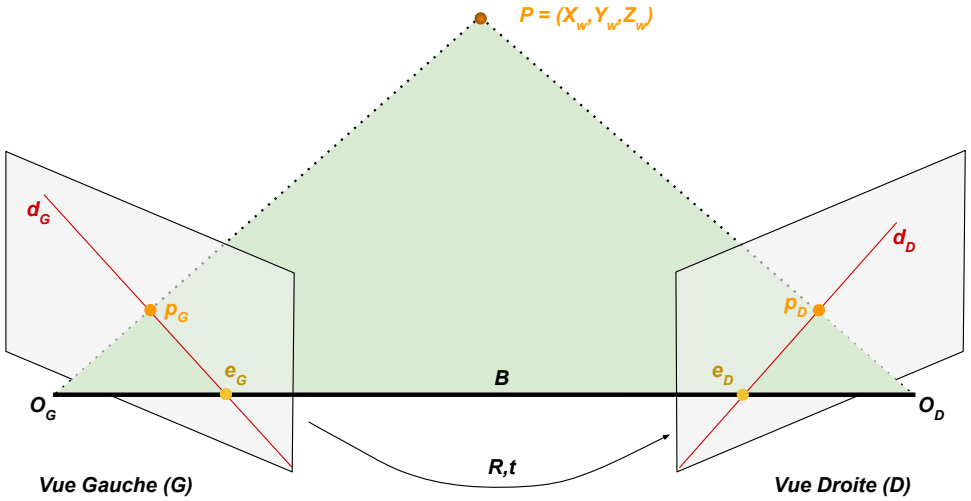


FIGURE 2.25 – Représentation schématique des contraintes épipolaires pour une caméra stéréoscopique. Un point tridimensionnel $\overrightarrow{P}$ dans le repère monde se projette sur les plans image des caméras gauche et droite en $\overrightarrow{p_G}$ et $\overrightarrow{p_D}$ respectivement. Source : auteur.

La relation épipolaire peut être formalisée par la matrice fondamentale F, une matrice 3×3 qui encode l'ensemble des contraintes épipolaires indépendamment des paramètres intrinsèques des caméras, selon :

$$\overrightarrow{p_D}^T F \overrightarrow{p_G} = 0 \tag{2.7}$$

Lorsque les matrices intrinsèques K_G et K_D des caméras gauche et droite sont connues, cette relation peut être réécrite à l'aide de la matrice essentielle E, qui intègre explicitement la pose relative des deux caméras. Elle est liée à F par la relation :

$$E = K_D^T F K_G \tag{2.8}$$

L'exploitation de la géométrie épipolaire permet de restreindre la recherche de correspondances à une seule dimension, le long des droites épipolaires. Dans le cas particulier où les caméras sont rectifiées, c'est-à-dire que leurs axes optiques sont parallèles et que leurs plans image sont coplanaires, les droites épipolaires deviennent horizontales. Cette condition est obtenue par l'application d'homographies de rectification H_G et H_D, respectivement aux vues gauche et droite. Dans ce cadre, la disparité est définie comme la différence entre les coordonnées horizontales des pixels homologues :

$$d = u_G - u_D \tag{2.9}$$

où u_G et u_D sont les abscisses des projections gauche et droite du point $\overrightarrow{P}$. La profondeur Z de ce point peut alors être déterminée géométriquement selon la relation :

$$Z = \frac{f \cdot B}{d} \tag{2.10}$$

où f est la distance focale exprimée en pixels et B la ligne de base entre les deux centres de projection O_G et O_D.

En pratique, l'estimation de la disparité dans un système stéréoscopique rectifié peut être réalisée à l'aide de méthodes géométriques comme le Block Matching (BM) ou le SGBM [Hamzah 2016, Hirschmuller 2007], qui produisent généralement des cartes de disparité clairsemées. Des méthodes plus récentes, basées sur l'apprentissage profond comme Stereo Any Video [Jing 2025], MVSAnywhere [Izquierdo 2025] ou Foundation Stereo [Wen 2025], permettent d'estimer des cartes de disparité denses.

Par ailleurs, des approches monoculaires de prédiction de profondeur ont été développées, exploitant des modèles d'apprentissage profond comme Depth Anything [Yang 2024a, Yang 2024b], UniDepthV2 [Piccinelli 2025], Video Depth Anything [Chen 2025], Depth Any Camera [Guo 2025] et Distill Any Depth [He 2025] ou encore EndoDAC [Cui 2024] et TRMDSV [Budd 2024] spécialisées pour les images chirurgicales. Dans le contexte de la CMI (cf. sections 1.2.4), ces méthodes ont été évaluées en termes d'exactitude et de robustesse aux perturbations chirurgicales (saignements, flou, projection sur l'optique) sur un jeu de données ex-vivo [Sharifian 2025b]. Les résultats montrent que malgré une bonne qualité visuelle, les méthodes monoculaires sont limitées en termes de justesse métrologique. Les méthodes spécifiques sont globalement plus robustes mais restent sensibles aux perturbations comme les projections ou le flou de mise au point.

Indépendamment de la méthode utilisée pour estimer la profondeur, un étalonnage stéréoscopique rigoureux reste indispensable. Celui-ci consiste à estimer non seulement les paramètres intrinsèques de chaque caméra, mais également la transformation rigide les reliant. Une approche usuelle consiste à étalonner individuellement chaque caméra selon une méthode monoculaire, puis à estimer les paramètres extrinsèques via l'observation conjointe d'une mire. Toutefois, comme dans le cas monoculaire, des variations non maîtrisées de la focale perturbent la stabilité des paramètres intrinsèques et extrinsèques.

Dans le domaine de l'endoscopie et plus particulièrement de la CMI (cf. sections 1.2.4), le modèle sténopé demeure généralement valable pour la modélisation géométrique des endoscopes rigides, à condition de tenir compte des spécificités optiques et mécaniques de ces systèmes [Barreto 2008, Yamaguchi 2004, Wu 2010, De Buck 2007, Liu 2017, Melo 2011]. Des approches alternatives, reposant sur la modélisation des rayons de lumière incidents sans hypothèse a priori sur la projection ou la distorsion, ont montré que les endoscopes se comportent de manière proche d'un système central. En conditions normales, voire en présence de torsion, le modèle central reste utilisable avec une précision satisfaisante. Par comparaison avec la méthode de Zhang [Zhang 1999] (implémentée notamment par Bouguet), l'erreur quadratique moyenne est réduite à 0,118 millimètre contre 0,35 millimètre. En revanche, les modèles de distorsion radiale polynomiale classiques comme ceux de Conrady [Conrady 1919] ou Brown [Brown 1996] s'avèrent peu adaptés aux optiques des endoscopes médicaux [Barreto 2008, Grossberg 2005].

La stabilité thermique de l'endoscope est un facteur souvent sous-estimé, mais déterminant. La température influe sur les paramètres intrinsèques via des micro déformations de l'assemblage. Un temps de stabilisation thermique de l'ordre de quinze minutes est souvent nécessaire. Il est donc recommandé de réaliser des contrôles réguliers de l'exactitude de l'étalonnage [Hardner 2022].

Dans le cas particulier des endoscopes obliques (30° à 45°), une rotation mécanique du tube optique autour de son axe est possible, introduisant un degré de liberté supplémentaire entre le tube et la tête de caméra. Cette rotation provoque un désaxement de l'axe optique, perceptible par un déplacement du point principal dans l'image et une variation des paramètres intrinsèques. Pour compenser ces effets, plusieurs stratégies ont été proposées, allant de la détection visuelle de repères angulaires à l'intégration de capteurs électromagnétiques ou optiques afin d'estimer l'angle de rotation en temps réel et d'ajuster dynamiquement les paramètres intrinsèques [Wu 2010, Yamaguchi 2004, De Buck 2007, Liu 2017, Melo 2011].

Dans les systèmes de télémanipulation robotisée comme le Da Vinci (Intuitive Surgical) (cf. section 1.2.4.3) ces problématiques sont atténuées car le tube optique et la caméra sont solidaires et indissociables. La rotation quasi-pure est pilotée électromécaniquement et est limitée à deux configurations : 30° vers le haut ou 30° vers le bas.

Afin de rendre l'étalonnage géométrique de l'endoscope réalisable au bloc opératoire dans un temps contraint (de l'ordre de la minute), un système d'étalonnage intéractif est généralement utilisé. Ce système combine trois modules : un module de sélection d'images qui s'appuie sur des critères comme la netteté (via la variance du Laplacien), l'angulation de l'axe optique par rapport à la normale de la mire et une distance temporelle ; un module d'étalonnage monoculaire ou stéréoscopique et un module d'évaluation pour valider la qualité de l'étalonnage à travers la bonne couverture du plan image et une erreur de reprojection inférieure à un seuil fixé en pixels (généralement un pixel). L'utilisateur peut être guidé dans ce processus, ce qui permet de garantir un étalonnage robuste, rapide et reproductible avec une erreur de reprojection typique de 0,19 pixel [Hardner 2022, Dowrick 2023].

Le tableau 2.5 présente, à titre de référence, les valeurs moyennes des paramètres d'étalonnage géométrique d'un endoscope stéréoscopique 30° Da Vinci Xi (Intuitive Surgical), issues des campagnes d'essais de faisabilité détaillées en section 2.2.3. Les distances focales ($f \simeq 1049$ pixels) présentent une faible variation relative ($< 0,5\,\%$), indiquant une bonne stabilité optique au sein d'un même dispositif comme entre différents dispositifs. Le centre optique montre une variabilité modérée (écart type d'environ dix pixels), avec une dispersion relative maximale de $4,6\,\%$. En revanche, les coefficients de distorsion radiale k_2 et k_3 affichent des écarts plus prononcés, avec des variations relatives atteignant respectivement $40,8\,\%$ et $32,1\,\%$. Du côté des paramètres extrinsèques, la composante de translation latérale T_x est remarquablement stable (coefficient de variation $CV = 0,035$).

La figure 2.26 illustre graphiquement ces variations relatives pour les principaux paramètres d'étalonnage (intrinsèques et extrinsèques), en les comparant à leur moyenne respective, pour six dispositifs différents nommés D1 à D6. Ces résultats confirment la stabilité globale de l'étalonnage dans le temps malgré une utilisation intensive et de nombreux cycles de nettoyage et stérilisation, avec des performances plus régulières observées pour les dispositifs D1, D2 et D5.

Caméra G.	fx (px)	fy (px)	cx (px)	xy (px)	k1	k2	k3
Min	1023,000	1023,000	575,000	486,000	-0,029	-0,187	-0,339
Max	1080,000	1080,000	612,000	528,000	0,013	0,080	0,352
Moyenne	1048,931	1048,931	592,724	509,517	-0,018	0,007	-0,022
Ecart-type	10,275	10,275	10,495	9,793	0,009	0,055	0,107
Var. relative	0,054	0,054	0,062	0,082	-2,329	40,796	-32,087
Coef. var.	0,010	0,010	0,018	0,019	-0,482	8,426	-4,944

Caméra D.	fx (px)	fy (px)	cx (px)	xy (px)	k1	k2	k3
Min	1026,000	1026,000	663,000	493,000	-0,028	-0,078	-0,128
Max	1061,000	1061,000	705,000	533,000	0,010	0,070	0,145
Moyenne	1048,750	1048,750	683,750	509,833	-0,019	0,020	-0,017
Ecart-type	8,818	8,818	9,821	9,955	0,009	0,038	0,077
Var. relative	0,033	0,033	0,061	0,078	-2,009	7,369	-16,457
Coef. var.	0,008	0,008	0,014	0,020	-0,451	1,884	-4,659

Extrinsèques	Rx (rad)	Ry (rad)	Rz (rad)	Tx (mm)	Ty (mm)	Tz (mm)
Min	-0,010	-0,010	0,000	-4,480	-0,220	-1,600
Max	0,020	0,010	0,000	-3,910	1,440	2,060
Moyenne	0,001	-0,001	0,000	-4,140	0,094	0,247
Ecart-type	0,005	0,004	0,000	0,146	0,316	0,790
Var. relative	36,000	-24,000	-	-0,138	17,660	14,845
Coef. var.	6,043	-4,899	-	-0,035	3,366	3,202

TABLE 2.5 – Paramètres intrinsèques et extrinsèques typiques pour un endoscope stéréoscopique 30° Da Vinci Xi (Intuitive Surgical), obtenus à partir de six dispositifs différents et 29 étalonnages par approche interactive. Source : auteur.

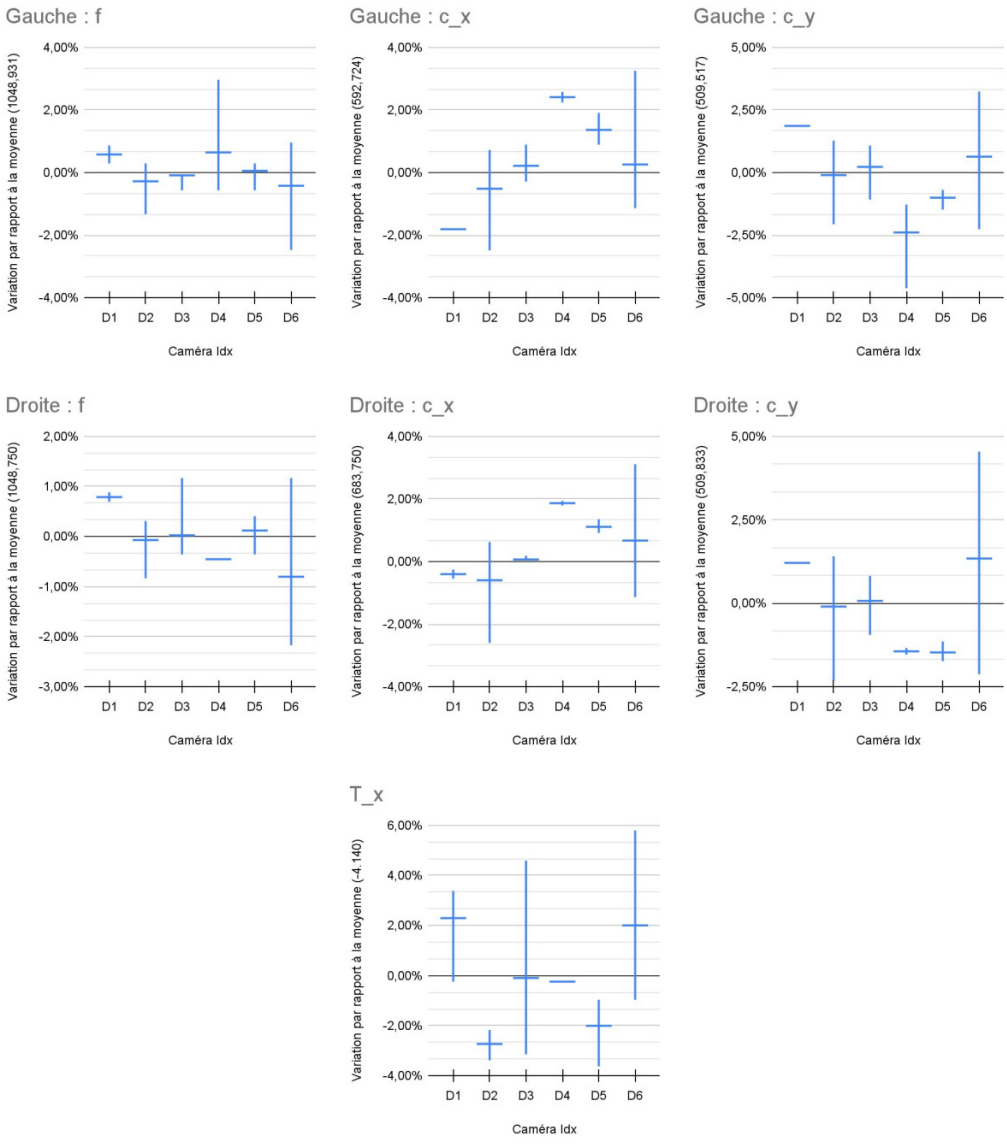


FIGURE 2.26 – Représentation graphique des variations par rapport à la moyenne de sept paramètres issus de l'étalonnage géométrique stéréoscopique de six endoscopes 30° Da Vinci Xi (Intuitive Surgical). Les paramètres considérés incluent : trois paramètres intrinsèques pour la vue gauche (f, c_x, c_y), trois paramètres intrinsèques pour la vue droite (f, c_x, c_y) ainsi qu'un paramètre extrinsèque (T_x). Source : auteur.

Chapitre 3

Synchronisation de flux vidéo gauche-droit en chirurgie endoscopique

---

Dans ce chapitre, nous nous intéressons plus particulièrement aux flux vidéo chirurgicaux stéréoscopiques, constitués de paires d'images gauche-droite capturées simultanément par une caméra endoscopique binoculaire. Bien que le système source assure un affichage synchronisé de ces paires d'images, les dispositifs d'acquisition vidéo tiers induisent fréquemment une désynchronisation gauche-droite entre les deux canaux visuels. Cette désynchronisation, d'origine matérielle ou logicielle, compromet la cohérence temporelle des paires stéréoscopiques et complique significativement les traitements ultérieurs impliquant la stéréovision, comme le calcul de cartes de disparité. Afin de remédier à ce problème, nous proposons une approche fondée sur l'exploitation de la composante de réflexion spéculaire très présente dans les images endoscopiques. Cette propriété, intrinsèquement liée aux caractéristiques optiques des tissus biologiques humides, est utilisée pour définir un SCGD. Ce score est obtenu à travers un enchaînement d'étapes comprenant la détection des taches spéculaires, leur caractérisation par des descripteurs invariants, puis leur appariement entre les vues gauche et droite. Les résultats expérimentaux démontrent la supériorité de notre méthode par rapport à 216 SCGD existants. Elle réduit la distance temporelle moyenne et maximale de synchronisation à $0,055$ et 1 image, avec un taux de synchronisation correcte de $94,1 \pm 3,6\,\%$, contre $0,3$ et 3 images et $81,2 \pm 6,4\,\%$ pour le meilleur SCGD existant. Ces performances mettent en évidence l'apport significatif de la réflexion spéculaire pour la synchronisation stéréoscopique passive en temps réel dans les systèmes d'acquisition tiers dépourvus d'informations de synchronisation explicites.

Sommaire

3.1 Introduction 117
3.2 Contexte technique général 118
3.3 Interactions lumière-matière en cœlioscopie 121
3.4 État de l'art 127
3.4.1 Méthodes de synchronisation vidéo actives 127
3.4.2 Méthodes de synchronisation vidéo passives 128
3.4.3 Perception visuelle tridimensionnelle 130
3.4.4 Détection automatique des taches spéculaires 131
3.5 Mécanisme général de synchronisation 133
3.6 Scores de comparaison Gauche-Droit 136
3.6.1 Exploitation de la composante spéculaire 136
3.6.2 Exploitation des composantes diffuse et spéculaire 139
3.7 Génération de données d'évaluation semi-synthétiques . . . 139
3.8 Expériences 142
3.8.1 Jeu de données pour l'évaluation 142
3.8.2 Mesure de la distance temporelle 143
3.8.3 Évaluation de la performance des scores proposés 145
3.8.4 Évaluation de l'impact de la segmentation des taches spéculaires sur la synchronisation 151
3.8.5 Évaluation de l'impact du score sur la synchronisation 154
3.8.6 Applications en chirurgie assistée par ordinateur 155
3.9 Conclusion 160

3.1 Introduction

Dans le contexte clinique de la CMI (cf. section 1.2.4) et plus particulièrement dans celui de la chirurgie endoscopique (cf. section 1.2.4.1), l'utilisation croissante d'endoscopes binoculaires, également appelés stéréoscopiques ou 3D dans le langage clinique, marque une évolution significative. Contrairement aux endoscopes conventionnels monoculaires (ou 2D), ces dispositifs offrent une perception tridimensionnelle de la scène opératoire, améliorant l'estimation de la profondeur et la compréhension des volumes anatomiques. Cette capacité de vision en relief facilite ainsi la gestuelle et l'ergonomie du geste chirurgical.

En chirurgie robotique (cf. section 1.2.4.3), les plateformes de télémanipulation intègrent nativement une chaîne de traitement binoculaire complète, depuis l'acquisition jusqu'à l'affichage stéréoscopique dans la console du chirurgien. Cette technologie permet une immersion visuelle optimale et une coordination œil-main améliorée (cf. section 1.2.4.3).

Les avancées en CAO et plus spécifiquement celles en CGI (cf. section 1.2.4.4), ont placé la vidéo peropératoire au cœur des dispositifs d'aide à la décision et à la navigation. De nombreux systèmes exploitent aujourd'hui le flux vidéo endoscopique en tant que source principale d'information, tant pour l'assistance en temps réel que pour l'analyse différée. C'est notamment le cas des dispositifs de RA basés sur des techniques de vision par ordinateur (cf. section 1.3.3), qui reposent entièrement sur ce flux pour superposer des informations virtuelles à la scène chirurgicale réelle. Ces applications imposent un traitement vidéo à faible latence, souvent en quasi temps réel, réalisé sur un système informatique dédié équipé d'une carte d'acquisition vidéo et dans certains cas d'un GPU pour répondre aux exigences computationnelles du traitement d'image et du rendu graphique.

La valeur ajoutée de la vidéo stéréoscopique repose entièrement sur la simultanéité de capture et d'affichage des deux images (gauche et droite) constituant la paire stéréoscopique. Lorsque celles-ci sont acquises et restituées au même instant, elles sont dites temporellement synchrones. À l'inverse, un décalage temporel qui induit une désynchronisation entre les vues gauche et droite compromet la perception stéréoscopique de l'utilisateur et peut perturber l'exploitation algorithmique de la paire d'images par un système de traitement tiers.

Ainsi, la synchronisation temporelle précise des flux binoculaires constitue un enjeu fondamental pour le développement de dispositifs de CGI exploitant la stéréovision. Elle conditionne non seulement la qualité de l'interaction visuomotrice du chirurgien avec le système mais aussi la fiabilité des analyses peropératoires ou postopératoires qui en découlent.

Dans le cadre d'un module de synchronisation passive, nous proposons d'exploiter la composante spéculaire afin de concevoir un SCGD spécifiquement adapté au contexte chirurgical. Notre SCGD spéculaire repose sur l'analyse des taches spéculaires, via un pipeline de traitement complet qui inclut leur identification, leur description ainsi que leur mise en correspondance. Nous avons validé notre approche expérimentalement en comparant la méthode proposée à 216 SCGD exis-

tants sur un jeu de données de 44 vidéos de NPRA et Résection Hépatique Robot-Assistée (RHRA) avec des désynchronisations simulées et réelles. Les résultats obtenus corroborent l'hypothèse selon laquelle l'exploitation conjointe des propriétés dynamiques des reflets spéculaires et des informations issues des composantes diffuses constitue une solution efficace au problème de la synchronisation stéréoscopique en chirurgie.

3.2 Contexte technique général

Les systèmes sources de flux vidéo, tels que les colonnes de cœlioscopie ou les télémanipulateurs chirurgicaux, assurent à la fois la captation de la scène intra-abdominale, son traitement primaire en vue d'un affichage ainsi que la mise à disposition des flux vidéo via des interfaces numériques normalisées (par exemple BNC, HDMI type A ou DVI-D) (cf. section 2.2.5). La lumière réfléchie par les objets de la scène est focalisée par un ensemble optique et transmise à un capteur d'image constitué d'un réseau de photosites, généralement disposés selon une matrice de Bayer (deux verts, un rouge et un bleu) qui permet la reconstruction colorimétrique d'une image par interpolation.

Le flux de photons est converti en intensité électrique via l'effet photoélectrique. Le signal analogique est ensuite amplifié et numérisé avant d'être intégré dans une chaîne de traitement vidéo numérique. La fréquence d'acquisition, généralement comprise entre 20 Hz et 60 Hz, est déterminée par un déclenchement régulé par l'unité de contrôle de la caméra ou CCU en Anglais (cf. section 1.2.4.1).

Dans la chaîne de traitement interne, les images sont horodatées afin de garantir un alignement temporel précis. Cette condition est essentielle au bon fonctionnement des systèmes stéréoscopiques, pour lesquels la cohérence temporelle gauche-droite ne peut souffrir d'approximation. Dans certains cas, les systèmes source intègrent également leur propre dispositif d'affichage comme la console chirurgien utilisée en chirurgie robotique (cf. section 1.2.4.3).

Deux contraintes techniques majeures régissent cette chaîne interne de traitement numérique. La première concerne la qualité d'image, qui dépend à la fois des propriétés optique du système, des caractéristiques du capteur (résolution spatiale, fidélité colorimétrique, sensibilité, rapport signal/bruit) et des paramètres d'encodage (taux de compression, débit binaire, profondeur de couleur). La seconde est la latence, qu'il est crucial de minimiser afin de garantir une adéquation entre le geste effectué et la perception visuelle de ce dernier par le chirurgien.

Une fois générées par la source, les données sont ensuite acheminées vers les différents équipements du bloc opératoire qui utilisent le flux visuel, comme les écrans chirurgicaux (aussi appelés écrans de rappel), les systèmes d'enregistrement, les serveurs de diffusion et les autres dispositifs tiers. Cette distribution repose quasi exclusivement sur des liaisons filaires, gage de stabilité, de débit soutenu et de fiabilité. Toutefois, le cheminement du signal est rarement direct. Il transite par une chaîne de traitement externe composée d'un ensemble hétérogène de dispositifs intermédiaires

assurant des fonctions de conversion, de duplication, d'adaptation de support ou d'amplification. Parmi ces composants, figurent notamment des distributeurs vidéo, des convertisseurs de formats, des transcodeurs fibre/cuivre, des buffers ainsi que des amplificateurs de signaux. Bien qu'indispensables au bon routage des données visuelles, ces dispositifs peuvent introduire des latences différentielles et dans le cas d'un flux multicanal, provoquer des désynchronisations temporelles inter-canaux, susceptibles de compromettre la restitution stéréoscopique.

À ce stade, le système source n'est plus en mesure d'assurer le maintien de l'intégrité temporelle et visuelle des flux. Cette responsabilité est alors transférée à la chaîne de traitement externe, dont la synchronisation devient un enjeu critique. En effet, la rupture du lien avec l'horodatage natif introduit une incertitude sur l'instant exact de captation des images, ouvrant la voie à des incohérences temporelles. Certaines stratégies de formatage des signaux binoculaires comme le Frame Packing (cf. section 2.2.5) permettent de préserver une synchronisation implicite lors du transport du signal entre la source et le dispositif de destination.

La difficulté s'intensifie lorsque les canaux gauche et droit sont transmis indépendamment, comme cela est le cas pour les télémanipulateurs chirurgicaux Da Vinci (Intuitive Surgical). Des distances temporelles fixes ou variables, systématiques ou aléatoires, peuvent apparaître entre les flux. Ces désynchronisations résultent de multiples facteurs, incluant les spécificités matérielles des cartes d'acquisition, les mécanismes internes de buffering ou encore les stratégies de parallélisation logicielle (voir figure 3.1). Ces écarts temporels entre canaux peuvent sérieusement altérer le fonctionnement des algorithmes de vision par ordinateur, notamment ceux fondés sur la stéréovision, pour lesquels une synchronisation rigoureuse est indispensable entre les deux vues.

À titre illustratif, les systèmes d'enregistrement embarqués bi-canaux comme le MVR Pro (Medicapture), qui s'appuient sur des architectures multithreadées, cumulent deux sources majeures de désynchronisation : (1) un décalage initial, introduisant un offset constant entre les deux canaux et (2) une dérive temporelle variable. Le premier résulte principalement de la latence inhérente à l'initialisation asynchrone des threads, tandis que le second découle des variations d'ordonnancement au sein du logiciel et des limitations de la bande passante ou de la mémoire tampon au niveau matériel.

Les systèmes tiers chargés d'exploiter ces flux, en dehors de certains environnements industriels où des mécanismes explicites de synchronisation sont disponibles, sont généralement dépourvus de toute information temporelle intégrée, tel qu'un signal d'horloge global ou des métadonnées d'horodatage associées à chaque image. Déconnectés de toute référence temporelle fiable, ces dispositifs doivent s'appuyer exclusivement sur le contenu visuel brut des images pour tenter d'estimer ou de compenser un éventuel désalignement temporel entre les vues gauche et droite. Ce type de synchronisation "aveugle" s'avère particulièrement complexe à implémenter, car elle dépend fortement du contenu de la scène, des dynamiques d'acquisition et des propriétés des algorithmes de traitement. En l'absence d'une synchronisation native, la mise en œuvre de traitements aval comme la génération de cartes de dis-

parité ou la reconstruction tridimensionnelle peropératoire devient plus fragile, voire inopérante en cas de désynchronisation marquée.

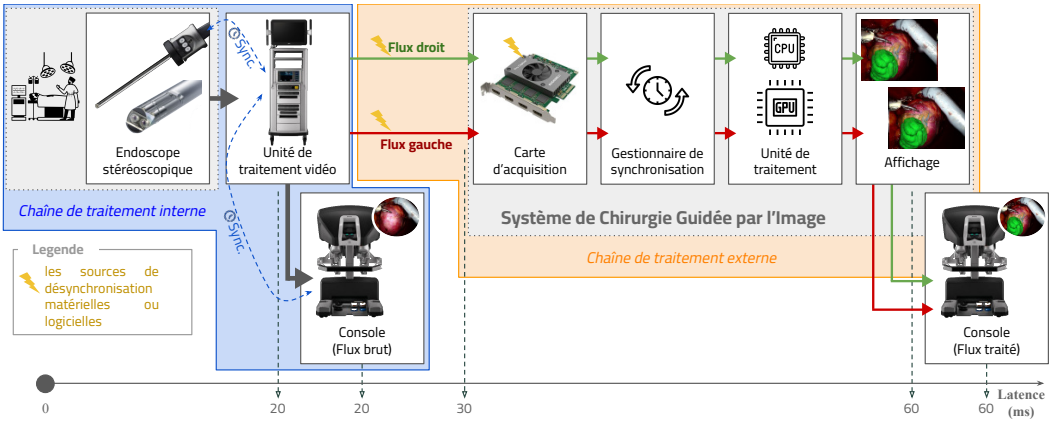


FIGURE 3.1 – Représentation schématique du flux de données chirurgicales, depuis l'endoscope stéréoscopique jusqu'à la console chirurgicale. Le flux initialement synchrone est exploité au sein d'un système de Chirurgie Guidée par l'Image (CGI). L'introduction de composants matériels ou logiciels dans la chaîne de traitement externe est susceptible d'engendrer des phénomènes de désynchronisation. Source : auteur.

La complexité de cette problématique est accrue par les contraintes réglementaires et normatives du bloc opératoire (cf. section 1.2.2.2). Les dispositifs d'acquisition comme les caméras endoscopiques doivent satisfaire des exigences strictes en matière de sécurité électrique, de biocompatibilité, de stérilité et être certifiés en tant que dispositifs médicaux au sens du règlement européen (MDR 2017/745) ou conformément aux exigences d'agences nationales telles que la Food and Drug Administration (FDA) ou l'Agência Nacional de VIgilância SAnitária (ANVISA). Ces obligations restreignent considérablement toute possibilité de modification à posteriori du matériel, y compris l'ajout de composants destinés à la synchronisation. En effet, l'intégration de capteurs ou modules externes dépourvus de finalité clinique directe soulève des problématiques réglementaires importantes et s'avère, dans la majorité des cas, incompatible avec les protocoles d'asepsie.

Dans ce contexte, les stratégies de synchronisation applicables à la chaîne de traitement externe doivent impérativement être non invasives, afin de ne pas compromettre l'intégrité du dispositif médical déjà en place ni perturber les conditions opératoires. Elles reposent ainsi sur des approches dites passives, basées sur l'analyse du flux vidéo en aval, postérieure à la capture. Ces méthodes permettent de détecter et si nécessaire de corriger les désalignements temporels entre les canaux sans modifier l'infrastructure matérielle d'acquisition, ni introduire de signaux auxiliaires susceptibles d'interférer avec le système existant. Il est également essentiel que ces techniques préservent la qualité visuelle et la faible latence du signal original, de

manière à ne pas altérer l'affichage peropératoire ni dégrader les performances des traitements d'image intégrés aux systèmes de CGI.

Ainsi, la synchronisation stéréoscopique en environnement chirurgical présente des défis spécifiques, étroitement liés à la structure visuelle des images acquises. Elle devient particulièrement délicate dans les situations où (C1) la scène demeure quasi-statique, limitant les opportunités d'identifier des correspondances fiables entre les vues gauche et droite en l'absence de mouvement exploitable. Elle est également fragilisée par (C2) les déplacements rapides de la caméra, qui introduisent des discontinuités temporelles marquées, ainsi que par (C3) les phénomènes de réflexion complexes résultant des interactions lumière-matière au niveau des instruments chirurgicaux et des tissus biologiques humides.

3.3 Interactions lumière-matière en cœlioscopie

La synchronisation passive de flux stéréoscopiques en environnement chirurgical, fondée sur l'analyse directe du contenu de l'image, impose une compréhension des mécanismes de formation des images sur la rétine d'un capteur optoélectronique. En cœlioscopie, contrairement à l'imagerie en champ ouvert, l'environnement est optiquement contrôlé. La cavité intra-abdominale explorée est intrinsèquement close et dépourvue d'éclairage ambiant. L'éclairage artificiel est assuré par une source polychromatique unique, colocalisée avec l'optique, qui peut être modélisée par un faisceau lumineux focalisé de type spot dont l'intensité lumineuse décroît radialement selon la distribution photométrique propre à l'optique d'émission suivant un profile Illuminating Engineering Society (IES) en Anglais donné.

Dans ce contexte, l'interaction entre les photons incidents et les structures biologiques suit les principes de l'optique physique. Le comportement optique des tissus dépend des propriétés directionnelles et spectrales du faisceau ainsi que de la constitution physico-chimique du matériau biologique irradié. À l'interface, la lumière peut être réfléchie, diffusée ou réfractée. Une partie est absorbée par les chromophores intracellulaires tels que l'hémoglobine ou la mélanine et une autre partie est transmise ou diffusée en profondeur. Ces interactions sont illustrées schématiquement en figure 3.2.

L'intensité lumineuse enregistrée par chaque pixel du capteur résulte de la combinaison des composantes spéculaire et diffuse de la réflexion. La réflexion spéculaire est très sensible à la géométrie locale et à l'angle de vue. Elle est prédominante sur les surfaces humides ou brillantes, typiques des environnements chirurgicaux. Elle induit des pics d'intensité souvent proches de la saturation du capteur, ce sont les taches spéculaires visibles en figure 3.4. En parallèle, la réflexion diffuse, plus isotrope, véhicule l'essentiel de l'information chromatique des textures superficielles. Elle constitue le principal vecteur de la perception des couleurs.

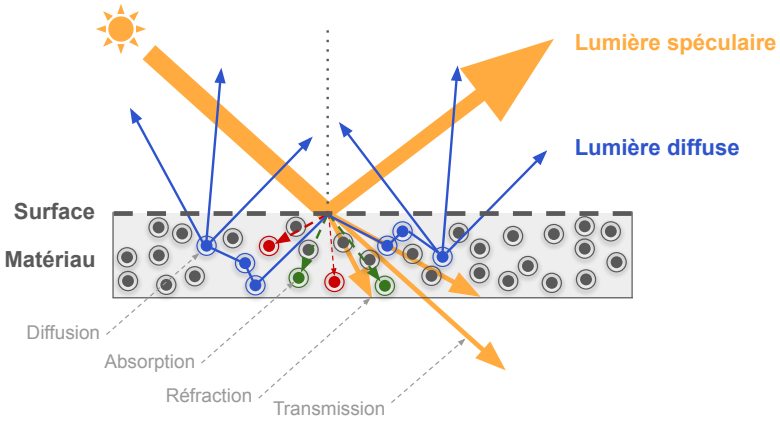


FIGURE 3.2 – Représentation schématique des interactions lumière-matière (réflexion, diffusion, absorption, réfraction et transmission) dans un échantillon de matière soumis à un faisceau de lumière incident. Source : auteur.

Pour modéliser les phénomènes perçus, résultants à la formation de ces images, il est possible de recourir à des modèles de rendu basés sur la physique ou Physically-Based Rendering (PBR) en Anglais. Ces derniers remplacent avantageusement les modèles de lumière empiriques tels que Blinn-Phong. L'approche PBR repose sur trois piliers fondamentaux : un modèle de surface microstructurée (microfacettes), la conservation de l'énergie lumineuse et l'usage d'une fonction de distribution de réflectance bidirectionnelle (Bidirectional Reflectance Distribution Function (BRDF) en Anglais) qui décrit la répartition de l'intensité réfléchie selon les directions d'incidence et d'observation pour un point p :

$$L_s(p, \omega_s) = \int_\Omega f_r(p, \omega_i, \omega_s) L_i(p, \omega_i)(n \cdot \omega_i) d\omega_i \tag{3.1}$$

où L_s désigne la radiance sortante, L_i la radiance incidente, ω_i, ω_s les directions d'illumination et de vision respectivement, n la normale locale de surface et f_r la BRDF.

Les matériaux biologiques présentent une surface optiquement irrégulière à l'échelle microscopique. Dans ce cadre, la théorie des microfacettes constitue un paradigme central. Elle postule que toute surface peut être modélisée par une collection de miroirs microscopiques. Chaque microfacette est orientée selon une distribution régie par la rugosité de la surface, caractérisée par un paramètre scalaire $\alpha \in [0, 1]$. Plus celle-ci est rugueuse, plus les facettes sont désalignées, entraînant une dispersion accrue de la réflexion spéculaire. À l'inverse, les surfaces lisses produisent des reflets plus directionnels et concentrés. Le vecteur demi-angle noté h, médian entre la direction de la lumière incidente l et celle de l'observateur v, est défini comme :

$$h = \frac{l + v}{\|l + v\|} \tag{3.2}$$

Le modèle spéculaire s'exprime par la combinaison de trois termes : $D(h)$ une fonction de distribution des microfacettes évaluée au demi-vecteur h, $G(l, v, h)$ une fonction de géométrie et une fonction de Fresnel $F(l, h)$, souvent modélisée en infographie par l'approximation de Schlick :

$$f_r(l, v) = \frac{F(l, h)D(h)G(l, v, h)}{4(n \cdot l)(n \cdot v)} \tag{3.3}$$

avec :

$$F_{Schlick}(F_0, l, h) = F_0 + (1 - F_0)(1 - (l \cdot h))^5 \tag{3.4}$$

où F_0 est l'indice de réflectance caractéristique du matériau.

La seule réflexion spéculaire ne suffit cependant pas à décrire les matériaux biologiques. Les modèles de BRDF diffus étendus, comme Oren–Nayar, permettent de modéliser la réflexion diffuse sur des surfaces rugueuses en tenant compte de l'auto-ombrage microscopique, contrairement au modèle Lambertien.

En complément, l'absorption et la diffusion interne sont prédominantes dans les tissus biologiques. Ces phénomènes peuvent être modélisés par la Bidirectional Surface Scattering Reflectance Distribution Function (BSSRDF) en Anglais, qui étend la BRDF à des interactions non-locales :

$$S(x_i, \omega_i, x_s, \omega_s) = \frac{dL_s(x_s, \omega_s)}{d\Phi_i(x_i, \omega_i)} \tag{3.5}$$

où $S(x_i, \omega_i, x_s, \omega_s)$ est la BSSRDF, $L_s(x_s, \omega_s)$ la radiance sortante et $\Phi_i(x_i, \omega_i)$ le flux incident.

La BSSRDF [Nicodemus 1977] tient compte de la distance entre les points d'entrée x_i et de sortie x_s, ce qui permet de décrire la diffusion sous la surface par transluminescence (SubSurface Scattering (SSS) en Anglais) notamment dans les tissus tels que la peau ou les muqueuses [Baranoski 2010].

Les tissus humains sont des milieux stratifiés, hétérogènes et hautement diffusifs [Welch 2011]. Cette structure peut être modélisée comme une superposition de couches (slabs en Anglais) [Haase 1992, Dorsey 2006] visibles en figure 3.3, chacune caractérisée par une fonction Bidirectional Scattering Distribution Function (BSDF) à ses interfaces et une Volume Distribution Function (VDF) en Anglais pour la propagation interne de la lumière. Une couche est définie par :

$$s = Slab(f, V) \tag{3.6}$$

où f est la BSDF de surface et V le médium volumique contenant des coefficients d'absorption σ_a, de diffusion σ_d et une fonction de phase $p(\omega_i, \omega_s)$. Le transport de la lumière au sein du matériau résulte d'un couplage entre réflexion, transmission et propagation interne selon :

$$T(\omega_i, \omega_s) = \int_V p(\omega_i, \omega_s) e^{-\tau(x)} dx \tag{3.7}$$

où $\tau(x)$ est la transmittance volumique.

Le modèle OpenPBR formalise cette architecture en couches comme une composition hiérarchique de BSDF, incluant des substrats diélectriques ou métalliques, des revêtements et des micro-structures superficielles, chacune décrite par une BSDF propre et soumise aux contraintes de conservation énergétique et de symétrie de réciprocité. Un représentation schématique du matériau physique idéalisé par le modèle OpenPBR est donnée en figure 3.3.

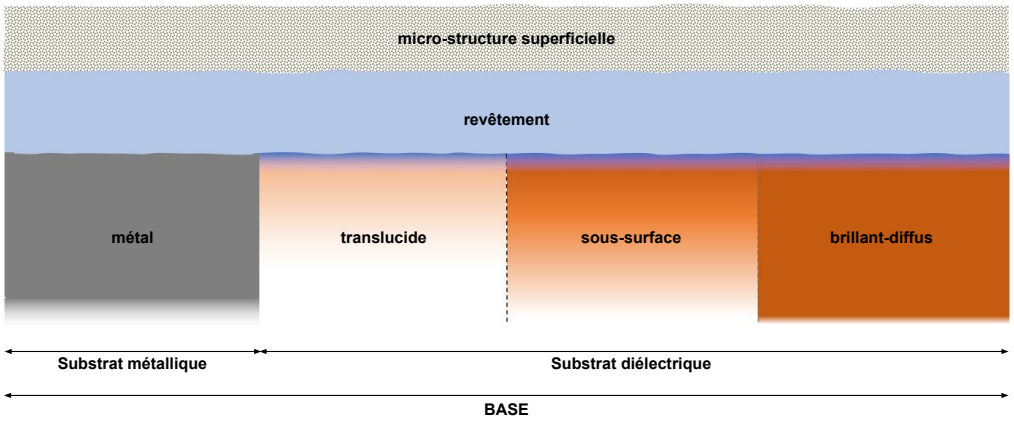


FIGURE 3.3 – Représentation schématique du matériau physique idéalisé par le modèle OpenPBR. La juxtaposition horizontale des couches représente le mélange statistique et l'empilement vertical représente la stratification. Source : générée par l'auteur à partir d'OpenPBR.

Cette modélisation permet de simuler visuellement des scènes chirurgicales via l'utilisation d'un moteur de rendu physique, tels que ceux basés sur le modèle Disney Principled BSDF ou son extension dans OpenPBR. Un exemple est donné en figure 3.4.

Ainsi, les phénomènes d'interaction lumière-matière sont des éléments discriminants majeurs dans l'analyse des images chirurgicales. La réflexion spéculaire, notamment, constitue un marqueur géométrique sensible à la position relative de la caméra et de la source d'éclairage ainsi qu'à la microtopographie des surfaces. Dans le cas de systèmes endoscopiques binoculaires à faible ligne de base (4 à 5 millimètres d'écartement entre les caméras gauche et droite) et optiques quasi-parallèles, les taches spéculaires apparaissent quasi-identiques dans les deux vues, tant en forme qu'en position. Cette propriété sera exploitée dans notre approche de synchronisation passive. Des exemples issus d'acquisitions réelles et de simulations sont présentés en figure 3.5.

Il est possible d'observer en figure 3.5, que pour une faible ligne de base (4 millimètres), les taches spéculaires apparaissent quasi-identiques dans les vues gauche et droite, tant en forme qu'en position, conformément aux acquisitions réelles. En revanche, lorsque la ligne de base est élargie (40 millimètres), un décalage perceptible des taches spéculaire apparaît.

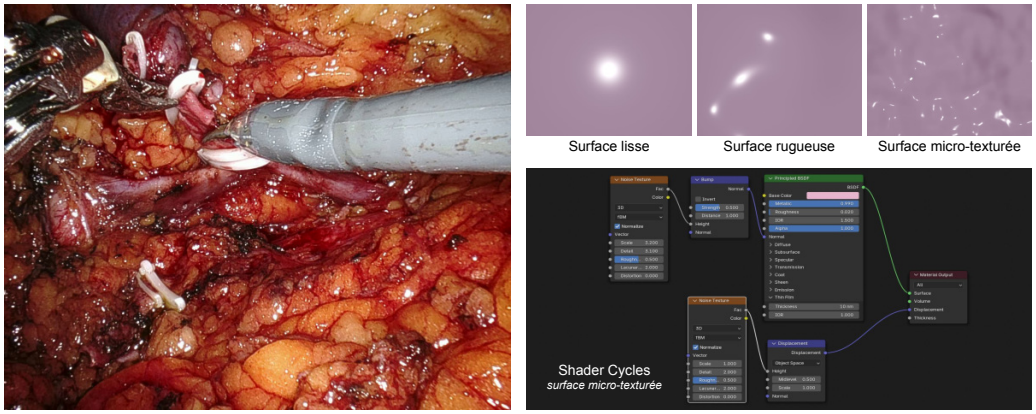


FIGURE 3.4 – (à gauche) Image peropératoire capturée par une caméra endoscopique Da Vinci (Intuitive Surgical). (à droite) Rendus graphiques générés à l'aide du moteur de rendu PBR Cycles (Blender Foundation), illustrant la perception de la composante spéculaire modélisée par une BSDF (Bidirectional Scattering Distribution Function) sur trois types de surfaces : lisse, rugueuse et micro-texturée. Source : auteur.

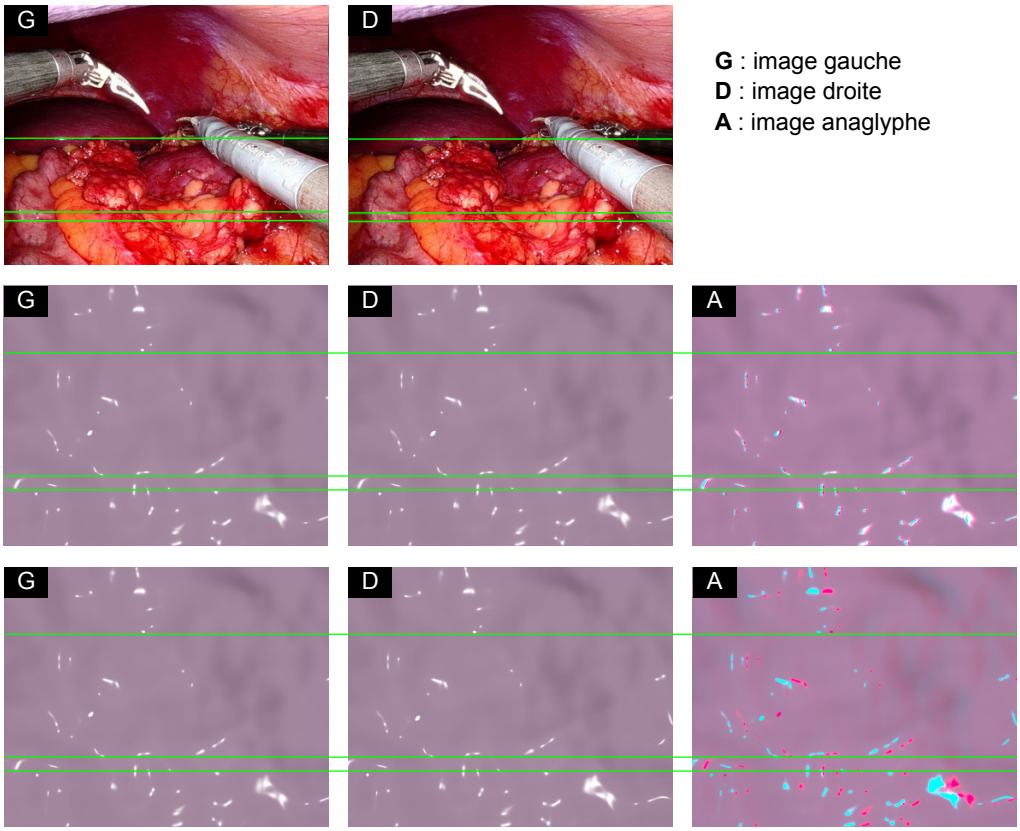


FIGURE 3.5 – (en haut) Paire stéréoscopique capturée par une caméra endoscopique Da Vinci (Intuitive Surgical) de ligne de base 4 millimètres, optiques quasi-parallèles. (Au milieu et en bas) Rendus graphiques générés à l'aide du moteur de rendu PBR Cycles (Blender Foundation), illustrant la perception de la composante spéculaire modélisée par une BSDF (Bidirectional Scattering Distribution Function). Dans les simulations, une lumière de type spot, colocalisée avec la caméra et placée à environ 3 centimètres de la surface, éclaire une zone hautement réfléchissante afin de reproduire l'aspect humide des tissus (paramètres BSDF : composant métallique 0,99 ; rugosité 0,02). La microtopographie tissulaire est simulée par micro-déplacements aléatoires via bump mapping. Ces rendus simulent une acquisition stéréoscopique avec une ligne de base de 4 millimètres (milieu) et de 40 millimètres (bas). Source : auteur.

3.4 État de l'art

Dans cette section, nous nous intéresserons dans un premier temps à présenter les approches de synchronisation de flux vidéo issues de l'état de l'art. Ces dernières se répartissent en deux grandes catégories : les méthodes actives qui reposent sur l'intégration de dispositifs matériels dédiés à la synchronisation et les méthodes passives qui exploitent l'analyse du contenu visuel des images pour estimer les désalignements temporels.

Dans un second temps, nous donnerons des éléments de la littérature scientifique portant sur les sujets de perception visuelle tridimensionnelle et de détection automatique des taches spéculaires dans les images.

3.4.1 Méthodes de synchronisation vidéo actives

La synchronisation active de flux vidéo repose sur l'introduction de références temporelles externes, indispensables à l'estimation et à la correction des décalages temporels entre séquences. Plusieurs stratégies ont été proposées dans la littérature pour répondre à cette problématique.

Certaines approches exploitent des stimuli lumineux visibles dans le champ des caméras, tels que des sources binaires ou des motifs optiques encodés. Les dispositifs de modulation temporelle d'illumination, notamment les matrices de DEL avec un codage de type Manchester [Hou 2013, Ding 2016, Hou 2010] ou les signaux stroboscopiques [Shen 2017], permettent de générer des séquences visuelles temporellement marquées jouant le rôle d'horloge optique [Zhao 2009, Mehrabi 2015].

D'autres méthodes s'appuient sur des objets oscillants pour créer une référence temporelle exploitable. L'utilisation d'un pendule [Pribanic 2015] ou d'un spot lumineux oscillant [Kim 2021] permet une synchronisation avec une précision inférieure à la période entre deux images. Dans le cas du pendule, le temps de passage de la masse par la position d'équilibre est estimé pour chaque caméra, la différence de ces instants fournit une estimation précise du décalage temporel.

L'alignement spatio-temporel peut également être fondé sur des données de géolocalisation. L'utilisation du Global Positioning System (GPS) en Anglais a ainsi été proposée pour synchroniser des séquences capturées par des caméras dont les champs de vision ne se recouvrent pas [Froner 2016]. Dans ce cas, les coordonnées spatio-temporelles d'un objet mobile sont exploitées pour reconstruire sa trajectoire et son accélération, lesquelles sont ensuite comparées aux données issues de la détection visuelle dans les vues caméra afin d'aligner temporellement les séquences issues de capteurs hétérogènes. Des architectures synchrones basées sur des topologies maître-esclave permettent également de synchroniser avec précision des caméras via des échanges de données d'horodatage (timestamps en Anglais) [Akhmetyanov 2021] par l'intermédiaire de protocoles standards tels que Universal Serial Bus (USB) en Anglais [Sousa 2015], Precision Time Protocol (PTP) en Anglais [Belhaoua 2016] ou encore General Purpose Input/Output (GPIO) en Anglais [Chen 2020].

Cependant, dans des environnements contraints tel que le bloc opératoire, l'introduction de nouveaux dispositifs matériels est fortement restreinte par les exigences réglementaires et normatives (cf. section 3.2). La modification des équipements existants est généralement proscrite. Les solutions de synchronisation doivent impérativement reposer sur des sources de données déjà présentes dans l'environnement, excluant de facto toute approche nécessitant un câblage ou un ajout de matériel dédié.

3.4.2 Méthodes de synchronisation vidéo passives

La synchronisation passive apparaît comme la stratégie la plus adaptée au contexte chirurgical car elle repose exclusivement sur l'analyse du contenu visuel des images, sans recourir à des signaux ou dispositifs externes. De nombreuses approches, tous domaines confondus, ont été proposées dans la littérature scientifique sous la forme de SCGD. Leur pertinence en milieu chirurgical est évaluée au regard des trois défis (C1), (C2) et (C3) détaillés en section 3.2.

Ces méthodes se déclinent en deux sous-catégories : celles qui exploitent une seule image et celles qui utilisent une séquence d'images. En d'autres termes, en présence ou en l'absence d'informations temporelles explicites.

Les méthodes basées sur des images isolées utilisent généralement des descripteurs visuels globaux ou locaux. Parmi les approches globales, il est possible de retrouver les matrices d'auto-similarité (Self-Similarity Matrix (SSM) en Anglais) [Dexter 2009], l'analyse fréquentielle et la corrélation de phase [Dai 2006b, Dai 2006a] ou encore l'empreinte numérique de l'image dérivée de cartes de caractéristiques extraites par des réseaux neurones profonds tels que VGG19 [Simonyan 2014b] ou Inception ResNet V2 [Szegedy 2017]. Des embeddings liés à des tâches spécifiques, comme la détection de silhouettes humaines [Yin 2022] peuvent également être utilisés. D'autres méthodes s'appuient sur des données géométriques issues de la géométrie épipolaire [Pooley 2003] ou sur la corrélation d'histogrammes [Polok 2015].

Les descripteurs locaux reposent, quant à eux, sur l'identification et la description de points clés à l'aide de techniques comme Scale Invariant Feature Transform (SIFT) en Anglais [Melloni 2015, Diego 2013, Huang 2019], des structures géométriques telles que les quads [Evangelidis 2011] ou encore des silhouettes humaines [Sinha 2004, Eichler 2022, Wang 2014a], parfois couplées à des contraintes épipolaires ou trifocales [Lei 2006, Whitehead 2005]. Dans ces approches, le problème de synchronisation équivaut à une tâche de recherche d'image de référence dans une base de données. Des techniques d'appariement exhaustives ou hiérarchisées (KD-Tree par exemple) sont alors employées, suivies d'une sélection du meilleur candidat via un système de vote ou un algorithme de plus court chemin comme celui de Dijkstra [Tripathi 2017], avec si besoin un raffinement spatio-temporel par des méthodes telles que Enhanced Correlation Coefficient (ECC) en Anglais [Evangelidis 2008].

Certaines approches considèrent la synchronisation comme une tâche de classification du désalignement temporel au sein d'une paire d'image, selon des classes telles que "MuchFaster", "Faster", "Regular", "Slower" et "MuchSlower". Les auteurs ont montré une précision de 98,54 % sur des scènes synthétiques simples [Kulawik 2021].

Les méthodes qui exploitent des séquences d'images, intègrent une dynamique temporelle en analysant par exemple les variations de luminosité induites par des changements environnementaux (tels que des flashs d'appareils photo lors de manifestations) [Caspi 2000, Ushizaki 2006, Shrestha 2006] ou par des artefacts intrinsèques au capteur comme l'effet rolling shutter [Bradley 2009, Smid 2019]. D'autres approches s'appuient sur le mouvement de la scène, d'objets ou d'acteurs humains, à travers le suivi de points clés (SIFT [Lowe 2004], SURF [Bay 2006]) ou de trajectoires 2D/3D observées par plusieurs caméras [Caspi 2002b, Sinha 2010, Albl 2017, Ukrainitz 2006, Raguse 2009, Goldmann 2010, Brito 2011, Plotz 2012, Kang 2002, Tresadern 2003, Whitehead 2005, Kang 2005, Wedge 2005, Wedge 2006, Caspi 2006, Meyer 2008, Liu 2012, Lu 2012, Shankar 2013b, Kuo 2013, Xie 2014, Benrhaiem 2015, Gaspar 2016, Morteza Safdarnejad 2017, Dhaya 2020, Shankar 2013a, Cao 2010, Wedge 2007, Cao 2006].

Lorsque les caméras sont rigidement liées ou déplacées conjointement, des homographies peuvent être estimées entre images successives [Caspi 2002a, Shakil 2006, Imre 2012] pour l'analyse du mouvement. Si les caméras sont indépendantes, la reconstruction de trajectoires tridimensionnelles s'effectue souvent par combinaison linéaire de trajectoires de bases Discrete Cosine Transform (DCT) en Anglais [Wang 2016b]. Les relations temporelles inter-caméras sont évaluées via des métriques de translation et rotation, représentées sous forme de matrices de corrélation ou de distance [Spencer 2004].

L'alignement temporel de séquences d'images peut être modélisé comme une tâche de comparaison de séries temporelles, pour laquelle des algorithmes tels que le Dynamic Time Warping (DTW) en Anglais [Berndt 1994, Tuytelaars 2004, Wang 2014b, Wieschollek 2017] sont fréquemment utilisés. Certaines méthodes exploitent des profils d'activité ou des cycles d'action [Lee 2013, Purushwalkam 2020], voire des variations dans le taux de compression binaire du flux vidéo [Schweiger 2012, Al-Nuaimi 2012] pour réaliser l'alignement. Les approches récentes à base d'apprentissage profond proposent des fonctions de coût telles que Soft-DTW, enrichies d'un terme de régularisation, atteignant des taux d'alignement compris entre 80 % et 85 % [Haresh 2021].

La majorité de ces techniques supposent des caméras fixes ou se déplaçant de manière conjointe. Un recouvrement visuel partiel est généralement requis pour une synchronisation efficace, sans pour autant imposer de fortes contraintes sur la scène ou le mouvement des objets. Par ailleurs, certaines méthodes nécessitent un étalonnage géométrique préalable des caméras (cf. section 2.2.6).

Cependant, l'ensemble de ces approches repose sur la réflexion diffuse, ce qui les rend particulièrement sensibles aux conditions spécifiques de l'environnement chirurgical. En particulier, les défis (C2) et (C3) mettent en échec la robustesse des points clés face aux reflets spéculaires et au flou de mouvement. L'absence de mou-

vement de la caméra (C1) rend également inopérantes la plupart des méthodes à image unique. Si les approches séquentielles offrent une meilleure résilience, leur efficacité décroît dans des environnements à éclairage constant, comme en endoscopie. De plus, certaines méthodes, telles que l'analyse du débit de compression, présentent une complexité computationnelle élevée, limitant leur applicabilité en temps réel.

Face à ces limitations, nous proposons le premier SCGD combinant les propriétés de réflexion diffuse et spéculaire, spécifiquement conçu pour répondre aux exigences de la synchronisation stéréoscopique en contexte chirurgical. Contrairement aux solutions existantes, cette approche surmonte simultanément les trois défis identifiés, tout en s'intégrant de manière transparente aux dispositifs du bloc opératoire.

3.4.3 Perception visuelle tridimensionnelle

La perception tridimensionnelle constitue une aptitude cognitive essentielle qui permet à l'être humain d'interpréter la profondeur, la forme et la distance des objets dans l'espace. Elle repose sur l'intégration fine de multiples indices visuels, tels que la stéréopsie (disparités entre les images rétiniennes issues de chaque œil), le mouvement, la convergence oculaire ainsi que les repères contextuels [Ijsselsteijn 2005, Winkler 2013].

Le système visuel humain exploite principalement la disparité binoculaire pour reconstruire une représentation de la profondeur. Toutefois, cette reconstruction optique et mentale peut être perturbée lorsque des signaux visuels contradictoires sont envoyés, comme cela est le cas dans les systèmes stéréoscopiques mal étalonnés. Un désalignement temporel entre les vues gauche et droite de ± 2 images demeure généralement acceptable en termes de qualité perçue ; au-delà de ce seuil, la désynchronisation impacte significativement la qualité de l'expérience ou Quality of Experience (QoE) en Anglais [Goldmann 2010].

Une exposition prolongée à des environnements immersifs tridimensionnels peut engendrer une fatigue visuelle, souvent désignée sous le terme Anglais de Digital eye Strain (DeS). Les dispositifs de RV sont particulièrement susceptibles d'induire un inconfort visuel du fait d'une dissociation entre l'accommodation et la vergence, phénomène connu sous le nom de conflit vergence-accommodation ou Vergence-Accommodation Conflict (VAC) en Anglais [Hirzle 2022, Segura 2018, Winkler 2013]. En effet, l'accommodation reste focalisée sur le plan de l'écran, tandis que la vergence s'ajuste à la profondeur virtuelle perçue [Ijsselsteijn 2005]. Cette inadéquation peut provoquer divers symptômes tels que des céphalées, une sécheresse oculaire, une vision floue ou un clignement excessif. Ces inconforts trouvent leurs origines dans des distorsions géométriques ou des artéfacts visuels (par exemple les effets de crosstalk ou de picket fence). Dans des contextes à forte exigence visuelle comme la chirurgie robotique (cf. section 1.2.4.3), ces limitations deviennent particulièrement critiques. Plusieurs études rapportent une prévalence élevée d'inconforts oculaires (33,9 %) et musculosquelettiques (74,6 %) chez les chirurgiens utilisant des systèmes de télémanipulation robotisée [Lee 2017, Winkler 2013]. Les systèmes de dernière génération ont toutefois permis des avancées notables, offrant non seule-

ment une amélioration de la visualisation chirurgicale, mais également un confort accru pour l'opérateur [Molle 2023].

Les environnements immersifs jouent un rôle déterminant dans la perception spatiale, en particulier dans des applications critiques telles que la chirurgie robotique (cf. section 1.2.4.3) ou la téléopération [Azarby 2022]. La perception de la distance y est influencée par divers facteurs, notamment l'écartement inter-caméras (ou ligne de base stéréoscopique). Une augmentation de cette distance (jusqu'à 20 millimètres selon certains auteurs) peut améliorer la restitution de la profondeur, bien qu'un écart trop important soit susceptible de générer un inconfort visuel [Avinash 2020].

De nombreuses avancées technologiques visent à optimiser la perception tridimensionnelle tout en atténuant les effets indésirables associés. L'utilisation de dispositifs autostéréoscopiques, ne nécessitant pas le port de lunettes, constitue par exemple une piste prometteuse pour réduire la charge cognitive et visuelle des utilisateurs en contexte chirurgical [Zhang 2022b].

L'intégration de technologies de vision tridimensionnelle continue de repousser les limites de la perception spatiale. Néanmoins, le contrôle précis de la profondeur perçue demeure un défi majeur. Des méthodes d'étalonnage sont proposées afin d'ajuster finement cette perception dans les images stéréoscopiques [Segura 2018], tandis que d'autres recherches s'attachent à détecter et corriger les artefacts susceptibles de la perturber [Knorr 2017, Silva 2020].

Ces travaux mettent en lumière la complexité que revêt la perception tridimensionnelle dans des environnements artificiels, en contraste avec son caractère intuitif dans le monde réel. Les altérations perceptives qui en résultent ne relèvent pas uniquement du confort subjectif, mais peuvent impacter de manière tangible la performance opérationnelle et la sécurité, notamment dans le domaine chirurgical.

3.4.4 Détection automatique des taches spéculaires

La détection des taches spéculaires, ou hautes lumières, représente une problématique clé en vision par ordinateur, tant dans les applications industrielles que médicales. Si ces artefacts lumineux ont longtemps été considérés comme des nuisances compromettant l'analyse de texture ou la reconnaissance d'objets, leur étude permet aujourd'hui de caractériser la nature des matériaux, leur géométrie et l'environnement lumineux de la scène [Fu 2020, Li 2024a, Makki 2023, Artusi 2011, Jiao 2021].

Les premières approches, développées dans les années 1990 jusqu'au début des années 2000, s'appuient principalement sur des modèles physiques et géométriques. Elles tiennent compte des interactions entre la lumière et les surfaces (cf. section 3.3), en exploitant des caractéristiques telles que les normales locales et la direction des sources lumineuses [Amanatides 1992]. Certaines méthodes utilisent des images prises selon des points de vue multiples afin de désambiguïser les zones de forte intensité lumineuse en s'appuyant sur les propriétés directionnelles des réflexions spéculaires [Ching 1993], bien que limitées par la présence d'occlusions.

La transformation d'espace colorimétrique, notamment dans l'espace Teinte-Saturation-Valeur (TSV) ou Hue-Saturation-Value (HSV) en Anglais, a permis d'améliorer la séparation des composantes spéculaires et diffuses. Certaines stratégies reposent sur des histogrammes standards ou bidimensionnels dans les espaces Intensité-Teinte ou Intensité-Saturation [Bajcsy 1996, Torres 2003, Ortiz 2005, Kim 2013]. D'autres méthodes adjoignent des techniques de seuillage automatique et d'ajustement du contraste pour une segmentation plus robuste des taches spéculaires [Morgand 2014, Xia 2019]. Ces approches ont été explorées pour les images chirurgicales [Saint-Pierre 2011, Yao 2010, Li 2019, Daher 2023, El Meslouhi 2011, Wang 2021b, Joseph 2023] mais souffrent de faux positifs avec les tissus clairs.

Les travaux de Feris et al. [Feris 2006] utilisent une collection de photos prises avec le flash de l'appareil suivant différents angles pour isoler les taches spéculaires via l'utilisation d'équations de Poisson. Tandis que Del Pozo et al. ont proposé une généralisation de cette dernière applicable à une image unique à travers l'analyse statique du flux spéculaire ou Static Specular Flow (SSF) en Anglais [DelPozo 2007]. Des méthodes alternatives ont été développées, s'appuyant sur des modèles statistiques [Zimmerman-Moreno 2006], morphologiques [Ortiz 2006] ou encore sur des mécanismes de clustering dans des espaces colorimétriques complexes [Xu 2005, Gao 2017].

L'émergence des réseaux de neurones convolutifs a transformé le champ de la détection et de la suppression des taches spéculaires. Des modèles comme MirrorNet [Yang 2019] ou Progressive Mirror Detection (PMD) en Anglais [Lin 2020] s'appuient sur des extracteurs contextuels bio-inspirés comme Contextual Contrast Feature Extraction (CCFE) en Anglais et des représentations structurelles pour segmenter avec précision les surfaces réfléchissantes. D'autres architectures telles que Mirror-YOLO [Li 2022a], dérivée de YOLOv4 [Bochkovskiy 2020], intègrent des modules d'attention comme Convolutional Block Attention Module (CBAM) en Anglais [Woo 2018] pour renforcer la connectivité de l'information dans l'image, bien que leur coût computationnel reste élevé.

La segmentation des réflexions spéculaires dans des zones à faible contraste a motivé le développement de modèles innovants comme VCNet, basé sur la chiralité visuelle [Tan 2022]. Plus récemment, des approches auto-supervisées ont été proposées pour pallier l'absence de données labellisées, en apprenant directement les caractéristiques intrinsèques des zones réfléchissantes [Lin 2023a].

L'intégration d'architectures hybrides combinant encodeur-décodeur UNet [Ronneberger 2015] et transformeur Swin [Liu 2021b] a permis de concevoir des systèmes dotés de doubles branches, capables à la fois de détecter et de supprimer les réflexions spéculaires [Wu 2023b, Rodríguez-Sánchez 2017]. De leur côté, les modèles récents comme CSFwinformer [Xie 2024b] exploitent simultanément les caractéristiques fréquentielles et spatiales à travers divers modules d'attention comme Dilated Window Attention (DWA) en Anglais et Spatial Attention Module (SAMo) en Anglais pour une localisation fine des zones de réflexion.

L'apport de la profondeur, dans le contexte des images Red Green Blue Depth (RGB-D) en Anglais, est notable. Des modèles comme PDNet [Mei 2021] ou Mirror3DNet [Tan 2021] combinent vision géométrique et reconnaissance d'objets à travers des architectures hybrides incluant Mask-RCNN [He 2017] et Plane-RCNN [Liu 2019a]. La notion d'incertitude estimée dynamiquement peut également accroître la qualité de la segmentation [Zhou 2023]. L'intégration d'approches de distillation des connaissances [Zhang 2024a, Zhou 2024] permet de maintenir des performances élevées avec des modèles légers même dans des contextes ambigus.

Dans le domaine vidéo, la dimension temporelle est exploitée pour améliorer la cohérence et la robustesse de la détection [Jiao 2021]. Des modèles tels que VMD-Net [Lin 2023b], ZOOM [Xu 2024a] ou MG-VMD [Warren 2024] utilisent la continuité temporelle des séquences pour affiner la segmentation des régions spéculaires.

Dans le cadre de nos travaux sur la synchronisation de flux vidéo chirurgicaux, l'absence de capteur RGB-D impose une contrainte forte. Par ailleurs, l'interdépendance entre estimation de profondeur et alignement temporel rend l'exploitation de la stéréoscopie inappropriée dans un contexte d'acquisition asynchrone.

3.5 Mécanisme général de synchronisation

Sur le plan technique, la synchronisation stéréoscopique repose sur un gestionnaire de synchronisation de haut niveau, chargé d'aligner temporellement deux flux d'images initialement désynchronisés. Les canaux gauche et droit peuvent présenter un désalignement temporel fixe ou variable noté δ (cf. section 3.2), qui perturbe la correspondance image à image nécessaire à une perception stéréoscopique cohérente (cf. section 3.4.3). Le gestionnaire compare les images gauches et droites issues des deux flux afin d'identifier les paires les mieux appariées via une analyse de leur contenu visuel. En sortie, deux nouveaux canaux synchrones sont générés, en d'autres termes la distance temporelle entre les images appariées est nulle ($\delta = 0$).

Ce processus s'appuie sur une recherche asymétrique de type 1 parmi N : pour chaque image d'un canal, le système identifie, au sein d'un intervalle temporel glissant, l'image correspondante optimale dans l'autre canal. Ce problème peut être formulé comme une tâche d'optimisation, visant à minimiser la distance temporelle globale tout en garantissant la cohérence séquentielle du flux, c'est-à-dire le respect de l'ordre temporel des images. Les images qui ne trouvent pas de correspondance adéquate dans le canal opposé sont explicitement étiquetées. Le cœur de cette synchronisation repose sur la capacité à quantifier, de manière fiable, la distance temporelle δ entre deux images. Cette mesure est assurée par un SCGD, basé sur le contenu visuel des images. Or, dans le contexte du traitement de vidéos chirurgicales, les SCGD conventionnels montrent leurs limites. Ils sont sensibles aux propriétés intrinsèques des images chirurgicales telles que les zones surexposées ou la forte distribution de taches spéculaires, engendrant un taux de faux positifs significatif.

Le gestionnaire de synchronisation que nous utilisons est illustré en figure 3.6. Il implémente le mécanisme de synchronisation qui, à partir d'une entrée de paires d'images désynchronisées $[F_{G,t}, F_{D,t}]$, où t est l'indice temporel, produit des paires d'images synchronisées $[\hat{F}_{G,t}, \hat{F}_{D,t}]$. La fréquence de sortie f_s est équivalente à la fréquence d'entrée f_e, tout en maintenant une latence inférieure à la période $T = 1/f_e$ entre deux images consécutives.

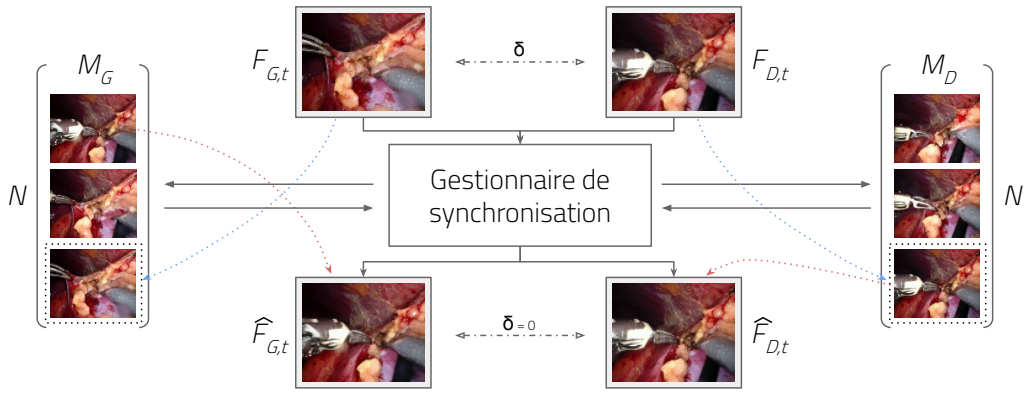


FIGURE 3.6 – Aperçu du gestionnaire de synchronisation, connectant une entrée stéréo désynchronisée $[F_{G,t}, F_{D,t}]$ à une sortie stéréo synchronisée $[\hat{F}_{G,t}, \hat{F}_{D,t}]$ en utilisant les mémoires tampons M_G et M_D de taille N et le Score de Comparaison Gauche-Droite (SCGD) qui évalue la distance temporelle δ. Source : auteur.

Le gestionnaire de synchronisation repose sur deux mémoires tampons circulaires M_G et M_D de taille N, qui conservent respectivement les N dernières images gauches et droites analysées. La taille de ces mémoires est un paramètre déterminant qui conditionne l'équilibre entre l'efficacité de la recherche de correspondances et la latence du système. En pratique, elle est généralement choisie dans un intervalle allant de $f_e/6$ à $f_e * 2$, en fonction des caractéristiques du flux vidéo traité et des ressources matérielles disponibles. Ce qui correspond en pratique à une fenêtre temporelle comprise entre 167 millisecondes et 2 secondes.

Le gestionnaire de synchronisation évalue la distance temporelle δ entre les paires d'images d'entrée et cherche à la minimiser. Pour y parvenir, trois étapes successives sont nécessaires : 1) détecter et décrire les indices visuels des images gauche et droite, 2) effectuer des tests de correspondance des indices visuels en évaluant le SCGD et 3) choisir la paire d'images qui minimise la distance temporelle δ et respecte la chronologie. À chaque itération à la fréquence f_e, les images d'entrée $F_{G,t}$ et $F_{D,t}$ sont intégrées respectivement dans les mémoires tampons M_G et M_D. Chaque nouvelle image est ensuite comparée à l'ensemble des images présentes dans la mémoire opposée. En d'autres termes, l'image $F_{G,t}$ est comparée à toutes les images dans M_D et $F_{D,t}$ à toutes les images dans M_G. Ce processus est parallélisé sur plusieurs fils d'exécution (threads en Anglais) pour améliorer l'efficacité computationnelle. Les comparaisons s'appuient sur un calcul de SCGD, dont les résultats sont stockés

dans deux vecteurs C_G et C_D. La paire synchronisée finale $[\hat{F}_{G,t}, \hat{F}_{D,t}]$ est déterminée comme celle qui minimise simultanément les deux vecteurs de score, sous contrainte de cohérence temporelle. Pour garantir le non-retour dans le temps, les images antérieures des appariements validés sont éliminées des mémoires tampons à chaque nouvelle paire d'images $[\hat{F}_{G,t}, \hat{F}_{D,t}]$ générée.

Afin de stabiliser les appariements dans le temps via une correction locale des potentielles erreurs de δ, une moyenne mobile exponentielle pondérée ou Exponentially Weighted Moving Average (EWMA) en Anglais est appliquée. Cette étape agit comme un filtre temporel, atténuant les fluctuations abruptes improbables tout en préservant la dynamique globale du flux. Le pseudo-code du gestionnaire de synchronisation est donné dans l'algorithme 1 pour un SCGD qui exploite les composantes diffuse et spéculaire des images chirurgicales.

Algorithme 1 Gestionnaire de synchronisation

Précondition : $F_{G,t}, F_{D,t}$

%*Boucle parallèle sur les couples de frames*
pour tout chaque $\langle F_{G,t}, F_{D,t} \rangle$ **faire**
  $H_G, H_D \leftarrow$ detecter_taches_speculaires$(F_{G,t}, F_{D,t})$
  $d0_G, d0_D \leftarrow$ calculer_descripteurs_speculaires(H_G, H_D)
  $d1_G, d1_D \leftarrow$ calculer_descripteurs_daisy$(H_G, H_D, F_{G,t}, F_{D,t})$
  $d2_G, d2_D \leftarrow$ calculer_sift$(F_{G,t}, F_{D,t})$
  $D_G \leftarrow$ concatener$(d0_G, d1_G, d2_G)$
  $D_D \leftarrow$ concatener$(d0_D, d1_D, d2_D)$
  si $|M_G| > 0$ **alors**
    $C_G \leftarrow$ calculer_SGGD(D_G, M_G)
    $C_D \leftarrow$ calculer_SGGD(D_D, M_D)
    $I, S \leftarrow$ selection$(\max C_G,\ \max C_D)$
    $I \leftarrow \mathrm{MMEXP}(I,\ \mathrm{prec}I)$
    Ajouter $F_{G,t}$ à M_G
    Ajouter $F_{D,t}$ à M_D
    Ajouter I à prec$I$
    si $S =$ gauche **alors**
      $\hat{F}_{G,t} \leftarrow M_G(I)$
    sinon
      $\hat{F}_{G,t} \leftarrow F_{G,t}$
    si $S =$ droite **alors**
      $\hat{F}_{D,t} \leftarrow M_D(I)$
    sinon
      $\hat{F}_{D,t} \leftarrow F_{D,t}$
retour $\langle \hat{F}_{G,t}, \hat{F}_{D,t} \rangle$

3.6 Scores de comparaison Gauche-Droit

Le SCGD constitue l'élément central du processus de synchronisation. Sa robustesse et sa pertinence conditionnent directement la qualité de l'appariement stéréoscopique. Il doit donc être spécifiquement conçu pour s'adapter aux caractéristiques des données traitées.

Dans cette optique, nous introduisons deux nouveaux SCGD conçus pour exploiter de manière ciblée les propriétés optiques des scènes chirurgicales. Le premier repose exclusivement sur la composante de réflexion spéculaire, tandis que le second combine les composantes diffuse et spéculaire. Cette hybridation permet de tirer parti de la stabilité des textures diffuses tout en conservant l'information dynamique portée par les réflexions spéculaires.

3.6.1 Exploitation de la composante spéculaire

Dans le cadre de l'analyse des réflexions purement spéculaires, nous proposons une méthode de détection et de description individualisée des taches spéculaires, visant à permettre un appariement robuste par comparaison de leurs descripteurs respectifs. À cet effet, les deux mémoires tampons M_G et M_D servent à stocker les images gauche et droite, chacune est associée à une liste structurée de taches spéculaires et à leurs descripteurs correspondants.

Sur le plan technique, le repérage spatial automatique des taches spéculaires repose sur une segmentation sémantique appliquée indépendamment à chaque image gauche et droite à l'aide d'un réseau de neurones entièrement convolutionnel (Fully Convolutional Network (FCN) en Anglais), illustré en figure 3.7. Ce modèle génère, pour toute image d'entrée F, un masque binaire M qui identifie les pixels correspondant à des réflexions spéculaires. L'architecture du réseau est constituée de dix-huit couches de convolution, intégrant chacune 64 filtres de taille $(3, 3)$ et un pas $(1, 1)$, avec une initialisation orthogonale des noyaux. L'activation Rectified Linear Unit (ReLU) en Anglais est utilisée après chaque couche, accompagnée d'une normalisation par lot (batch normalisation en Anglais) avec un momentum de 0,1 et une valeur epsilon de 0,0001. La sortie finale est produite par une fonction d'activation sigmoïde. Le réseau compte un total de 594 240 paramètres entraînables et 2 048 paramètres non entraînables. En pratique, deux instances du modèle de segmentation fonctionnent en parallèle pour les canaux gauche et droit respectivement, suivi d'un post-traitement.

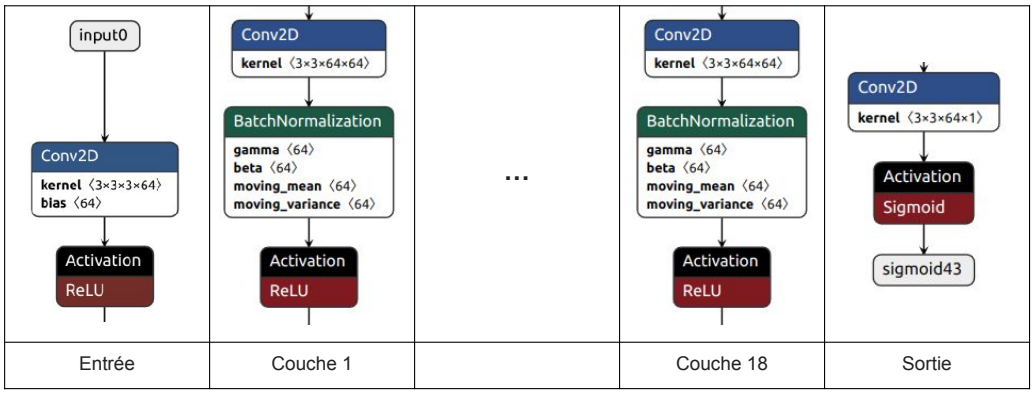


FIGURE 3.7 – Architecture du réseau de neurones convolutionnel dédié à la segmentation des taches spéculaires en endoscopie. Source : auteur.

L'entraînement du modèle a été réalisé à partir d'un jeu de données constitué de 1 000 images extraites du canal gauche de dix NPRA, collectées dans le cadre des essais de faisabilité mentionnés en section 2.2.3. L'ensemble des images a été labellisé manuellement à l'aide d'un seuillage adaptatif utilisé comme outil de prélabellisation. Quelques exemples sont présentés en figure 3.8. Le jeu de données a été divisé en 800 images pour l'entraînement et 200 pour la validation. Pour améliorer la robustesse du modèle face aux variations visuelles, des techniques classiques d'augmentation de données ont été appliquées, à savoir des retournements horizontaux et verticaux ainsi que des ajustements réalistes de luminosité dans une plage de $[-0,3, 0,3]$. L'optimiseur Adam a été utilisé pour l'apprentissage, avec un taux d'apprentissage (learning rate en Anglais) initial de 0,001, une taille de lot (batch size en Anglais) de 32 et un nombre total de 20 itérations (epochs en Anglais). La fonction de perte adoptée est l'entropie croisée binaire :

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right] \tag{3.8}$$

où N est le nombre total d'observations, y_i le label manuel aussi appelé vérité terrain pour l'observation i et $\hat{y}_i$ le label prédit par le modèle pour cette même observation i.

Une fois les taches spéculaires automatiquement segmentées, une étape d'étiquetage est effectuée en identifiant chaque composante connexe dans le masque M, selon la méthode décrite par Bolelli et al. [Bolelli 2019]. Dans une boucle itérative parallèle, chaque tache est ensuite caractérisée à travers cinq composantes distinctes : trois attributs géométriques (position du centroïde, surface de la tache en pixels et premier moment) et deux attributs d'intensité calculés sur une région carrée de dix pixels centrée sur le centroïde, à savoir un histogramme de couleur en dix intervalles ainsi que le descripteur DAISY pur [Tola 2009, Melloni 2015]. L'appariement des taches est ensuite effectué par un algorithme de type brute-force avec

validation croisée, tel qu'implémenté dans les bibliothèques de vision par ordinateur comme OpenCV (BFMatcher). Le SCGD est défini comme le nombre de correspondances validées par le Test de Ratio de Lowe (TRL) [Lowe 2004], avec un seuil fixé à $d = 0,8$. Le pseudo-code du traitement pour une image donnée est présenté dans l'algorithme 2.

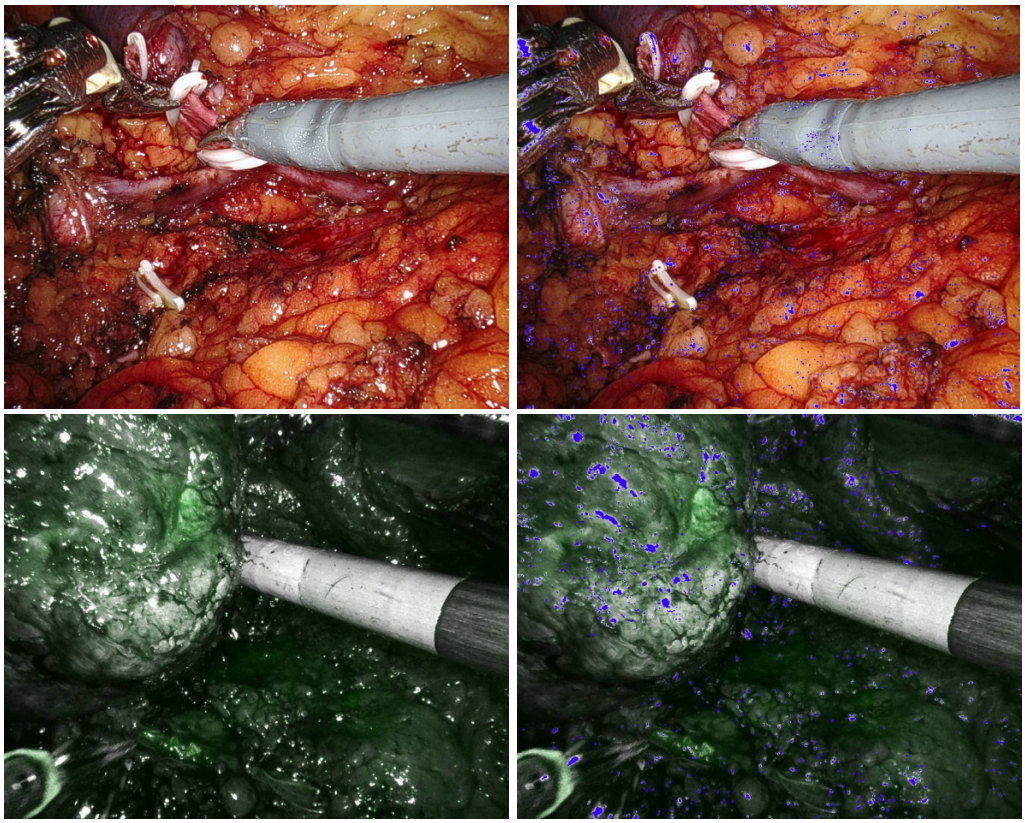


FIGURE 3.8 – Exemples d'images de Néphrectomie Partielle Robot-Assistée (NPRA) labellisées manuellement avec la classe "tache spéculaire" visible en bleu. (en haut) Image en couleur RVB (Rouge-Vert-Bleu) totalisant 1 719 taches spéculaires isolées, (en bas) image de fluorescence avec l'indocyanine verte (ICG) totalisant 969 taches spéculaires isolées. Source : auteur.

Algorithme 2 Traitement de la composante spéculaire d'une image

Précondition : F, modèle FCN

$M \leftarrow \text{FCN}(F)$ %*Segmentation du masque spéculaire*
$T \leftarrow$ Extraction des composantes connexes dans M
$D \leftarrow \emptyset$ %*Initialisation des descripteurs*
pour tout $t \in T$ **en parallèle faire**
    $c \leftarrow \text{Centroïde}(t)$
    $s \leftarrow \text{Surface}(t)$
    $m_1 \leftarrow \text{Premier moment}(t)$
    $R \leftarrow$ Patch centré sur c de taille 10×10
    $h \leftarrow \text{Histogramme couleur}(R, 10 \text{ bins})$
    $d \leftarrow \text{Descripteur DAISY}(R)$
    $D \leftarrow D \cup \{(c, s, m_1, h, d)\}$
retour D

3.6.2 Exploitation des composantes diffuse et spéculaire

Dans le but d'exploiter conjointement les composantes de réflexion spéculaire et diffuse afin d'accroître la stabilité et la précision de la synchronisation passive, nous introduisons une mesure de synchronisation combinée, dite SCGD diffus-spéculaire. Cette mesure fusionne, pour chaque paire d'images stéréoscopiques, la contribution du SCGD spéculaire proposé (cf. section 3.6.1) avec celle d'un SCGD fondé sur les informations diffuses. Pour ce dernier, nous retenons le nombre de correspondances obtenues à l'aide du détecteur-descripteur SIFT [Lowe 2004], en raison de son bon compromis entre invariance aux transformations locales et efficacité computationnelle.

3.7 Génération de données d'évaluation semi-synthétiques

Nous avons conçu un simulateur de désynchronisation capable de transformer une séquence stéréoscopique initialement synchrone en une séquence semi-synthétique asynchrone, tout en offrant un contrôle fin sur les paramètres de désynchronisation. Le simulateur permet notamment de moduler la fréquence d'images, qu'elle soit constante (CFR) ou variable (VFR), de simuler la perte ou la duplication stochastique d'images ainsi que d'introduire des décalages temporels initiaux indépendants pour chaque canal. Ce cadre de simulation constitue un environnement expérimental rigoureusement contrôlé, particulièrement adapté à l'évaluation des algorithmes de synchronisation.

Sur le plan opérationnel, le simulateur exploite des indices d'images extraits au préalable du flux stéréoscopique d'origine, représentés sous forme de fichiers image individuels. Chaque canal est modélisé comme une séquence ordonnée de fichiers indexés de 0 à $n-1$, où n correspond au nombre total d'images disponibles. Soit $\mathcal{L}_G$ et

$\mathcal{L}_D$ les séquences ordonnées d'indices pour les canaux gauche et droit respectivement. Trois paramètres définis par l'utilisateur régissent la simulation : le décalage initial appliqué respectivement aux canaux gauche (δ_G) et droit (δ_D), ainsi que l'amplitude de variation de la fréquence d'échantillonnage (VFR) vfr_range. Il est à noter que les fichiers sources ne sont ni altérés ni directement chargés, ce qui garantit l'intégrité des données et une exécution efficiente.

Pour simuler des conditions réalistes, comparables à celles des systèmes d'imagerie stéréoscopique fonctionnant à 60 Hz, les décalages initiaux typiques varient de 0 à 120 images, tandis que la plage de VFR est généralement comprise entre $[-1, 1]$ image.

Lors de la simulation, les décalages initiaux δ_G et δ_D sont appliqués en éliminant les indices correspondants dans chaque séquence $\mathcal{L}_G$ et $\mathcal{L}_D$. Si la modulation de fréquence (VFR) est activée, les séquences d'indices précédemment obtenues font l'objet d'un sous-échantillonnage pour permettre de modéliser la suppression stochastique d'images. Le facteur de sous-échantillonnage ν est défini par $\nu = 2\max(|\texttt{vfr_range}[0]|, |\texttt{vfr_range}[1]|) + 1$. Il représente l'intervalle minimal entre deux images consécutives afin de prévenir toute redondance ou chevauchement non maîtrisé.

Les indices désynchronisés sont alors générés en introduisant un bruit stochastique, contraint par la plage de VFR, appliqué aux indices des séquences initiales $\mathcal{L}_G$ et $\mathcal{L}_D$ sous-échantillonnées. À l'issue de ce processus, les séquences d'indices gauche $\mathcal{L}_G$ et droite $\mathcal{L}_D$ sont tronquées de manière à conserver une longueur identique. Les fichiers image sont ensuite dupliqués selon ces séquences d'indices modifiées $\mathcal{L}_G$ et $\mathcal{L}_D$, en préservant la convention de nommage, laquelle encode explicitement la position temporelle dans le flux synchrone initial, fournissant ainsi un repère de vérité terrain pour l'évaluation.

Les données générées sont dès lors prêtes à être intégrées dans des pipelines de traitement automatisé. Le pseudo-code détaillé du simulateur de désynchronisation est présenté dans l'algorithme 3.

Algorithme 3 Simulateur de désynchronisation

Précondition : $\mathcal{S}_G, \mathcal{S}_D, \delta_G, \delta_D, \text{VFR}, [r_{min}, r_{max}]$

$\mathcal{L}_G \leftarrow$ Liste triée des images de $\mathcal{S}_G$
$\mathcal{L}_D \leftarrow$ Liste triée des images de $\mathcal{S}_D$
assert $\mathcal{L}_G = \mathcal{L}_D$ %*Vérification de synchronisation initiale*
%*1 - Application des décalages initiaux*
$\mathcal{L}_G \leftarrow \mathcal{L}_G[\delta_G :]$
$\mathcal{L}_D \leftarrow \mathcal{L}_D[\delta_D :]$
%*2 - Simulation du VFR (Variable Frame Rate)*
si VFR activé **alors**
    $\nu \leftarrow 2 \cdot \max(|r_{min}|, |r_{max}|) + 1$ %*Facteur de sous-échantillonnage*
    $\mathcal{I}_G \leftarrow$ Indices de $\mathcal{L}_G$ sous-échantillonnés avec pas ν
    $\mathcal{I}_D \leftarrow$ Indices de $\mathcal{L}_D$ sous-échantillonnés avec pas ν
    $\varepsilon_G \sim \mathcal{U}[r_{min}, r_{max}]$ pour chaque indice dans $\mathcal{I}_G$
    $\varepsilon_D \sim \mathcal{U}[r_{min}, r_{max}]$ pour chaque indice dans $\mathcal{I}_D$
    $\mathcal{I}_G \leftarrow \mathcal{I}_G + \varepsilon_G$
    $\mathcal{I}_D \leftarrow \mathcal{I}_D + \varepsilon_D$
    $\mathcal{L}_G \leftarrow [\mathcal{L}_G[i]]$ pour chaque $i \in \mathcal{I}_G$
    $\mathcal{L}_D \leftarrow [\mathcal{L}_D[i]]$ pour chaque $i \in \mathcal{I}_D$
    assert Aucune duplication dans $\mathcal{L}_G$, $\mathcal{L}_D$
%*3 - Troncature pour égaliser les longueurs*
$n \leftarrow \min(|\mathcal{L}_G|, |\mathcal{L}_D|)$
$\mathcal{L}_G \leftarrow \mathcal{L}_G[: n]$
$\mathcal{L}_D \leftarrow \mathcal{L}_D[: n]$
%*4 - Export optionnel*
si export activé **alors**
    pour $(img_G, img_D) \in (\mathcal{L}_G, \mathcal{L}_D)$ **faire**
        Copier img_G dans le dossier de destination gauche
        Copier img_D dans le dossier de destination droit
retour $\mathcal{L}_G, \mathcal{L}_D$

3.8 Expériences

Dans cette section, nous avons conduit une série d'expérimentations sur données afin d'évaluer la performance et l'impact des SCGD proposés pour la synchronisation passive de flux vidéo stéréoscopique chirurgicaux. L'objectif est de valider leur pertinence dans le contexte chirurgical, en particulier dans le cadre exigeant de la CGI.

3.8.1 Jeu de données pour l'évaluation

Le jeu de données d'évaluation se compose de 44 séquences vidéo stéréoscopiques acquises en environnement chirurgical réel. Il inclut 38 séquences de NPRA collectées au cours des sessions d'essais de faisabilité décrites en section 2.2.3, ainsi que 6 séquences de RHRA réalisées à l'aide du télémanipulateur chirurgical Da Vinci Xi (Intuitive Surgical). L'ensemble est structuré en 26 séquences courtes de 21 images dont deux exemples sont illustrés en figure 3.9 et 18 séquences longues de 200 images.

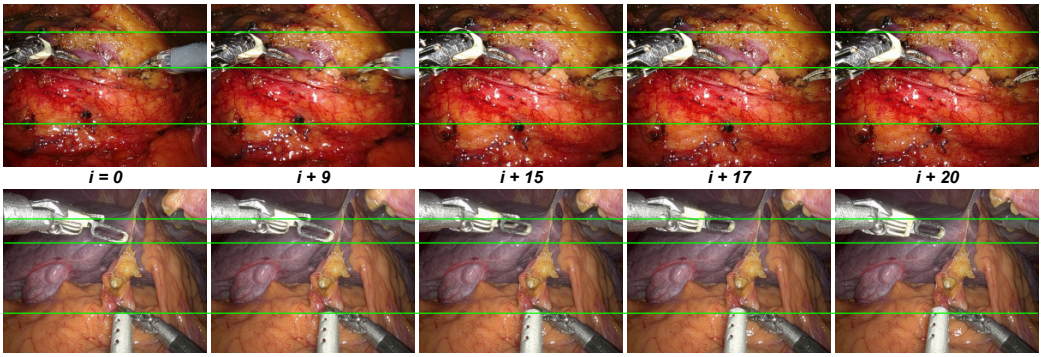


FIGURE 3.9 – Exemples de deux séquences courtes enregistrées à 60 Hz : (en haut) Néphrectomie Partielle Robot-Assistée (NPRA) et (en bas) Résection Hépatique Robot-Assistée (RHRA). Source : auteur.

Toutes les séquences ont été capturées en imagerie endoscopique couleur RVB, à l'exception de quatre séquences courtes de NPRA, enregistrées en mode fluorescence avec l'indocyanine verte (ICG) (cf. section 1.3.2.2). Les données reflètent la variabilité des contextes opératoires, avec des dynamiques allant de scènes quasi-statiques à des mouvements rapides de la caméra endoscopique, intégrant également les interactions instrumentales sur les tissus. Ce jeu de données constitue ainsi une base d'évaluation représentative et robuste pour l'analyse de scénarios chirurgicaux à cinématique variable.

La synchronisation des flux repose sur un système d'acquisition stéréoscopique dédié, doté d'une fréquence d'échantillonnage constante (CFR) et d'un mécanisme de déclenchement simultané entre les canaux gauche et droit, assurant une capture quasi-synchrone. Un alignement temporel fin est ensuite obtenu par une inspection

image-par-image. Des événements visuellement saillants, tels que les arcs d'électrochirurgie, sont exploités en tant que marqueurs temporels de référence. Cette procédure rigoureuse garantit une synchronisation précise, indispensable au développement et à l'évaluation fiable des algorithmes de synchronisation.

3.8.2 Mesure de la distance temporelle

La mesure précise de la distance temporelle entre les flux stéréoscopiques gauche et droit peut être effectuée selon diverses approches. Toutes reposent sur l'hypothèse que la caméra endoscopique filme une scène unique avec un certain recouvrement entre les vues gauche et droite. En endoscopie, des événements transitoires et imprévisibles comme les arcs électriques générés lors de l'utilisation d'instruments d'électrochirurgie, peuvent être utilisés comme repères temporels ponctuels. Toutefois, en l'absence de référence externe, il devient quasiment impossible d'assurer une évaluation temporelle continue et homogène pour chaque image du flux vidéo. Cette absence de repère constant compromet la quantification fiable des phénomènes de désynchronisation aléatoires ou variables au cours du temps.

Une approche robuste pour obtenir une mesure instantanée et répétable consiste à intégrer une référence visuelle détectable par des techniques de vision par ordinateur. Lorsque l'environnement le permet, l'exploitation de la reconnaissance optique de caractères ou Optical Character Recognition (OCR) en Anglais peut être envisagée, notamment si un chronomètre est visible au sein du champ filmé. Une alternative plus contrôlable et performante repose sur l'utilisation de marqueurs synthétiques, tels que les marqueurs AruCo. Ces marqueurs bidimensionnels, caractérisés par une bordure noire épaisse entourant une matrice binaire unique, fournissent une identification explicite (index, noté idx) à chaque instance. La complexité du code interne dépend directement de la dimension du marqueur, par exemple une structure 6×6 encapsule 36 bits d'information.

L'implémentation de cette méthode nécessite la génération préalable des marqueurs AruCo selon un dictionnaire prédéfini. Ces marqueurs sont destinés à être affichés puis capturés par le système optique, tandis qu'un algorithme de détection dédié assure l'extraction de l'identifiant unique (idx) de chaque marqueur à partir des images acquises, en se référant au même dictionnaire.

Afin de garantir que chaque image acquise contienne un marqueur distinct, il est impératif que la fréquence de mise à jour des marqueurs, notée f_t, soit suffisante par rapport à la fréquence d'acquisition du flux vidéo. Contrairement au principe classique du théorème d'échantillonnage de Nyquist-Shannon, où l'objectif est d'éviter le repliement spectral en échantillonnant à une fréquence au moins deux fois supérieure à celle du signal, la présente approche vise à s'assurer que le contenu affiché (le marqueur) soit renouvelé à chaque échantillon, c'est-à-dire à chaque image capturée.

Ainsi, la fréquence d'actualisation des marqueurs f_t doit être au minimum égale à la fréquence maximale du flux vidéo $f_{\max}$, voire supérieure si une meilleure résolution temporelle est requise :

$$f_t \geq f_{\max} \tag{3.9}$$

La différence d'indices détectés entre les canaux gauche et droit, notée $\delta = idx_G - idx_D$, permet alors de caractériser la distance temporelle instantanée. L'analyse statistique de cette variable offre des indicateurs standards comme les distances temporelles minimale, moyenne et maximale. Une valeur $\delta > 0$ indique une avance du canal gauche sur le droit, tandis qu'une valeur $\delta < 0$ révèle l'inverse. La condition $\delta = 0$ traduit une synchronisation parfaite entre les deux flux. Il est également possible de quantifier les erreurs d'appariement, c'est-à-dire les cas où l'identification des marqueurs ne correspond pas entre les deux canaux. Un exemple d'implémentation de cette méthode est présenté en figure 3.10.

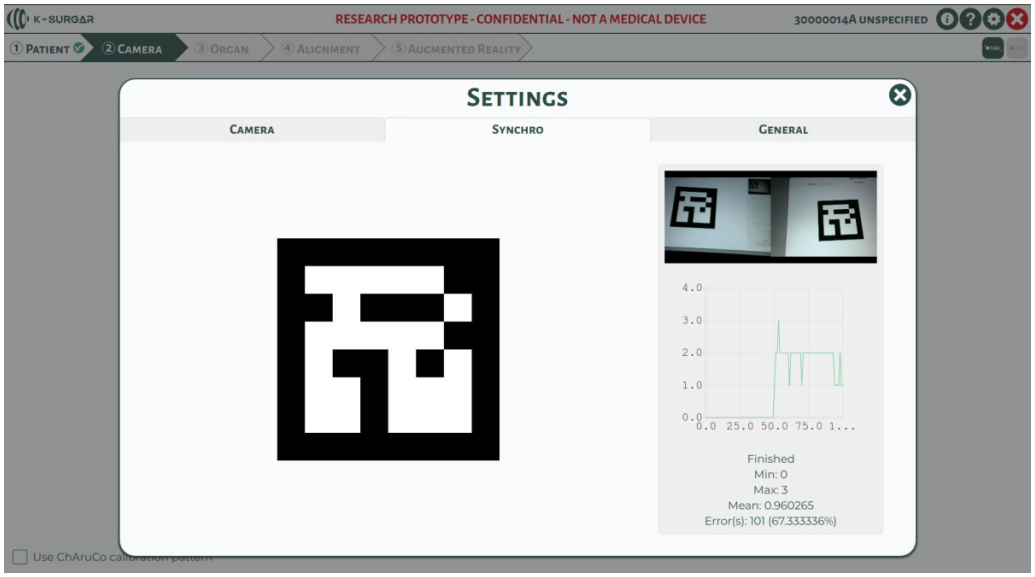


FIGURE 3.10 – Exemple d'implémentation d'une méthode de mesure de la distance temporelle entre deux canaux d'un système d'acquisition stéréoscopique à l'aide d'une référence externe comme un marqueur AruCo. Source : auteur.

Sur un ordinateur équipé d'Ubuntu 24.04 et d'un processeur Intel Core™ i7-13620H, le temps de traitement nécessaire à la détection d'un marqueur AruCo de dimension 6×6, intégré dans une image de résolution 512×512 pixels, s'élève à deux millisecondes. Cette performance est largement compatible avec une fréquence d'échantillonnage de 60 Hz, correspondant à une période de 16,67 millisecondes.

3.8.3 Évaluation de la performance des scores proposés

Une série d'expériences est réalisée afin d'évaluer la performance brute des deux SCGD proposés, à savoir celui exploitant uniquement la composante spéculaire nommé SPEC (cf. section 3.6.1) et celui utilisant les composantes diffuse et spéculaire conjointement nommé DIFF_SPEC (cf. section 3.6.2) au regard de 216 SCGD de la littérature. Pour cela, une analyse quantitative et qualitative des courbes de réponse pour une tâche d'appariement est réalisée, afin de caractériser la sensibilité des méthodes lors de la recherche de la meilleure paire composée par défaut d'une image gauche pré-sélectionnée et $K = 21$ images droites candidates. Tâche qui incombe au SCGD dans le contexte de la synchronisation passive de flux vidéo stéréoscopiques. Le score résultant permet in fine au gestionnaire de synchronisation de réaliser la minimisation de la distance temporelle et donc de proposer une nouvelle paire d'image synchrone (cf. section 3.5).

Les expériences ont été conduites pour les 26 séquences courtes extraites de notre jeu de données d'évaluation présenté en section 3.8.1.

Les 216 SCGD pertinents utilisables en chirurgie peuvent être regroupés en trois catégories de méthodes.

Une première catégorie concerne les méthodes de mise en correspondance de caractéristiques locales, qui s'appuient essentiellement sur la détection et la description de points de clés (keypoints en Anglais), via des méthodes issues de la littérature listées en annexe L. La qualité et la robustesse de la détection et de la description de ces dernières conditionne l'invariance de la méthode face aux changements visuels comme la colorimétrie mais également le point de vue.

Nous avons choisi 201 combinaisons de détecteurs-descripteurs parmi les plus utilisés dans le domaine. Chaque combinaison doit comprendre une méthode de détection et une méthode de description parmi 9 détecteurs locaux purs : Adaptive and Generic Accelerated Segment Test (AGAST) en Anglais [Mair 2010], Canny [Canny 2009], FAST [Rosten 2006], Good Feature To Track (GFTT) en Anglais [Shi 1994], Harris-Laplace (HL) [Mikolajczyk 2004], Maximal Self-Dissimilarity (MSD) en Anglais [Tombari 2014], Maximally Stable Extremal Regions (MSER) en Anglais [Matas 2004], Simple Learned Keypoints (SiLK) en Anglais [Gleize 2023], STAR [Agrawal 2008] ; 9 descripteurs locaux purs Boosted Efficient Binary Local Image Descriptor (BEBLID) en Anglais [Suárez 2020], BOOST DESC [Trzcinski 2013], Binary Robust Independent Elementary Features (BRIEF) en Anglais [Calonder 2010], DAISY [Tola 2009], Fast Retina Keypoint (FREAK) en Anglais [Alahi 2012], Learned Arrangements of Three Patch Codes (LATCH) en Anglais [Levi 2016], Locally Uniform Comparison Image Descriptor (LUCID) en Anglais [Ziegler 2012], TFEAT [Balntas 2016], Oxford Visual Geometry Group (VGG) [Simonyan 2014a] et 12 détecteurs-descripteurs locaux Accelerated-KAZE (AKAZE) en Anglais [Alcantarilla 2011], Accurate and LIghtweight KEypoint detection and descriptor (ALIKE) en Anglais [Zhao 2022b], A LIghter Keypoint and descriptor Extraction network via Deformable transformation (ALIKED) en Anglais [Zhao 2023b], Binary

Robust Invariant Scalable Keypoints (BRISK) en Anglais [Leutenegger 2011], D2-Net [Dusmanu 2019], Detect, Don't Describe - Describe, Don't Detect (DeDoDe) en Anglais [Edstedt 2024a], DIScrete Keypoints (DISK) en Anglais [Tyszkiewicz 2020], KAZE [Alcantarilla 2012], ORB [Rublee 2011], SIFT [Lowe 2004], SuperPoint (SP) [DeTone 2018], SURF [Bay 2006]. Le nombre 201 combinaisons de détecteurs-descripteurs est obtenu par le calcul suivant : 12 détecteurs-descripteurs + 12 détecteurs-descripteurs (descripteurs non utilisés) × 9 descripteurs + 9 détecteurs × 9 descripteurs.

Une méthode d'appariement exhaustive (Brute-Force Matcher) avec vérification croisée est utilisée pour mettre en correspondance le descripteur gauche avec les K descripteurs droits. Pour cela, un calcul de distance entre descripteurs est effectué via une norme L2 par défaut et une distance de Hamming pour les descripteurs basés sur des chaînes de caractères binaires comme ORB, BRIEF et BRISK. Le tout suivi d'un filtrage standard via le TRL [Lowe 2004].

Cette même catégorie est complétée par dix méthodes de détection, description et appariement end-to-end utilisant les modèles d'apprentissage profond suivants : ASpanFormer [Chen 2022], CO-Attention for conditioned image Matching (COAM) en Anglais [Wiles 2021], Efficient LoFTR [Wang 2024c], GlueStick [Pautrat 2023], LightGlue [Lindenberger 2023], Local Feature TRansformer (LoFTR) en Anglais [Sun 2021], Patch Area Transportation with Subdivision (PATS) en Anglais [Ni 2023], Robust dense feature Matching (RoMa) en Anglais [Edstedt 2024b], SuperGlue [Sarlin 2020] et XFeat [Potje 2024].

Pour cette première catégorie, le SCGD correspond à la valeur numérique représentant le nombre de points clés appariés, après application du TRL, comme illustré pour une paire d'images gauche-droite en figure 3.12.

Une seconde catégorie composée de trois SCGD, utilise des caractéristiques globales de l'image via la mise en œuvre d'un descripteur comme GIST [Oliva 2001], l'indice de la similarité structurelle Structural Similarity Index Measure (SSIM) en Anglais [Wang 2004] et un Descripteur Profond (DeP) à partir des cartes de caractéristiques du réseau de neurones AlbuNet initialement entraînées pour la segmentation sémantique du parenchyme rénal. En quelques mots, ce modèle utilise l'architecture UNet [Ronneberger 2015] avec un encodeur Resnet34 [He 2016] pré-entraîné sur le jeu de données ImageNet [Deng 2009]. Pour l'entraînement, nous avons créé un ensemble de données de 15 683 images extraites de 18 procédures NPRA que nous avons annotées manuellement. Nous avons entraîné le modèle avec l'optimiseur Stochastic Gradient Descent (SGD) en Anglais pendant 30 itérations avec une taille de lot de 8, un taux d'apprentissage de 0,001 et une fonction de perte par entropie croisée.

Pour cette deuxième catégorie, les SCGD composés de descripteurs globaux utilisent la distance L2 comme score résultant.

Une troisième et dernière catégorie exploite des informations temporelles collectées à partir d'une série d'images successives. Pour cela l'intégralité des 21 images gauches et des 21 images droites des séquences d'évaluation est utilisée, afin de calculer un SCGD représentatif de l'alignement temporel global. Les méthodes d'analyse

du profil de débit binaire de la vidéo (bitrate en Anglais) ou l'analyse du profil de trajectoire bidimensionnelle sont mises en œuvre. Dans ce second cas, des homographies sont estimées à partir de points SIFT détectés entre les images consécutives d'un même canal et ceux à gauche et à droite, afin d'évaluer une distance temporelle selon la méthode proposée par Caspi et al. [Caspi 2002a].

Comme mentionné précédemment, l'évaluation de la performance brute des SCGD est réalisée à partir des courbes de réponse. Le graphique de réponse idéal serait une courbe en cloche centrée sur une distance temporelle nulle avec une forte décroissance de part et d'autre de cette dernière. Par souci de clarté, seuls les meilleurs représentants de chaque catégorie sont analysés. Les temps de calcul ou d'inférence ne sont pas pris en compte pour juger de la performance du score. Les figures 3.13 et 3.11 ainsi que le tableau 3.1 montrent que dans les séquences contenant des mouvements rapides de la caméra endoscopique (figure 3.13 a, c, e, g), les extracteurs de caractéristiques locales avec support spatial diffus, les descripteurs globaux, l'indice de similarité structurelle, les cartes de caractéristiques du réseau neuronal entraîné et les méthodes proposées révèlent une sensibilité suffisante pour re-synchroniser une paire d'images gauche-droite. Cependant, les méthodes globales deviennent inutilisables dans les scénarios quasi-statiques (figure 3.13 b, d, f, h) et les méthodes utilisant des descripteurs locaux diffus présentent des taux d'erreur dépassant deux images. La méthode proposée se distingue par sa capacité à re-synchroniser à l'image près grâce à sa haute sensibilité aux micro-mouvements. Cela la différencie des descripteurs locaux diffus, qui sont intrinsèquement insensibles aux petits mouvements en raison de leur invariance.

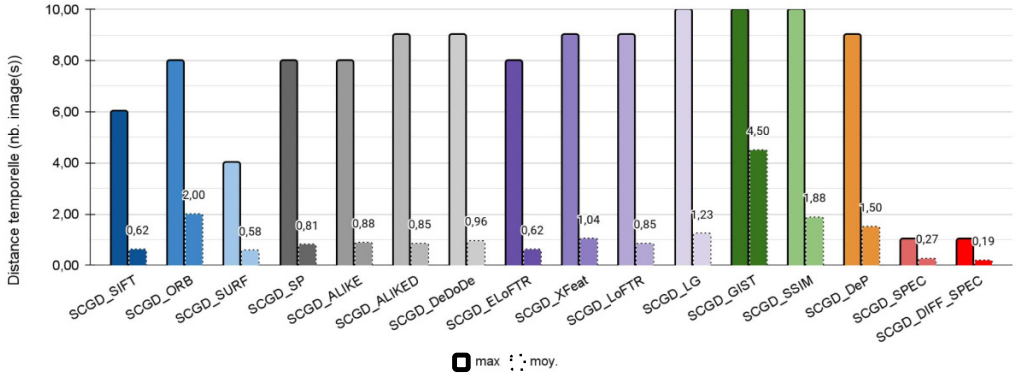


FIGURE 3.11 – Distance temporelle moyenne et maximale pour l'évaluation du décalage de synchronisation, estimées sur les 26 séquences courtes du jeu de données, pour différents Score de Comparaison Gauche-Droite (SCGD) utilisant les méthodes suivantes : SIFT, ORB, SURF, SuperPoint (SP), ALIKE, ALIKED, DeDoDe, Efficient LoFTR (ELoFTR), XFeat, LoFTR, LightGlue (LG), GIST, SSIM, les cartes de caractéristiques du réseau neuronal entraîné (DeP) et les méthodes proposées SPEC et DIFF_SPEC. Source : auteur.

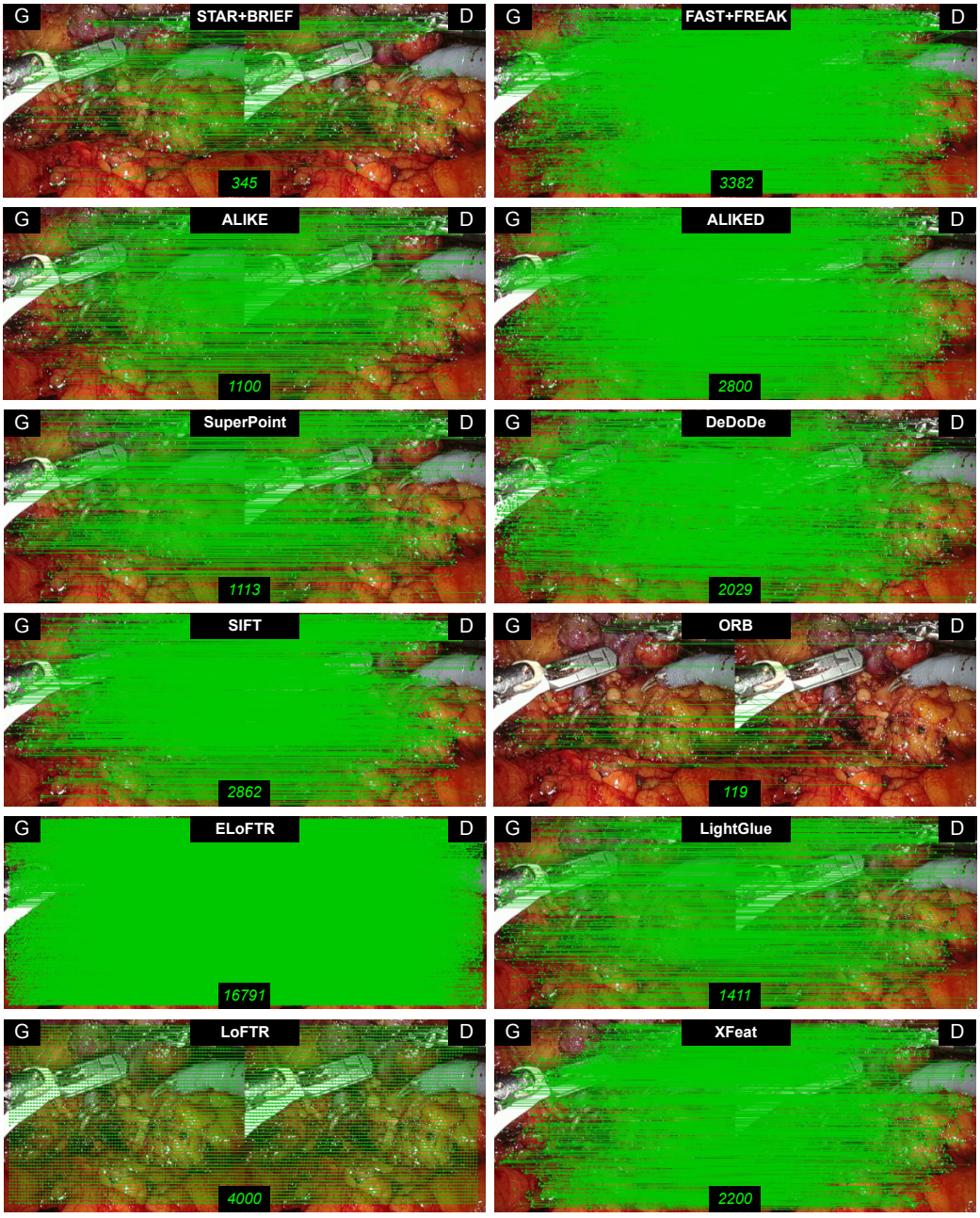


FIGURE 3.12 – Mise en correspondance de caractéristiques locales des images d'une paire gauche-droite synchrone via diverses méthodes de la littérature : STAR+BRIEF, FAST+FREAK, ALIKE, ALIKED, SuperPoint, DeDoDe, SIFT, ORB, Efficient LoFTR (ELoFTR), LightGlue, LoFTR et XFeat. Source : auteur.

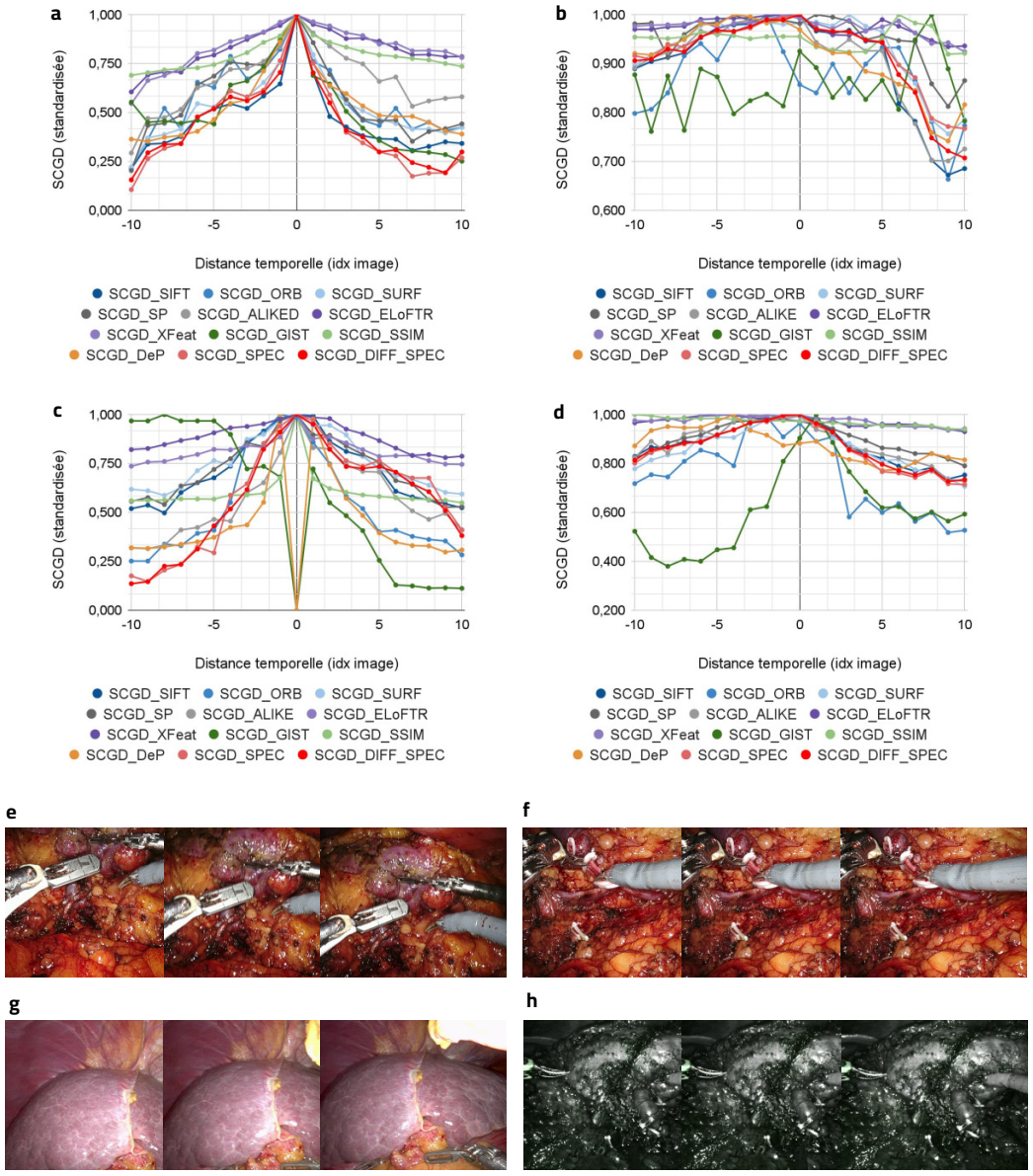


FIGURE 3.13 – (a) Courbes de réponse pour la séquence n°8 avec mouvement rapide de la caméra lors d'une Néphrectomie Partielle Robot-Assistée (NPRA). (b) Courbes de réponse pour la séquence n°12 quasi-statique lors d'une NPRA. (c) Courbes de réponse pour la séquence n°22 avec mouvement rapide de la caméra lors d'une Résection Hépatique Robot-Assistée (RHRA). (d) Courbes de réponse pour la séquence n°23 quasi-statique lors d'une NPRA, capturée en mode fluorescence. (e) Images d'exemple de la séquence n°8 avec mouvement rapide de la caméra lors d'une NPRA. (f) Images d'exemple de la séquence n°12 quasi-statique lors d'une NPRA. (g) Images d'exemple de la séquence n°22 avec mouvement rapide de la caméra lors d'une RHRA. (h) Images d'exemple de la séquence n°23 quasi-statique lors d'une NPRA, capturées en mode fluorescence. Source : auteur.

			Distance temporelle en nombre d'image(s)																		
Chir.	Image	Mvt. Caméra	SCGD SIFT	SCGD ORB	SCGD SURF	SCGD SP	SCGD ALIKE	SCGD ALIKED	SCGD DeDoDe	SCGD ELoFTR	SCGD XFeat	SCGD LoFTR	SCGD LG	SCGD GIST	SCGD SSIM	SCGD DeP	SCGD SPEC	SCGD DIFF_SPEC
NPRA	RVB	Lent	0	1	0	2	0	0	0	0	1	0	0	4	1	1	0	0
NPRA	RVB	-	0	8	0	0	0	0	0	0	0	0	1	3	1	1	0	0
NPRA	RVB	Rapide	0	0	0	1	0	0	1	0	1	0	0	5	1	1	0	0
NPRA	RVB	-	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0
NPRA	RVB	-	1	1	0	0	0	0	0	0	1	2	2	10	1	0	0	0
NPRA	RVB	-	0	0	0	1	0	0	0	0	0	0	0	3	0	1	0	0
NPRA	RVB	-	0	0	0	0	7	9	1	0	1	1	0	4	4	3	0	0
NPRA	RVB	Rapide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NPRA	RVB	-	0	2	0	0	0	0	0	0	0	0	0	10	0	0	0	0
NPRA	RVB	-	0	8	0	1	0	0	9	1	0	0	1	0	10	0	1	0
NPRA	RVB	-	0	0	0	0	0	0	0	0	2	0	0	9	0	0	0	0
NPRA	RVB	-	1	2	3	1	0	0	0	0	1	2	10	8	6	4	0	0
NPRA	RVB	Rapide	0	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0
NPRA	RVB	Rapide	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0
NPRA	RVB	-	4	4	4	3	3	4	3	0	3	0	7	10	1	2	1	1
NPRA	RVB	-	0	1	1	1	1	0	1	0	1	0	1	0	2	0	0	0
NPRA	RVB	-	0	0	0	0	0	0	0	0	0	0	0	8	10	7	0	0
NPRA	RVB	-	0	1	0	1	0	0	0	0	0	0	0	4	0	4	0	0
RHRA	RVB	-	6	2	4	8	8	8	5	8	9	9	6	1	10	9	1	1
RHRA	RVB	-	1	8	0	0	0	0	0	0	0	0	1	2	1	0	1	1
RHRA	RVB	-	1	1	1	1	1	1	1	1	2	2	2	8	0	1	1	1
RHRA	RVB	Rapide	0	1	0	0	0	0	0	0	0	0	0	8	0	1	0	0
RHRA	ICG	-	0	3	0	0	3	0	3	5	4	5	0	1	0	1	1	0
NPRA	ICG	Rapide	0	2	1	0	0	0	0	0	0	0	0	1	0	1	0	0
NPRA	ICG	-	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0
NPRA	ICG	-	2	4	1	1	0	0	1	1	1	1	1	1	0	1	1	1
		Moyenne	0,62	2,00	0,58	0,81	0,88	0,85	0,96	0,62	1,04	0,85	1,23	4,50	1,88	1,50	0,27	0,19
		Maximum	6,00	8,00	4,00	8,00	8,00	9,00	9,00	8,00	9,00	9,00	10,00	10,00	10,00	9,00	1,00	1,00
		Ecart-type	1,42	2,48	1,21	1,65	2,12	2,39	2,05	1,81	1,93	2,01	2,52	3,56	3,29	2,25	0,45	0,40

TABLE 3.1 – Distances temporelles pour l'évaluation du décalage de synchronisation, estimées sur les 26 séquences courtes du jeu de données, pour différents Score de Comparaison Gauche-Droite (SCGD) utilisant les méthodes suivantes : SIFT, ORB, SURF, SuperPoint (SP), ALIKE, ALIKED, DeDoDe, Efficient LoFTR (ELoFTR), XFeat, LoFTR, LightGlue (LG), GIST, SSIM, les cartes de caractéristiques du réseau neuronal entraîné (DeP) et les méthodes proposées SPEC et DIFF_SPEC. Source : auteur.

3.8.4 Évaluation de l'impact de la segmentation des taches spéculaires sur la synchronisation

L'évaluation du module de segmentation automatique des taches spéculaires utilisé par les deux SCGD proposés est cruciale afin de quantifier son impact sur la méthode de synchronisation. Pour ce faire, un jeu de données soigneusement sélectionné de trente images a été créé, avec une distribution uniforme entre les jeux de données de NPRA, de RHRA et de coloscopie disponibles. Ces données constituent un point de référence essentiel pour évaluer quantitativement la précision de la segmentation des taches spéculaires et son impact sur la formulation des descripteurs.

Pour cette évaluation analytique, nous avons systématiquement comparé les composants des descripteurs extraits des images de référence (Ground Truth (GT) en Anglais) dans ce jeu de données avec ceux dérivés des masques de segmentation prédits par le modèle de segmentation (cf. section 3.4.4). Un ensemble complet de métriques a été utilisé, incluant le taux de détection des taches spéculaires en nombre d'instances segmentées, l'indice de Dice de la mesure de similarité entre deux masques binaires de segmentation, l'erreur de localisation du centroïde dans le plan image, la déviation angulaire du premier moment des taches spéculaires, les distances de Bhattacharyya D_B et Hellinger D_H pour la congruence des histogrammes (distributions discrètes), les distances L2 des descripteurs DAISY et les distances L2 globales des descripteurs spéculaires.

$$\text{Dice} = \frac{2|P \bigcap G|}{|P + G|} \tag{3.10}$$

où P et G sont respectivement les ensembles finis relatifs à la prédiction et à la référence. Pour des données binaires, l'indice de Dice s'écrit :

$$\text{Dice} = \frac{2VP}{2VP + FP + FN} \tag{3.11}$$

où VP, FP et FN sont respectivement les éléments/pixels associés aux vrais positifs, faux positifs et faux négatifs.

$$D_B(P, Q) = -\ln\left(BC(P, Q)\right) \quad \text{avec} \quad BC(P, Q) = \sum_{x \in X} \sqrt{P(x) \cdot Q(x)} \tag{3.12}$$

où BC est le coefficient de Bhattacharyya et $P(x)$ et $Q(x)$ deux distributions de probabilités discrètes définies sur l'ensemble X.

$$D_H(P, Q) = \sqrt{1 - BC(P, Q)} \tag{3.13}$$

Les résultats de cette expérience montrent un taux de détection des taches spéculaires moyen de $79,85 \pm 23,60\,\%$, avec un indice de Dice moyen de $81,34 \pm 18,11\,\%$, une erreur de localisation du centroïde moyenne dans le plan image de $1,75 \pm 1,60$ pixel, une déviation angulaire moyenne de $18,45 \pm 12,98\,°$, une distance de Bhattacharyya moyenne de $0,06 \pm 0,05$, une distance de Hellinger moyenne de $0,22 \pm 0,11$,

une distance moyenne entre les descripteurs DAISY de $0,02 \pm 0,02$ et une distance moyenne globale entre les descripteurs spéculaires de $31,36 \pm 34,35$. Ces valeurs ont été calculées comme la moyenne et l'écart-type sur l'ensemble du jeu de données.

Les observations préliminaires, comme illustrées en figure 3.14a, ont révélé une tendance à la sous-segmentation des taches spéculaires par le modèle. Cela peut conduire à l'omission de certaines taches spéculaires durant le processus de segmentation ou lors de l'appariement à un correspondant manquant car non segmenté.

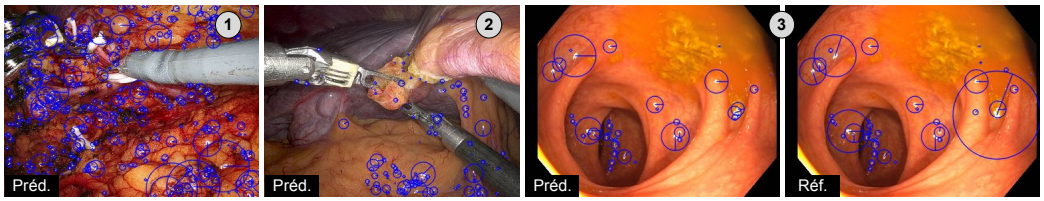
Pour élucider les effets de certains phénomènes sur le processus de synchronisation, des simulations ont été réalisées. Elles visent à évaluer l'influence de la détection et de la segmentation sous-optimales des taches spéculaires sur les courbes de réponse de synchronisation. Trois scénarios distincts ont été mis en œuvre en utilisant une séquence de référence courte, composée de 21 images, extraite de notre jeu de données d'évaluation présenté en section 3.8.1. Cette séquence brute, numérotée 8, est précédemment référencée dans la figure 3.13a. Les scénarios sont les suivants :

- Scénario 1 : implique la détection imprécise d'un sous-ensemble aléatoire de 10 %, 20 % et 30 % des taches spéculaires.
- Scénario 2 : se concentre sur la segmentation inexacte d'un sous-ensemble aléatoire de 10 %, 20 % et 30 % des taches spéculaires.
- Scénario 3 : combine à la fois la détection et la segmentation sous-optimales d'un sous-ensemble aléatoire de 10 %, 20 % et 30 % des taches spéculaires.

Les courbes de réponse, représentées dans la figure 3.14b, ont été obtenues en exécutant chacune des 9 expériences 1 000 fois. En outre, la figure 3.14 inclut l'erreur de synchronisation dans des conditions d'évaluation de la sensibilité (les courbes de réponse) en fonction de la dégradation de la segmentation des taches spéculaires. Ces résultats suggèrent que la synchronisation est peu affectée tant que la spécificité et la répétabilité de la détection des taches sont maintenues.

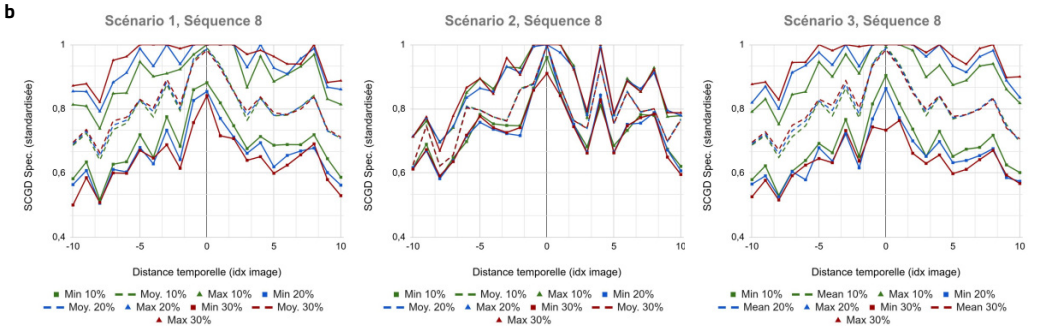
Pour garantir la capacité du système global à resynchroniser les flux vidéo stéréoscopiques, il est crucial de respecter un critère de réponse supérieur à 77 % pour la segmentation des taches spéculaires. Nos expériences démontrent que le modèle proposé dépasse ce seuil, validant ainsi son efficacité dans des conditions difficiles.

a



	Score DICE	Taux de détection	Erreur de Localisation (px)	Erreur Angulaire (°)	Distance de Bhattacharyya	Distance de Hellinger	Distance DAISY	Distance Globale
①	0,997	0,996	0,072	0,665	0,003	0,059	0,002	1,208
②	0,924	0,947	0,684	7,888	0,022	0,146	0,010	10,344
③	0,879	0,986	0,411	25,786	0,027	0,162	0,013	32,677

b



c

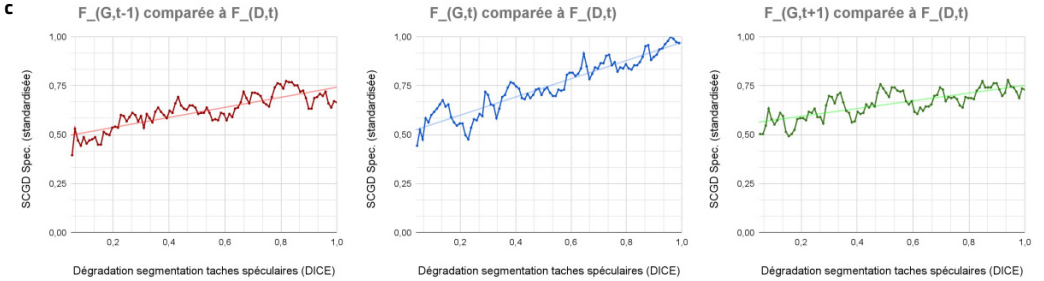


FIGURE 3.14 – (a) Évaluation complète de la détection et de la segmentation des taches spéculaires dans les images chirurgicales : une analyse à la fois qualitative et quantitative. (b) Représentation graphique des courbes de réponse à travers trois scénarios simulés : 1 - détection imprécise d'un sous-ensemble aléatoire de taches spéculaires, 2 - segmentation inexacte d'un sous-ensemble aléatoire de taches spéculaires et 3 - détection et segmentation sous-optimales d'un sous-ensemble aléatoire de taches spéculaires. (c) Analyse de l'erreur de synchronisation en relation avec la dégradation de la segmentation des taches spéculaires dans des conditions d'évaluation de la sensibilité. Source : auteur.

3.8.5 Évaluation de l'impact du score sur la synchronisation

Les performances de la méthode de synchronisation sont examinées en détail sur 18 séquences désynchronisées, soit 3 600 images générées par le simulateur de désynchronisation (cf. section 3.7) basé sur les séquences de références longues, composée de 200 images, extraites de notre jeu de données d'évaluation présenté en section 3.8.1. Les paramètres de simulation sont adaptés au flux de 60 Hz. Sans lissage temporel, nous atteignons un décalage temporel moyen de $0,055$ image et un décalage temporel maximal de 1 image pour le SCGD spéculaire proposé, avec $94,1 \pm 3,6\,\%$ des images correctement synchronisées. Le SCGD existant le plus performant atteint un décalage temporel moyen et maximal de $0,3$ et 3 images, avec $81,2 \pm 6,4\,\%$ d'images correctement synchronisées. L'utilisation du SCGD diffus-spéculaire proposé et de la fonction de lissage EWMA augmente les performances à $98,6 \pm 0,6\,\%$ d'images correctement synchronisées. Le temps nécessaire pour traiter chaque nouvelle paire d'images entrantes, depuis la réception des images jusqu'à la sortie de la paire synchronisée, en incluant toutes les étapes de détection d'image (3 millisecondes), de description (3 millisecondes), de correspondance (7 millisecondes) et de sélection (1 milliseconde), est inférieur à 14 millisecondes en utilisant le multithreading sur un ordinateur équipé d'Ubuntu 20.04, d'un processeur Intel Core™ i9-10900K et une carte graphique Nvidia RTX 2080Ti, en dessous du taux de rafraîchissement de 60 Hz, correspondant à une période de 16,67 millisecondes.

Une étude statistique vise à évaluer l'efficacité du SCGD spéculaire ou du SCGD diffus-spéculaire proposés (Méthode A) et d'un SCGD existant de haute performance (Méthode B). La comparaison se concentre sur deux métriques principales : le décalage temporel moyen et maximal ainsi que le pourcentage d'images correctement synchronisées. L'hypothèse nulle (H0) est "Il n'y a pas de différence significative entre la méthode proposée et la méthode de référence en termes de décalage temporel moyen, de décalage temporel maximal et de taux de succès de la synchronisation". L'hypothèse alternative (H1) est "Il y a une différence significative entre la méthode proposée et la méthode de référence pour ces métriques". Pour l'analyse, un niveau de risque est fixé à la valeur commune de $\alpha = 0,05$. Si $p < \alpha$, l'hypothèse nulle est rejetée, suggérant une différence significative entre les méthodes. Si $p \geq \alpha$, l'hypothèse nulle n'est pas rejetée, suggérant aucune différence significative.

Dans cette étude, des méthodes statistiques appariées ont été utilisées pour évaluer les mêmes flux vidéo avec la Méthode A et la Méthode B, facilitant ainsi une comparaison directe des performances. Un t-test apparié est utilisé pour l'analyse des décalages temporels, tandis qu'un z-test est employé pour l'analyse des taux de succès, avec des valeurs p respectives de $5,47 \times 10^{-9}$ et $9,10 \times 10^{-14}$ indiquant la significativité des résultats et de l'acceptation de l'hypothèse H1. Ces analyses ont fourni des informations détaillées sur les différences de performance, confirmant la supériorité statistique de la Méthode A sur la Méthode B.

3.8.6 Applications en chirurgie assistée par ordinateur

Le système de synchronisation passive présenté dans ce chapitre a été conçu pour un usage en contexte chirurgical (cf. section 3.2), dans le but d'optimiser la synchronisation des flux stéréoscopiques exploités par des systèmes tiers de CAO et CGI (cf. section 1.2.4.4).

La qualité de cette synchronisation, quantifiée par la distance temporelle δ entre les vues gauche et droite d'une caméra stéréoscopique, constitue un paramètre critique influençant directement la performance et la fiabilité des tâches de vision par ordinateur. Afin d'évaluer de manière quantitative les effets d'une désynchronisation temporelle sur la chaîne de traitement, une série d'expériences a été menée sur trois applications représentatives : l'étalonnage géométrique des caméras stéréoscopiques, l'estimation de la disparité et la reconstruction de surfaces peropératoires par stéréovision.

Dans le cadre de l'évaluation de l'étalonnage géométrique, dix jeux de données semi-synthétiques ont été générés à partir d'une séquence d'étalonnage initialement synchrone, capturée à une fréquence constante de 60 Hz. Le simulateur de désynchronisation (cf. section 3.7) a permis d'introduire un décalage temporel δ de une à dix images entre les vues gauche et droite. Chaque séquence comporte 20 images gauches et 20 images droites. Ces dernières sont identiques pour toutes les séquences, tandis que les images gauches sont avancées temporellement selon la valeur de δ, avec $\delta < 0$. Pour chaque jeu de données, dix étalonnages stéréoscopiques successifs ont été réalisés selon la méthode décrite en section 2.2.6.

Les résultats illustrés en figure 3.15 montrent l'évolution des paramètres internes et externes d'une caméra endoscopique Da Vinci Xi (Intuitive Surgical) en fonction de la distance temporelle δ entre les images gauche et droite d'une paire stéréoscopique.

Les paramètres internes ne sont pas affectés par la désynchronisation, les valeurs de distance focale et de position du centre optique sont quasi-identiques avec une variation relative inférieure à 0,03 % à gauche et 0,18 % à droite. Cette légère asymétrie résulte de la nature fixe de la séquence droite, tandis que la séquence gauche varie selon le décalage simulé. L'erreur maximale de reprojection associée à l'estimation des paramètres internes reste faible, elle n'excède pas 0,011 pixel à gauche et 0,055 pixel à droite. En revanche, l'estimation des paramètres externes est fortement perturbée par la désynchronisation, avec des erreurs absolues maximales atteignant 1 degré pour la rotation et 7 millimètres pour la translation. Pour référence, la ligne de base stéréoscopique d'un endoscope Da Vinci Xi (Intuitive Surgical) est d'environ 4 millimètres. L'erreur de reprojection globale du système stéréoscopique varie de 2 pixels pour une distance temporelle d'une image, soit 16,67 millisecondes, à 11 pixels pour une distance temporelle de dix images.

Les paramètres externes d'un système stéréoscopique se révèlent donc particulièrement sensibles aux effets de désynchronisation. Cette instabilité se propage ensuite à l'ensemble des traitements aval. La validité des paramètres géométriques estimés lors de l'étalonnage est donc essentielle.

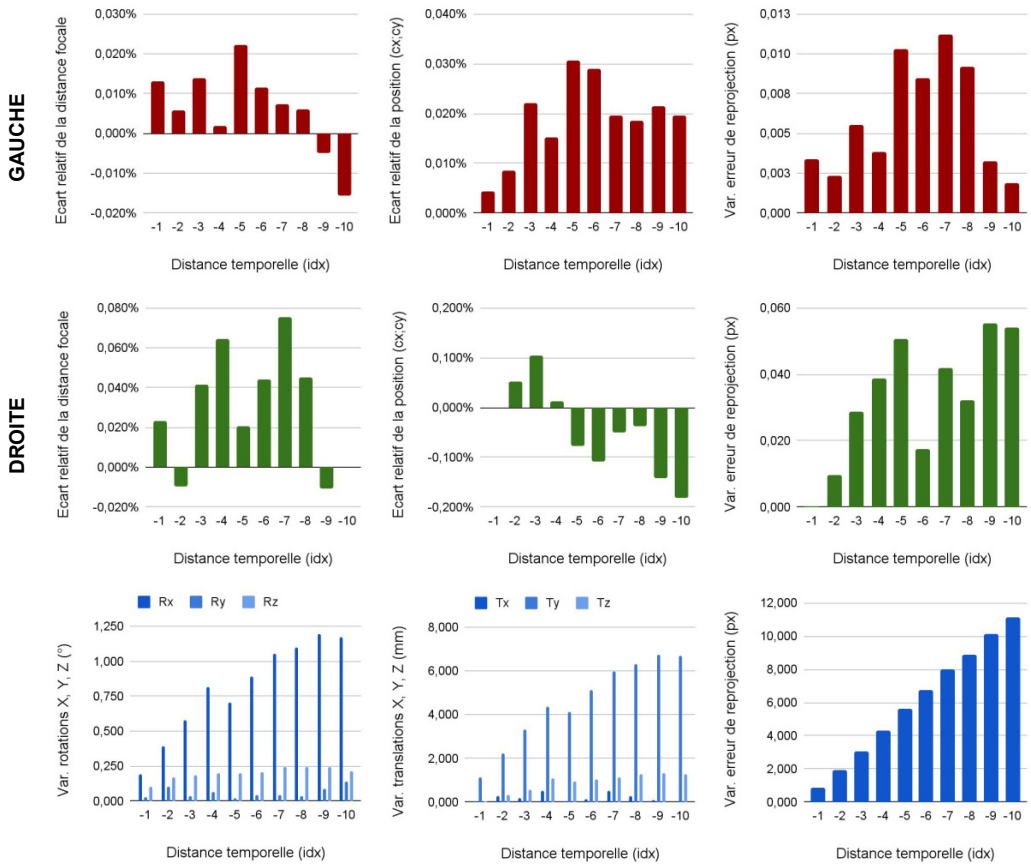


FIGURE 3.15 – Évaluation quantitative de la variation des paramètres d'étalonnage géométrique d'une caméra endoscopique Da Vinci Xi (Intuitive Surgical) en fonction de la distance temporelle entre les images gauche et droite d'une paire stéréoscopique. Source : auteur.

L'étude de la seconde application, à savoir l'estimation des cartes de disparité, a été menée en considérant un système stéréoscopique parfaitement étalonné en condition synchrone. L'introduction d'un simple retard d'une image ($|\delta = 1|$, soit 16,67 millisecondes) induit plusieurs inexactitudes dans l'estimation. Elles se manifestent visuellement par des discontinuités, des artéfacts topographiques (trous, pics) et une perte de fidélité dans les détails de surface. Ces dégradations affectent à la fois les approches géométriques comme le SGBM [Hirschmuller 2005, Hirschmuller 2007] et les méthodes basées sur l'apprentissage profond comme FoundationStereo [Wen 2025], classée première parmi 272 approches sur le benchmark Middlebury Stereo [Scharstein 2002]. Des exemples comparatifs en scénarios quasi-statiques et dynamiques sont présentés en figure 3.16 pour une caméra endoscopique Da Vinci Xi (Intuitive Surgical) dans des conditions semi-synthétiques générées avec le simulateur de désynchronisation (cf. section 3.7).

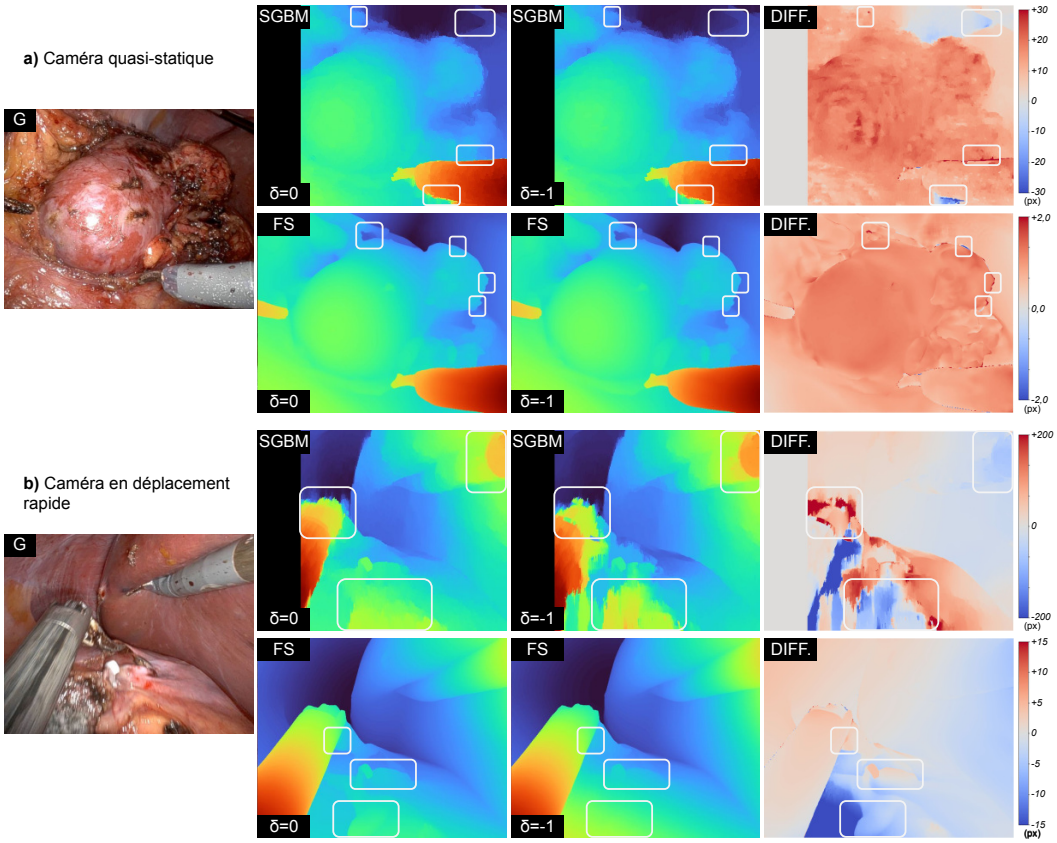


FIGURE 3.16 – Évaluation qualitative et quantitative de l'estimation de la disparité par Semi-Global Block Matching (SGBM) et FondationStereo (FS) en conditions synchrone (distance temporelle $\delta = 0$) et asynchrone (distance temporelle $\delta = -1$). Deux scénarios sont considérés : a) caméra quasi-statique et b) caméra en déplacement rapide. Source : auteur.

Pour un corpus de 184 660 paires d'images rectifiées, l'erreur moyenne absolue entre les cartes de disparité générées en conditions synchrones et asynchrones est évaluée. Les séquences stéréoscopiques asynchrones sont générées à l'aide du simulateur de désynchronisation (cf. section 3.7) pour trois distances temporelles $|\delta|$ de une, deux et dix images. Comme précédemment, les méthodes d'estimation de la disparité SGBM et FoundationStereo sont utilisées. Les résultats sont présentés dans le tableau 3.2.

Distance temporelle δ (nb. images)	Erreur moyenne absolue de disparité (pixels)	
	SGBM	Modèle FS
-1	20,09 ± 63,13	3,71 ± 14,12
-2	32,07 ± 75,37	6,05 ± 20,04
-10	93,81 ± 132,71	17,69 ± 35,85

TABLE 3.2 – Erreur moyenne absolue (en pixels) entre les 184 660 cartes de disparité générées via Semi-Global Block Matching (SGBM) et FoundationStereo (FS) en conditions synchrones et asynchrones pour différentes distances temporelles δ. Source : Auteur.

Du point de vue qualitatif, le modèle FoundationStereo permet de part sa construction d'estimer de manière dense les cartes de disparité. En revanche, pour obtenir des résultats visuellement similaires avec le SGBM un post-traitement est nécessaire via un filtre Weighted Least Squares (WLS) par exemple. Sur le plan quantitatif, l'erreur moyenne absolue de disparité tend à croître proportionnellement à l'amplitude des mouvements observés ainsi qu'à l'augmentation de la distance temporelle entre les vues. FoundationStereo se montre généralement plus robuste à la désynchronisation comparativement au SGBM.

La dernière d'application concerne la reconstruction de surface peropératoire fondée sur la stéréovision. Comme précédemment, les paramètres d'étalonnage sont considérés fiables car l'étalonnage géométrique a été réalisé en condition synchrone. Les cartes de disparité sont issues des méthodes SGBM et FoundationStereo, puis sont converties en cartes de profondeur selon la méthode décrite en section 2.2.6. Les nuages de points correspondant sont ensuite générés. Des exemples comparatifs sont illustrés en figure 3.17 pour une caméra endoscopique Da Vinci Xi (Intuitive Surgical) quasi-statique et en déplacement rapide, dans des conditions semi-synthétiques générées avec le simulateur de désynchronisation (cf. section 3.7).

L'analyse quantitative, conduite à nouveau sur 184 660 paires d'images rectifiées, met en évidence une augmentation de l'erreur moyenne absolue de profondeur avec l'augmentation de la distance temporelle $|\delta|$, comme présenté dans le tableau 3.3. Les séquences asynchrones simulées montrent que la méthode géométrique SGBM conserve une précision métrique élevée (erreur moyenne inférieure à 1 millimètre même pour $\delta = -10$), tandis que FoundationStereo présente une erreur systématique

plus importante (supérieure à 2 millimètres). Ces résultats soulignent que, malgré leurs performances apparentes sur des benchmarks standards, les approches basées sur l'apprentissage profond restent peu adaptées aux modules d'estimation de la profondeur en conditions asynchrones difficiles.

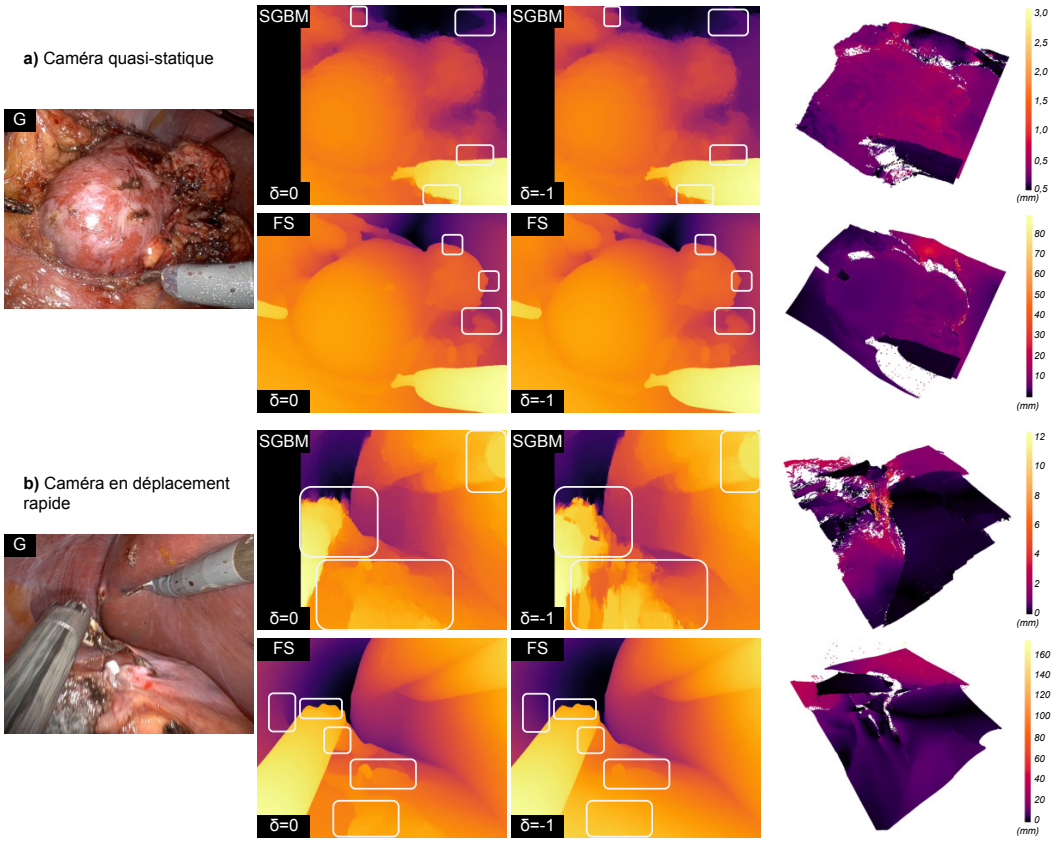


FIGURE 3.17 – Évaluation qualitative et quantitative de la reconstruction de surface peropératoire à partir de cartes de disparité obtenues en conditions synchrone (distance temporelle $\delta = 0$) et asynchrone (distance temporelle $\delta = -1$). Deux scénarios sont considérés : a) caméra quasi-statique et b) caméra en déplacement rapide. Les cartes de profondeur et les nuages de points sont dérivés des disparités obtenues par Semi-Global Block Matching (SGBM) et FoundationStereo (FS). Source : auteur.

Ainsi, la synchronisation temporelle précise entre les vues stéréoscopiques est un prérequis indispensable à toute chaîne de traitement tridimensionnel fiable. La désynchronisation, même faible ($|\delta = 1|$), induit une dégradation cumulative des performances, affectant l'ensemble des modules algorithmiques depuis l'étalonnage jusqu'à la reconstruction géométrique. La sensibilité de ces fonctionnalités souvent mises en œuvre dans les systèmes de CGI a donc un impact direct sur la performance du dispositif et par extension un impact potentiel sur la performance clinique.

Distance temporelle δ (nb. images)	Erreur moyenne absolue de profondeur (mm)	
	SGBM	Modèle FS
-1	0,15 ± 0,34	2,37 ± 41,53
-2	0,25 ± 0,48	3,16 ± 41,94
-10	0,76 ± 0,88	6,98 ± 43,20

TABLE 3.3 – Erreur moyenne absolue (en millimètres) entre les 184 660 cartes de profondeur générées à partir de disparités calculées via Semi-Global Block Matching (SGBM) et FoundationStereo (FS) en conditions synchrones et asynchrones pour différentes distances temporelles δ. Source : Auteur.

3.9 Conclusion

Dans ce chapitre, nous avons proposé une chaîne de traitement pour synchroniser les flux vidéo stéréoscopiques chirurgicaux de manière passive via la mise en œuvre d'un gestionnaire de synchronisation. Notre solution tire pleinement parti du contenu visuel des images chirurgicales en exploitant à la fois les réflexions diffuses et spéculaires avec les nouveaux SCGD Spéculaire et Diffus-Spéculaire proposés. Contrairement aux approches actives qui nécessitent une instrumentation spécifique, la méthode proposée est facilement intégrable dans l'environnement chirurgical.

Les différentes expériences menées ont montré que l'utilisation de la réflexion spéculaire apporte un avantage important pour la résolution du problème de synchronisation des flux vidéo stéréoscopiques chirurgicaux en temps réel grâce à ses propriétés hautement discriminantes. Cela est tout particulièrement vrai dans les scénarios où la caméra endoscopique est quasi-statique. Actuellement, la méthode est principalement limitée par le nombre de taches spéculaires détectées sur les surfaces tissulaires humides, condition sine qua non pour tirer parti de la composante de réflexion spéculaire.

L'ensemble des résultats expérimentaux met en lumière le rôle central de la synchronisation temporelle dans les traitements stéréoscopiques peropératoires. Les effets mesurés, tant au niveau de l'étalonnage, de l'estimation de la disparité que de la reconstruction de surface, soulignent un point fondamental de vigilance. Cette observation peut justifier la nécessité d'intégrer un mécanisme de synchronisation, même passif, au sein des chaînes de traitement utilisées en CGI.

Dans les travaux futurs, nous nous concentrerons sur l'intégration de la méthode proposée à un système CGI pour la NP et sur les tests en conditions réelles en salle d'opération, afin de démontrer l'efficacité du système en temps réel pour récupérer des flux vidéo stéréoscopique synchronisés de manière fiable.

L'omniprésence de l'apprentissage profond et ses capacités croissantes pourront nous conduire à explorer l'apprentissage de la fonction de comparaison gauche-droite. Pour cela un modèle de type siamois pourra être mise en œuvre pour l'encodage des caractéristiques des flux gauche et droit, suivi d'un module de régression pour l'inférence du SCGD. Les données nécessaires à l'entraînement du modèle pourront être générées par notre simulateur de désynchronisation.

Chapitre 4

Estimation automatique d'une pose globale par apprentissage profond

---

Dans ce chapitre, nous nous intéressons plus particulièrement à la problématique centrale du recalage 3D-2D, rencontrée lors de l'implémentation d'un système de RA en contexte chirurgical. Ce processus vise à aligner avec précision le jumeau numérique préopératoire du patient avec les images peropératoires acquises par la caméra endoscopique. L'application à la chirurgie du rein demeure un défi non résolu, principalement en raison de l'absence de repères anatomiques clairement identifiables à la fois dans les données préopératoires et peropératoires. Afin de surmonter cette limitation, nous proposons une méthode fondée sur un modèle générique de rein combiné à un réseau de neurones artificiels end-to-end. Ce réseau est entraîné sur un ensemble de données spécifiquement généré pour régresser directement la pose de la caméra à partir d'une image endoscopique. L'évaluation quantitative réalisée sur sept cas cliniques met en évidence une forte concordance avec les recalages manuels réalisés par des experts chirurgiens, malgré la variabilité anatomique inter-patient. Les résultats obtenus révèlent une erreur moyenne d'alignement du contour de la tumeur de $7,3 \pm 4,1$ millimètres, pour un temps d'inférence de $9,4 \pm 0,2$ millisecondes. Cette approche de recalage 3D-2D, exempte de tout repère externe, satisfait les exigences cliniques en matière de précision, de rapidité d'exécution et de contraintes en ressources matérielles et humaines rencontrées dans la pratique clinique. Elle représente ainsi un outil prometteur pour la NP assistée par RA.

Sommaire

4.1 Introduction 165
4.2 Contexte technique général 167
4.3 État de l'art 170
4.3.1 Méthodes de recalage 3D-2D en chirurgie abdomino-pelvienne 170
4.3.2 Recalage 3D-2D en néphrectomie partielle 176
4.3.3 Récupération de la pose humaine à partir d'images monoculaires 178
4.4 Analyse de la méthodologie de recalage 179
4.5 Approches de recalage 3D-2D basées sur des primitives . . 184
4.5.1 Méthodes de recalage 3D-2D manuelles 184
4.5.2 Méthodes de recalage 3D-2D semi-automatiques 189
4.5.3 Méthodes de recalage 3D-2D automatiques 190
4.6 Approche de recalage 3D-2D basée sur l'apprentissage automatique 193
4.6.1 Généralités 193
4.6.2 Collecte des données 194
4.6.3 Labellisation des données 195
4.6.4 Préparation des données 197
4.6.5 Entraînement du réseau de neurones 201
4.6.6 Inférence 203
4.7 Expériences 203
4.7.1 Évaluation de la performance brute 203
4.7.2 Évaluation dans le contexte clinique 204
4.7.3 Application 206
4.8 Conclusion 208

4.1 Introduction

Dans les systèmes de CAO est plus spécifiquement en CGI (cf. section 1.2.4.4), la fusion de données unimodales ou multimodales constitue un levier important pour guider et sécuriser le geste chirurgical. Ce processus, communément désigné par le terme de recalage, permet d'établir une correspondance spatialement et/ou temporellement entre des jeux de données numériques hétérogènes, acquis à différentes étapes de la prise en charge, qu'elles soient préopératoires ou peropératoires.

En RA, le recalage 3D-2D est un module clé pour la génération d'images peropératoires augmentées. Dans cette application particulière, le jumeau numérique tridimensionnel du patient construit lors de la planification chirurgicale (cf. section 1.3.1) est aligné avec la vue bidimensionnelle peropératoire acquise en temps réel par une caméra chirurgicale, qu'il s'agisse d'un endoscope en CMI ou d'une caméra montée dans le scialytique en CO. Cet alignement conduit dans l'idéal à une parfaite superposition du jumeau numérique avec l'image et permet au chirurgien de visualiser par transparence les structures anatomiques à protéger ou à réséquer, comme illustré en figure 4.1.

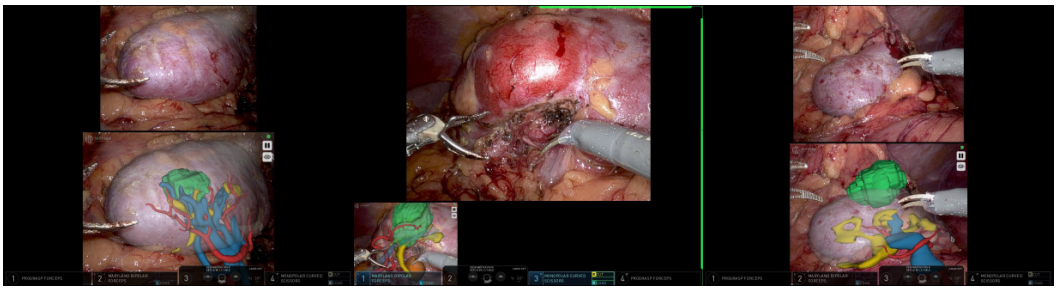


FIGURE 4.1 – Exemples de vues peropératoires augmentées en Néphrectomie Partielle Robot-Assistée (NPRA), affichées en temps réel dans la console du robot Da Vinci Xi (Intuitive Surgical). Les structures anatomiques augmentées sont au nombre de cinq : le parenchyme rénal en gris, la tumeur en vert, la veine rénale en bleu, l'artère rénale en rouge et le système collecteur en jaune. Source : auteur.

Dans le contexte spécifique de la chirurgie rénale et plus particulièrement de la NP, le guidage par RA présente un potentiel clinique significatif, notamment lors des phases chirurgicales complexes ou à risque élevé (cf. section 2.1.4). Contrairement à d'autres organes de la cavité abdomino-pelvienne comme l'utérus, le foie et la rate, le rein se distingue par sa localisation rétropéritonéale (cf. section 2.1.1) qui, lors d'une approche transpéritonéale (cf. section 2.1.4), ne le rend pas directement accessible, comme illustré en figure 4.2. Cette configuration anatomique, associée à la présence de tissu adipeux périrénal, limite la visibilité directe du rein et rend l'identification de repères anatomiques fiables particulièrement ardue.

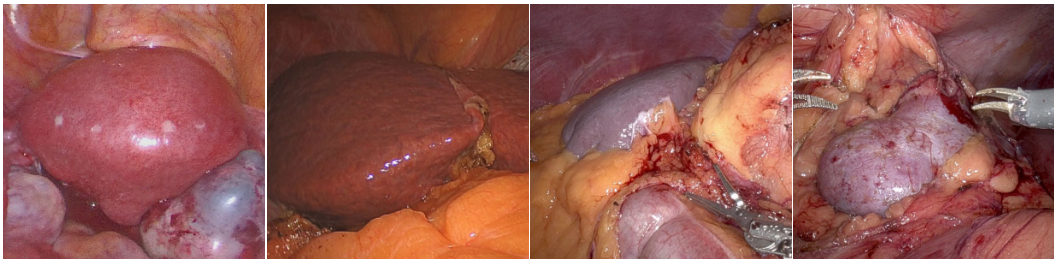


FIGURE 4.2 – (de gauche à droite) Vues peropératoires de quatre organes abdomino-pelviens : l'utérus, le foie, la rate et le rein droit après dégraissage. Source : auteur.

Les primitives anatomiques, qui servent de points d'appui pour un recalage 3D-2D robuste, sont définies comme des éléments reconnaissables qui permettent une correspondance 3D-2D fiable entre le jumeau numérique et la vue peropératoire. Leur identification est fortement tributaire du type d'intervention chirurgicale et de la technique chirurgicale employée. En NP, ces primitives sont souvent absentes. Dans de rares cas, des éléments tels que le hile rénal ou des irrégularités de surface peuvent être visibles. L'analyse des images et de la méthodologie de positionnement manuelle du jumeau numérique sur l'image peropératoire (cf. section 4.4) révèle que seuls quelques indices visuels comme la surface du parenchyme rénal ou les contours des tumeurs exophytiques sont exploitables à ce jour. Ces primitives ne suffisent pas à résoudre complètement le problème de recalage 3D-2D par manque de contraintes.

Face à l'ensemble de ces éléments clinico-techniques, nous proposons une méthode de recalage 3D-2D automatique sans primitive et applicable à la chirurgie du rein. Cette approche s'appuie sur les capacités de l'apprentissage profond, en particulier sur les réseaux de neurones artificiels end-to-end, dont l'entrée est constituée d'une image peropératoire issue de la caméra endoscopique et la sortie est la pose souhaitée, résolvant ainsi le problème de recalage 3D-2D. Nous nous affranchissons ainsi de l'utilisation explicite de primitives anatomiques habituellement nécessaires au recalage 3D-2D.

Les contributions techniques de ce travail sont de quatre ordres : 1) le développement d'une approche de recalage 3D-2D sans primitive basée sur l'utilisation d'un prototype de forme adapté à une solution end-to-end ; 2) l'élaboration d'une stratégie de génération de jeux de données labellisés avec des poses du prototype de forme, afin de pallier à la rareté des données d'entraînement ; 3) l'introduction d'une méthode de détection automatique du repère canonique du rein basée sur l'analyse de la géométrie du jumeau numérique ; et 4) la conception et l'entraînement d'un réseau de neurones artificiels end-to-end pour l'estimation de la pose dans un contexte clinique réaliste. L'évaluation de la méthode repose à la fois sur une analyse quantitative des erreurs d'alignement et sur une validation de la pertinence clinique des résultats obtenus, en particulier lors du transfert de la pose du prototype de forme vers le jumeau numérique du patient. L'analyse se concentre sur les structures anatomiques spécifiques au patient et utiles à la NP, notamment les

tumeurs, les vaisseaux et les voies excrétrices, dans le but de garantir un recalage adéquat et cliniquement pertinent.

4.2 Contexte technique général

L'un des verrous technologiques en vision par ordinateur dans le contexte de la RA pour la chirurgie réside dans le problème du recalage 3D-2D. Celui-ci consiste à estimer la pose d'un objet tridimensionnel, exprimée par une translation t et une rotation R dans l'espace $\mathbb{M}$, à partir d'observations bidimensionnelles. Bien que cette problématique soit largement étudiée dans la littérature, elle revêt une importance critique dans le contexte chirurgical, où la justesse de l'estimation conditionne directement la pertinence et la sécurité des informations superposées à la vue peropératoire.

Historiquement, deux grandes familles d'approches ont été proposées pour résoudre ce problème. La première repose sur l'utilisation de descripteurs explicites : elles consistent à extraire des caractéristiques géométriques ou photométriques (points d'intérêt, arêtes, contours, etc.) du jumeau numérique et dans l'image peropératoire, afin d'établir des correspondances. L'estimation de pose s'appuie ensuite sur la géométrie projective. Ces méthodes peuvent être performantes mais dépendent fortement de la qualité des primitives extraites et de leur stabilité dans des environnements visuellement complexes.

La seconde famille, plus récente, repose sur les méthodes d'apprentissage profond. Ces approches exploitent les capacités des réseaux de neurones convolutifs ou Convolutional Neural Network (CNN) en Anglais, à apprendre des représentations latentes à partir des données brutes, sans recourir nécessairement à une étape explicite de correspondance. Ces représentations peuvent ensuite être utilisées pour régresser directement la pose de l'objet ou inférer des correspondances implicites robustes à la variabilité des données médicales.

Un enjeu central dans ces deux familles d'approches réside dans la nécessité ou non d'une initialisation préalable. Certaines méthodes supposent une estimation approximative initiale de la pose, afin de raffiner la solution jusqu'à résoudre le problème. Cependant, cette dépendance à l'initialisation est souvent source d'erreur car la qualité du recalage final est souvent corrélée à la justesse de cette première estimation.

Sur le plan de l'interaction avec l'utilisateur, souvent le chirurgien ou un assistant, ou en d'autres termes vis-à-vis de l'aptitude à l'utilisation du système de RA, trois types d'approches se distinguent : les méthodes de recalage 3D-2D manuelles, semi-automatiques et automatiques.

L'approche manuelle consiste à laisser à l'utilisateur la charge d'ajuster la position, l'orientation et l'échelle du jumeau numérique tridimensionnel à chaque mouvement de la caméra ou de l'organe, en s'appuyant sur sa connaissance propre (expérience, notions anatomiques). Cette approche est sous présente afin de constituer une mesure de contrôle des risques, en participant à la mitigation du risque résí-

duel d'un mauvais recalage 3D-2D semi-automatique ou automatique dans le cas des dispositifs médicaux (cf. norme ISO 14971 :2019).

Les approches semi-automatiques et automatiques utilisées en chirurgie s'appuient quant à elles sur des primitives anatomiques visibles conjointement sur le jumeau numérique tridimensionnel et dans les images peropératoires. Ces repères, tels que les bifurcations vasculaires ou des contours d'organes, servent à contraindre et initialiser le processus de recalage 3D-2D. Dans le cas semi-automatique, l'identification des primitives de recalage est réalisée manuellement par l'utilisateur, tandis qu'en automatique, cette tâche est déléguée à module de détection dédié.

Sur le plan algorithmique, le recalage peut être formulé selon deux natures différentes : rigide ou déformable. Dans le cas rigide, le problème se limite à l'estimation d'une transformation composée d'une rotation et d'une translation. Dans le cas déformable, des composantes de déformation sont estimées en plus de la rotation et de la translation. Elles peuvent prendre diverses formes comme des déplacements locaux ou des paramètres d'un modèle de déformation. Ce dernier va nécessiter de modéliser le comportement des structures anatomiques sous contraintes pour différentes hypothèses. Le modèle de déformation peut tenir compte des propriétés biomécaniques des structures (module de Young et coefficient de Poisson) et de l'environnement (gravité, pression, surface de contact). Le recours à la simulation par éléments finis est alors très fréquent. Dans la suite de ce chapitre, nous nous intéressons au cas rigide avec l'objectif de proposer dans des travaux futurs une solution déformable.

Sur le plan purement technique, les méthodes d'estimation de pose de type PnP ne sont pas directement applicables, en raison de l'absence de correspondances 3D-2D connues entre le jumeau tridimensionnel et l'image peropératoire. Une solution semi-automatique consiste à faire intervenir l'utilisateur pour sélectionner manuellement des paires 3D-2D de points homologues, comme illustré en figure 4.3.

La robustesse des méthodes d'estimation de pose peut être améliorée via l'intégration de stratégies de rejet des valeurs aberrantes, comme RANSAC, qui permet d'identifier un sous-ensemble cohérent de correspondances [Fischler 1981]. Dans ce cadre, la pose est estimée à partir d'un sous-ensemble de points pour lequel le modèle ajusté produit le plus grand ensemble consensuel. Dans la méthode PnP-RANSAC, l'ensemble consensuel peut être défini comme le sous-ensemble des points pour lesquels l'erreur de reprojection est inférieure à un seuil de distance prédéfini. L'erreur de reprojection peut être définie comme le carré de la distance entre les points observés et les points projetés via le modèle.

Dans le cadre spécifique de la NP, le recalage 3D-2D présente des difficultés accrues. Les structures anatomiques impliquées (parenchyme rénal, tumeur, vaisseaux) offrent peu de repères visuels distinctifs. Les organes environnants tels que le foie, la rate ou le péritoine, bien qu'éventuellement visibles, sont peu utiles en raison de leur mobilité indépendante de celle du rein. Par ailleurs, contrairement aux objets techniques ou architecturaux, le rein ne présente ni coins nets ni arêtes régulières. Ces difficultés empêchent l'utilisation des méthodes développées pour d'autres organes et motivent le développement d'approches basées sur l'apprentissage automatique.

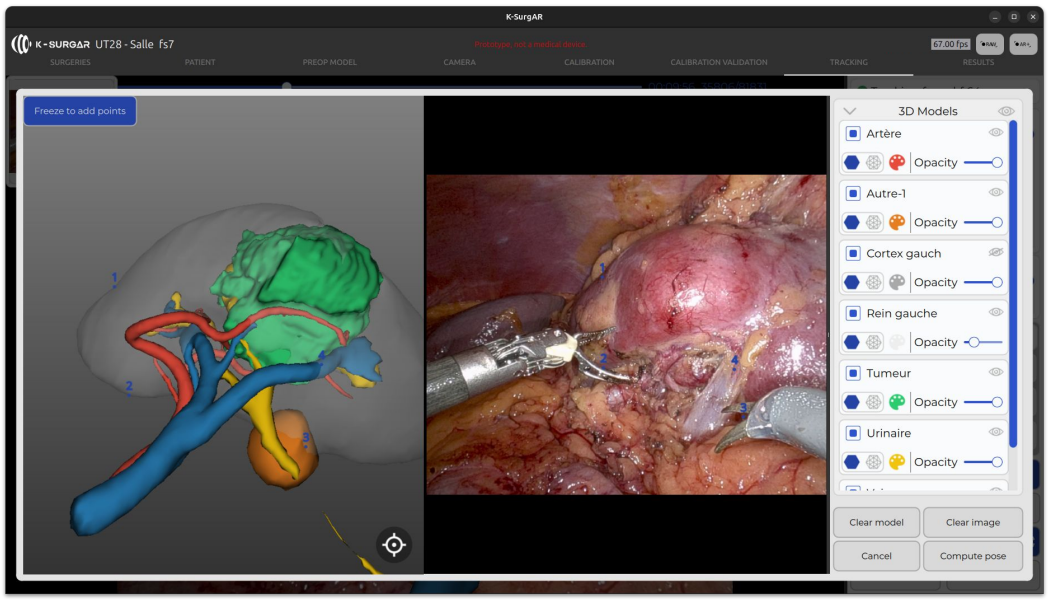


FIGURE 4.3 – Exemple d'implémentation d'une méthode semi-automatique de recalage 3D-2D par sélection manuelle de correspondance 3D-2D (en bleu) utilisées dans l'estimation de la pose de la caméra via Perspective-n-Point (PnP). Source : auteur.

Si les exigences de précision spécifiques à la chirurgie rénale ne sont pas aussi strictement formalisées que dans d'autres spécialités, certaines analogies permettent néanmoins de fixer des objectifs pertinents. Par exemple, en hépatectomie, une marge oncologique d'un centimètre est imposée pour les tumeurs hépatiques postérieures qui ne peuvent pas être localisées de manière fiable en peropératoire [Zhong 2017, Khojasteh 2025]. De même, en myomectomie, les myomes de moins de un centimètre échappent souvent à la détection et peuvent nécessiter une reprise chirurgicale en cas d'oubli [Desai 2011, Khojasteh 2025]. Ces éléments suggèrent qu'un recalage dont l'erreur est inférieure à dix millimètres peut être jugée cliniquement acceptable. Nous choisissons de conserver ce seuil pour la NP, notamment pour le traitement chirurgical des petites masses rénales localisées de moins de quatre centimètres, pour lesquels la NP est la modalité de traitement recommandée selon les directives de l'Association Européenne d'Urologie (EAU) [Bex 2025].

La méthode proposée dans ce chapitre repose sur un modèle générique de rein, appelé prototype de forme, utilisé comme support pour l'entraînement d'un réseau de neurones dédié à l'estimation de pose. L'un des principaux obstacles rencontrés est l'indisponibilité des données d'entraînement, car la vérité terrain (GT) n'est pas directement mesurable sur le patient en peropératoire. Pour remédier à cette limitation, nous avons recours à un proxy qui s'appuie sur des recalages labellisés par des experts et réalisés hors ligne, en dehors du bloc opératoire, par le biais d'une analyse rétrospective des séquences vidéo enregistrées. Nous avons, pour cela,

développé un système avancé de labellisation des données, qui permet aux experts chirurgiens de visualiser l'intégralité des images peropératoires enregistrées et de naviguer dans le flux vidéo en amont et en aval afin d'affiner l'étiquetage des poses de référence. Le jeu de données se compose ainsi de données cohérentes temporellement pour l'entraînement du réseau de recalage.

4.3 État de l'art

Dans cette section, nous nous intéresserons aux travaux existants relatifs au recalage 3D-2D et à l'estimation de pose dans le contexte chirurgical. Dans un premier temps nous nous focalisons sur les méthodes de recalage 3D-2D du jumeau numérique, obtenue lors de la planification préopératoire, sur une vue peropératoire acquise via une caméra endoscopique, spécifiquement dans les applications en chirurgie abdomino-pelvienne. Dans un second temps, nous rediscutons des approches spécifiques à la chirurgie rénale précédemment évoquée en section 2.2.4. Enfin, nous nous intéressons au problème de l'estimation de la pose humaine, en raison des similitudes qui ont inspirées la méthode proposée.

4.3.1 Méthodes de recalage 3D-2D en chirurgie abdomino-pelvienne

Le recalage 3D-2D des organes mous et déformables de la région abdomino-pelvienne constitue un champ de recherche actif. Les travaux existants se concentrent principalement sur des organes tels que l'utérus, le foie, le pancréas ou encore la prostate.

Dans ces applications, les jumeaux numériques tridimensionnels du patient sont dérivés de la phase de planification préopératoire à partir de modalités d'imagerie médicales comme l'IRM ou la TDM (cf. sections 1.3.1.1 et 1.3.1.2). Tandis que les images peropératoires sont capturées à partir d'une caméra fixe en chirurgie ouverte, tenue à la main par l'opérateur en cœlioscopie ou intégrée dans un système de chirurgie robotique (cf. sections 1.2.4.2, 1.2.4.3 et 2.2.5).

L'objectif principal en RA peropératoire est de permettre un affichage en temps réel, anatomiquement cohérent, des structures à protéger ou à réséquer. Cela vise à assister le geste chirurgical et, idéalement, à améliorer les résultats oncologiques et fonctionnels postopératoires. Un des exemples peut être la prostatectomie totale avec préservation des bandes neurovasculaires, pendant laquelle la visualisation peropératoire de l'envahissement tumoral capsulaire permet de guider la phase de dissection nerveuse [Alchin 2016]. Dans le cas de l'utérus, la visualisation directionnelle des fibres musculaires issues de la tractographie IRM peut permettre le guidage des incisions [Chauvet 2020, Bourdel 2019].

Tel que mentionné en section 4.2, le recalage 3D-2D manuel est historiquement la première catégorie de méthodes mise en œuvre pour réaliser l'alignement en peropératoire. Elle est très largement adoptée car aucune difficulté technique liée à la détection de primitives de recalage ou au calcul de transformations n'existe. L'intégralité de l'analyse repose sur la connaissance propre de l'opérateur. Cette solution est donc purement du rendu graphique intéractif afin de rendre possible la superposition visuelle et la mobilisation du jumeau numérique sur l'image. L'opérateur doit donc résoudre le problème de recalage en ajustant manuellement les trois rotations, les trois translations, l'échelle et dans certains cas un sous-ensemble de paramètres de déformation linéaires ou non-linéaires, souvent exprimés sur les grands axes du jumeau numérique. Le modèle élastique mobilisable en torsion, traction et compression peut utiliser les formules de Barr [Barr 1984] afin de simuler ces transformations sur les maillages avec peu d'effort computationnel et selon certains auteurs de bons résultats visuels [Porpiglia 2019]. Concrètement, l'opérateur, généralement une personne dédiée à l'utilisation de la solution de RA, utilise une IHM, comme un écran tactile, une souris ou une souris 3D type SpaceMouse (3DConnexion) à six degrés de liberté, pour réaliser le recalage 3D-2D.

Cette approche a été mise en œuvre dans de nombreux travaux de la littérature, avec des applications pour l'utérus [Kupas 2019], le foie [Acidi 2023, Ntourakis 2016, Pessaux 2015], le pancréas [Boul-Atarass 2025, Wu 2023a] et la prostate [Della Corte 2024, Schiavina 2021b, Bianchi 2021], dont un exemple est illustré en figure 4.4.

L'approche par recalage 3D-2D manuel est cependant chronophage et complexe, ce qui limite très fortement son application en contexte clinique de routine dû au manque de temps ou de personnel dédié. Son intégration en tant qu'approche principale dans un dispositif médical utilisable en routine clinique est donc peu probable.

Pour pallier ces limitations, une seconde catégorie de méthodes, désignées sous le nom de semi-automatiques a été développée. Celles-ci nécessitent une phase d'initialisation manuelle par l'opérateur, suivi d'un ajustement automatique. Les données d'initialisation fournies peuvent prendre différentes formes, il peut s'agir de points de repère appareillés entre le jumeau numérique et la vue caméra ou entre le jumeau numérique et un point physique tridimensionnel sur l'organe capté via un stylet suivi optiquement [Thompson 2018, Prevost 2020, Tang 2017, Okamoto 2015]. Ou encore de contours visibles sur le jumeau numérique et la vue peropératoire.

La dernière catégorie rassemble les méthodes dites automatiques, ne nécessitant aucune interaction avec le système.

Une première approche consiste à utiliser un modèle tridimensionnel peropératoire généré via des techniques de reconstruction tridimensionnelle passives [Richter 2024, Song 2017] de type Structure from Motion (SfM) en Anglais [Özyeşil 2017], stéréovision dans le cas des endoscopes stéréoscopiques ou estimation de profondeur monoculaire par apprentissage profond comme Endo-DAC [Cui 2024, Sharifian 2025b]. Le modèle peropératoire ainsi reconstruit peut être aligné avec le jumeau numérique préopératoire via recalage 3D-3D rigide ou déformable. Les méthodes algorithmiques généralement utilisées sont l'ICP ou le

CPD [Ravikumar 2017]. Cette approche est utilisée par certains auteurs en chirurgie hépatique [Onda 2013, Luo 2020a, Kang 2014, Haouchine 2014].

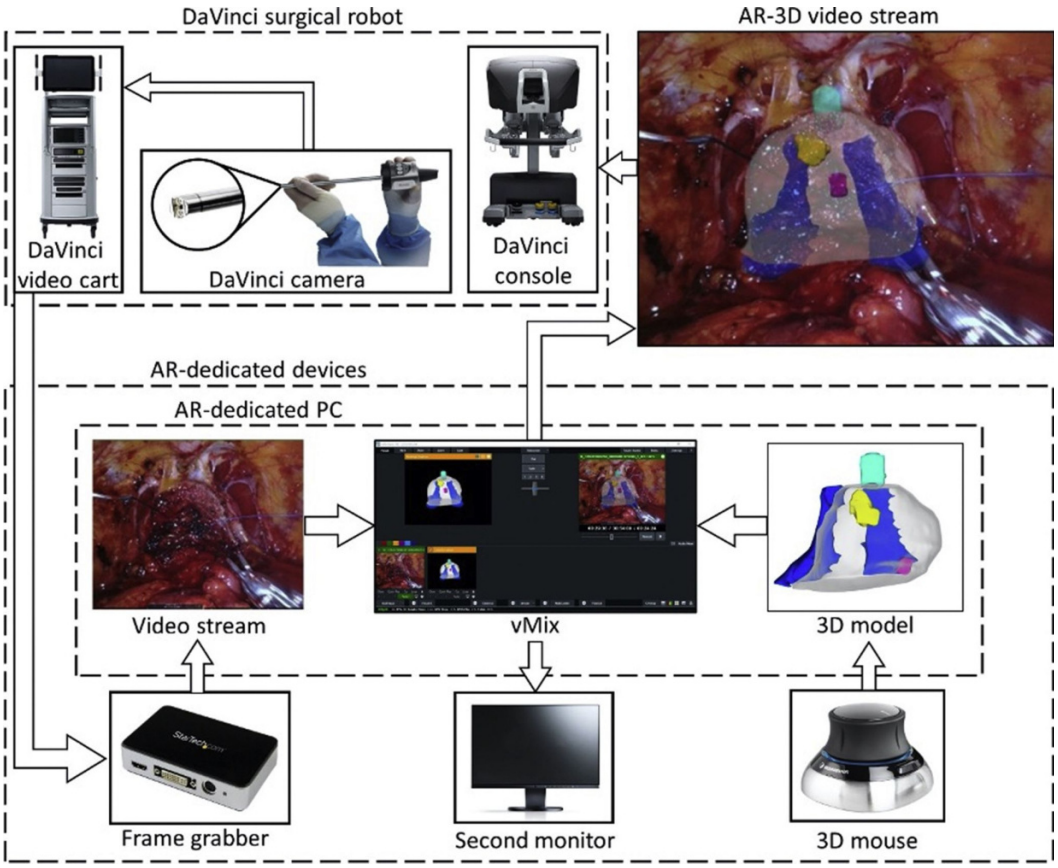


FIGURE 4.4 – Représentation schématique de la mise en œuvre d'un système de Réalité Augmentée (RA) en chirurgie de la prostate, basé sur une approche de recalage 3D-2D manuel via une souris 3D de type SpaceMouse (3DConnexion). Source : [Schiavina 2021b].

Cependant, un défaut de reconstruction ou une surface trop petite, liée à une vue partielle, peut conduire à un échec du recalage 3D-3D par manque de données ou de contraintes [Schneider 2020]. Pour renforcer les contraintes de recalage et guider l'optimisation numérique, certaines approches intègrent des primitives anatomiques robustes, comme la crête du foie [Espinel 2021, Haouchine 2016, Robu 2018, Modrzejewski 2019, Espinel 2020, Le Roy 2019, Bertrand 2020] ou les contours de l'utérus [Collins 2020, François 2020], ainsi illustrés en figure 4.5.

Les primitives anatomiques et plusieurs observations de la scène sous différents points de vue permettent d'accroître la précision du recalage et surtout d'améliorer l'estimation de la déformation avec des erreurs d'alignement inférieures à dix millimètres selon certains auteurs [Espinel 2021, Adagolodjo 2017, Koo 2017].

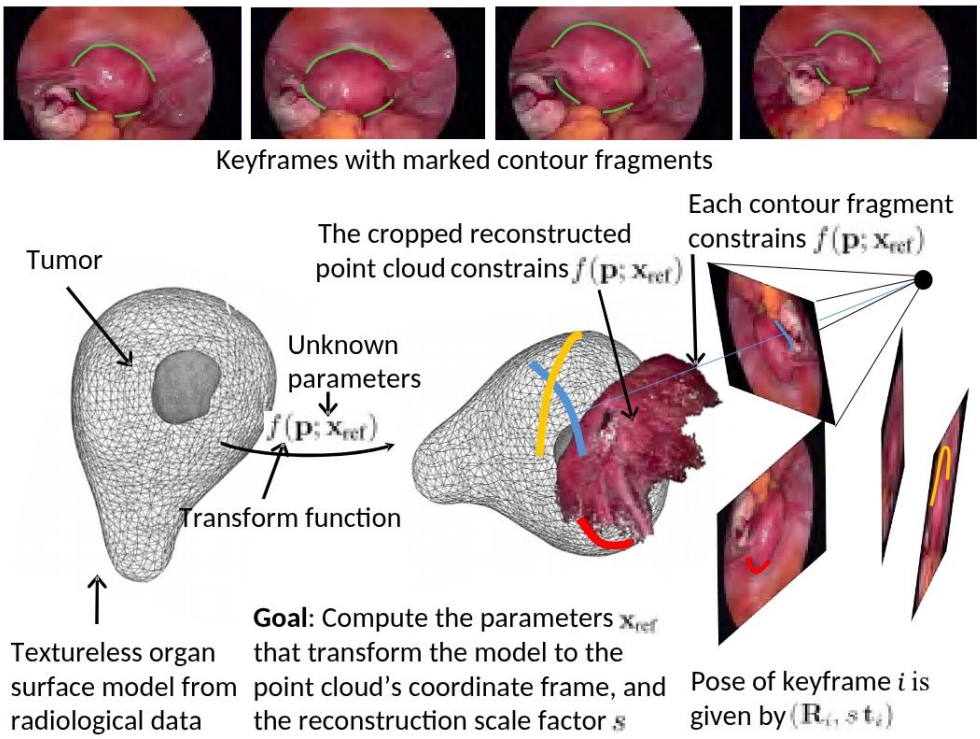


FIGURE 4.5 – Représentation schématique du processus de recalage 3D-3D déformable contraint par les primitives visuelles anatomiques de l'utérus. Source : [Collins 2020].

Le recours à un modèle de déformation biomécanique permet une modélisation réaliste du comportement des tissus sous contraintes. Ces modèles peuvent être de nature élastique ou visco-élastique, isotrope ou anisotrope. Les déplacements locaux sont alors contraints par le modèle de déformation [Özgür 2018, Collins 2020].

Une fonction de coût, exprimée comme la somme pondérée des résidus carrés non linéaires combinant les résidus de distance ainsi que les résidus d'énergie en fonction des paramètres du modèle de déformation peut être utilisée afin de déformer le jumeau numérique à la réalité observée [Collins 2020]. Cette quantité est alors minimisée par des méthodes d'optimisation numériques itératives, dont le coût calculatoire est élevé. L'utilisation en temps réel n'est donc possible qu'en réalisant d'importantes hypothèses et simplifications qui conduisent à une perte de précision.

Pour répondre à ces contraintes, des approches basées sur l'apprentissage profond ont récemment émergé. Le réseau Liver Mesh Recovery (LMR) en Anglais, illustré en figure 4.6, apprend une correspondance entre des primitives bidimensionnels hépatiques détectées dans l'image et la pose du jumeau numérique ainsi que les paramètres de déformation souhaités [Labrunie 2023].

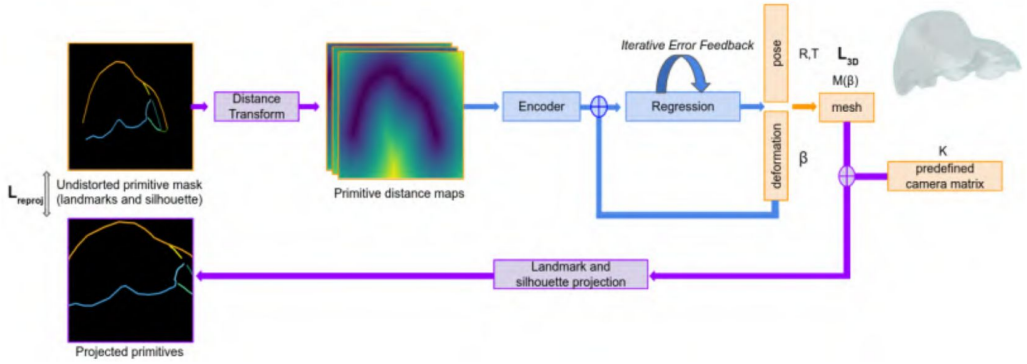


FIGURE 4.6 – Méthode de recalage 3D-2D automatique patient spécifique LMR (Liver Mesh Recovery en Anglais), basé sur des primitives anatomiques bidimensionnelles visibles à la surface du foie. Source : [Labrunie 2023].

Dans la version patient-spécifique, le réseau de neurones est entraîné sur des données synthétiques générées à partir de simulations. Un modèle déformable du foie, spécifique au patient est mobilisé et déformé, puis rendu sous de multiples points de vue ($n = 40\ 000$). Les primitives sont ensuite générées à partir des images simulées et converties en cartes de distances utilisées en entrée du réseau de neurones. Cette méthode nécessite 20 heures de calcul préopératoire pour chaque patient et permet d'obtenir selon les auteurs une erreur de recalage ou TRE moyenne de $17,9 \pm 2,8$ millimètres [Labrunie 2023].

Pour une généralisation inter-patients, un modèle générique construit à partir de 71 foies est utilisé. Les jumeaux patients sont recalés selon une méthode coarse-to-fine suivant 14 points anatomiques puis via l'utilisation de Radial Basis Functions (RBF) en Anglais [Buhmann 2000] et de Non-Rigid Iterative Closest Point (NR-ICP) en Anglais [Amberg 2007]. Le modèle générique de foie est alors construit via Statistical Shape Modelling (SSMo) en Anglais [Cootes 1992] ou Gaussian Process Morphable Model (GPMM) en Anglais [Lüthi 2017]. Ce nouveau modèle remplace le jumeau numérique du patient pour l'entraînement du réseau de neurones, via 120 000 vues synthétiques générées. L'erreur de recalage ou TRE moyen est évaluée à $28,9 \pm 18,2$ millimètres [Labrunie 2024].

Certains travaux ont exploré une approche appelée "position-based dynamics" [Müller 2007] d'un matériau continu [Bender 2014] qui n'emploie pas de modèle de déformation. En revanche, il tente de déplacer chaque sommet du maillage volumétrique en fonction de l'intégration numérique explicite de l'équation du mouvement de Newton, ainsi que des contraintes. Ces dernières peuvent être des cibles positionnelles, utilisées comme contraintes d'attachement, ainsi que des contraintes élastiques et d'ombrage [Koo 2017].

Lorsque les primitives de recalage ne sont pas clairement visibles, des stratégies alternatives ont été développées via l'utilisation de repères artificiels aussi appelés fiduciaires.

Pour la CO du pancréas, des repères artificiels de type code QR sont fixés sur la peau du patient au niveau du sternum avant la réalisation de l'imagerie préopératoire, comme illustré sur la figure 4.7. En peropératoire un marqueur monté sur un pointeur permet de localiser la veine cave qui est une primitive de recalage raisonnablement fixe. L'image augmentée peut alors être affichée sur un dispositif de visualisation comme des lunettes de réalité mixte Hololens (Microsoft) [Javaheri 2025, Tang 2021].

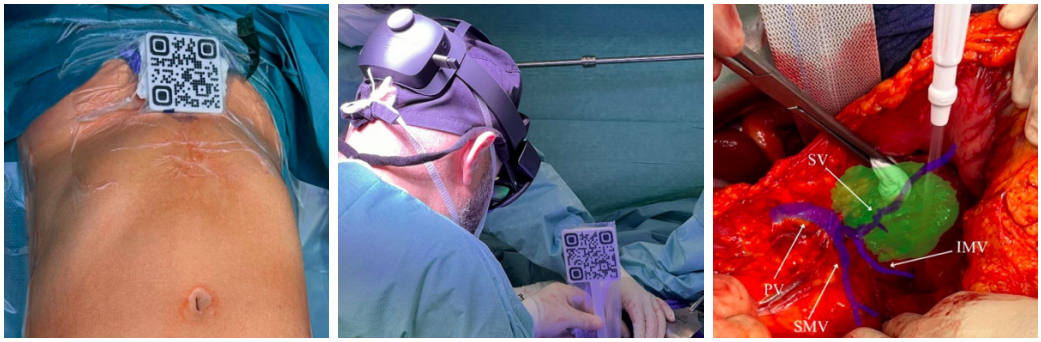


FIGURE 4.7 – Système de Réalité Augmentée (RA) en Chirurgie Ouverte (CO) du pancréas basé sur des marqueurs. Source : générée par l'auteur à partir de [Javaheri 2025].

Dans le cas de la prostate, l'échographie peropératoire peut être utilisée comme modalité intermédiaire de recalage pour composer les transformations entre le jumeau numérique et le vue endoscopique [Müller 2021].

D'autres auteurs ont proposé d'utiliser l'instrumentation chirurgicale mise en œuvre en prostatectomie comme primitive de recalage 3D-2D. Le cathéter est alors inclus dans le jumeau numérique et permet le recalage via une méthode en deux étapes, illustrée en figure 4.8. Une première étape consiste à détecter le cathéter introduit dans la loge prostatique pendant l'intervention. Une seconde étape utilise un réseau de neurones entraîné sur des images semi-synthétiques pour inférer la rotation du cathéter et in fine l'augmentation de l'image [Checcucci 2023].

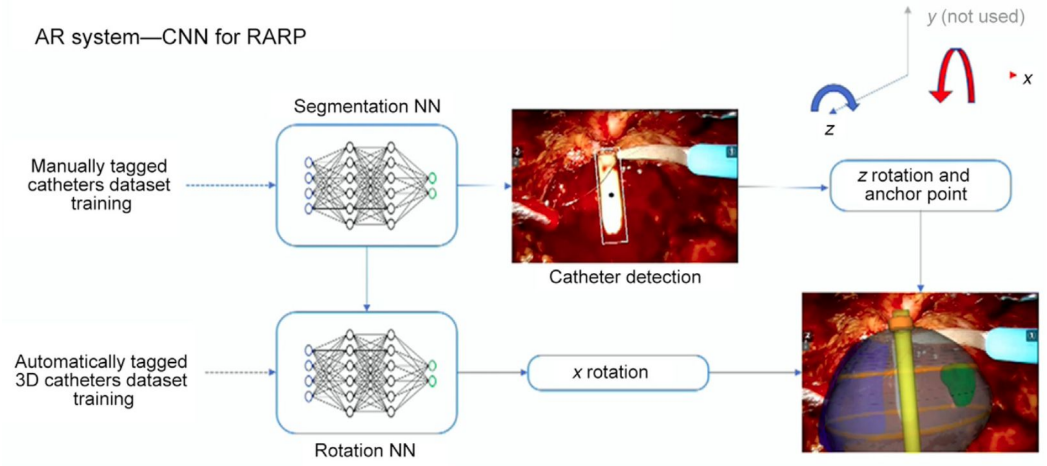


FIGURE 4.8 – Système de Réalité Augmentée (RA) en Chirurgie Mini-Invasive (CMI) de la prostate basé sur une méthode de recalage automatique utilisant l'instrumentation chirurgicale (cathéter) comme primitive de recalage 3D-2D. Source : [Checcucci 2023].

4.3.2 Recalage 3D-2D en néphrectomie partielle

Le recalage 3D-2D dans le contexte de la NP a déjà fait l'objet de la section 2.2.5 de ce manuscrit. Seuls les éléments clés de compréhension et de comparaison sont rappelés ci-après.

De manière analogue à d'autres interventions en chirurgie abdomino-pelvienne (cf. section 4.3.1), le recalage 3D-2D en NP repose majoritairement sur des approches manuelles rigides et déformables [Khaddad 2023, Della Corte 2024, Su 2009, Nakamura 2010, Pratt 2012, Chen 2014, Wang 2015, Porpiglia 2018b, Porpiglia 2020, Schiavina 2021a, Amparore 2022b].

Certaines méthodes semi-automatiques ou automatiques, notamment celles basées sur la reconstruction peropératoire [Altamar 2011, Amparore 2022a, Zhang 2019b], présentent des limitations significatives en raison de leur complexité technique et leur faible adaptabilité aux contraintes opérationnelles en milieu cliniques.

En l'absence de primitives anatomiques rénales visibles, tels que des points de référence ou une silhouette distincte de l'organe, les méthodes semi-automatiques ou automatiques de la littérature échouent à résoudre le problème d'alignement.

Une adaptation d'une méthode de recalage 3D-2D automatique présentée précédemment pour la prostate [Checcucci 2023], a récemment été proposée pour la NP. Cette méthode fondée sur l'apprentissage profond, procède en deux étapes distinctes [Piana 2024], comme illustré en figure 4.9. La première étape consiste à segmenter la surface rénale à partir d'une image peropératoire à l'aide d'un premier réseau de neurones. La position tridimensionnelle du rein et son échelle sont alors calculées en ajustant une ellipse sur le masque de segmentation binaire. La seconde étape est dédiée à la régression de l'orientation via un second réseau de neurones appelé "RotationCNN".

Plusieurs différences distinguent cette approche de la méthode proposée dans ce chapitre. Le modèle "RotationCNN" est entraîné exclusivement à partir de données synthétiques issues de la simulation de vues du jumeau numérique. Il est donc spécifique au patient et requiert un réentrainement pour chaque nouveau cas clinique. De plus, la méthode nécessite plusieurs images avec un déplacement régulier de la caméra. Lorsqu'il a été testé sur des images réelles, il a été nécessaire pour les auteurs d'ajouter à la méthodologie une troisième étape, basée sur le flux optique, pour atténuer le manque de robustesse du réseau "RotationCNN". Enfin, les auteurs soulignent que cette approche ne s'avère pas fonctionnelle en pratique clinique. Lorsque l'organe cible présente une structure complexe, comme le rein, l'initialisation du recalage doit être effectuée manuellement. Ce constat peut s'expliquer par l'écart important de domaine entre les images synthétiques d'entraînement et les images réelles lors de l'inférence.

Par conséquent, malgré les avancées significatives observées au cours de la dernière décennie, le développement d'une méthode de recalage 3D-2D cliniquement viable pour la NP reste difficile à mettre en œuvre.

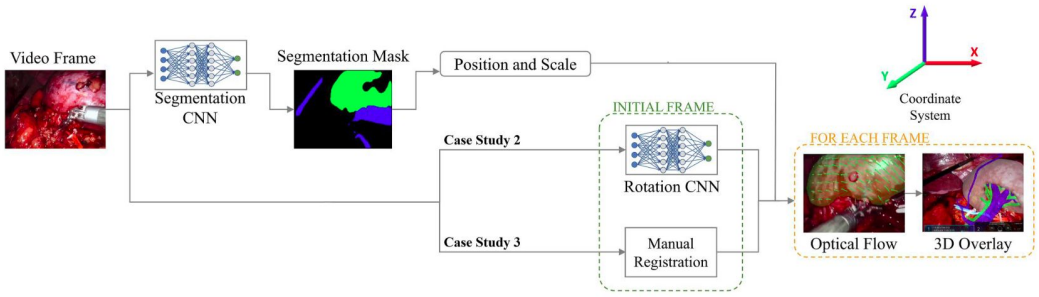


FIGURE 4.9 – Système de Réalité Augmentée (RA) en Néphrectomie Partielle Robot Assistée (NPRA) basé sur une méthode de recalage automatique utilisant la segmentation de la surface rénale comme primitive de recalage 3D-2D. Source : générée par l'auteur à partir de [Padovan 2022].

4.3.3 Récupération de la pose humaine à partir d'images monoculaires

La récupération de la pose humaine à partir d'une image monoculaire consiste a estimer un maillage tridimensionnel plausible du corps humain, cohérent à la fois avec les contraintes anatomiques humaines et les informations visuelles extraites de l'image d'entrée. Ce problème de vision par ordinateur présente des similitudes marquées avec celui du recalage 3D-2D qui nous occupe, notamment par la nécessité d'estimer une configuration tridimensionnelle à partir d'observations bidimensionnelles.

La plupart des méthodes contemporaines s'appuient sur la combinaison d'un modèle paramétrique de la forme du corps humain avec des représentations conjointes de la pose, comme SCAPE [Anguelov 2005] et SMPL [Loper 2023]. Pour résoudre ce problème, deux classes de méthodes existent : les méthodes basées sur l'optimisation et les méthodes basées sur l'apprentissage automatique [Tian 2023].

Les méthodes basées sur l'optimisation estiment initialement une pose et une forme grossières du corps et les affinent itérativement en ajustant les paramètres du modèle de forme pour les adapter à de multiples indices visuels tels que les silhouettes et les points de repère anatomique [Guan 2009]. Bien qu'efficaces, ces méthodes sont très consommatrices en ressources informatiques et sensibles à l'initialisation.

Les méthodes basées sur l'apprentissage ont connu des progrès significatifs suite au développement de framework end-to-end comme : Human Mesh Recovery (HMR) [Kanazawa 2018], SPIN [Kolotouros 2019] et les récents PyMAF [Zhang 2021b] et PyMAF-X [Zhang 2023a]. La méthode HMR, illustrée en figure 4.10, utilise un encodeur ResNet-50 chargé d'extraire les caractéristiques de l'image d'entrée qui sont concaténées avec la pose, les paramètres de forme et les paramètres de la caméra et introduites dans un réseau de régression composé de plusieurs couches entièrement connectées (typiquement trois pour HMR). Le régresseur exploite le retour d'erreur itératif ou Iterative Error Feedback (IEF) en Anglais, pour

affiner les prédictions de manière progressive. De plus, un discriminateur également composé de couches entièrement connectées, évalue la cohérence des estimations générées afin d'améliorer leur plausibilité en particulier dans les scénarios dépourvus de vérité de terrain tridimensionnelle.

Développée à partir des modèles de base, PyMAF intègre une structure pyramidale de caractéristiques et un mécanisme de rétroaction d'alignement pour l'affinage multi-échelle. PyMAF-X étend les capacités de PyMAF à la récupération du corps entier, obtenant des résultats à l'état de l'art dans des scénarios difficiles. Ces méthodes sans primitive constituent une source d'inspiration pour la problématique de recalage 3D-2D en NP. Leur transposition directe reste néanmoins difficile en raison de deux contraintes principales : d'une part, il n'existe pas de modèle paramétrique de forme rénale universellement accepté et d'autre part, les données d'entraînement sont rares et difficiles à collecter.

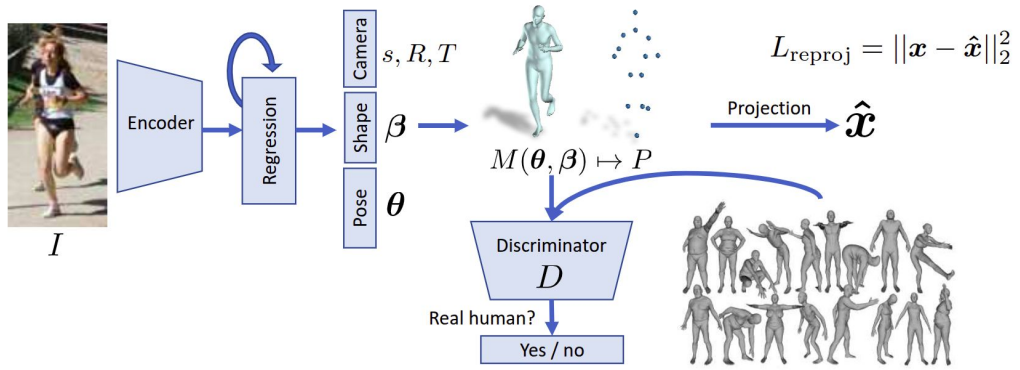


FIGURE 4.10 – Méthode de récupération de pose humaine HMR (Human Mesh Recovery en Anglais), basée sur l'analyse d'images monoculaires. Source : [Kanazawa 2018].

4.4 Analyse de la méthodologie de recalage

Comme mentionné en section 2.1.4, la NP se décompose en cinq phases chirurgicales. Le recours à la RA en peropératoire ne s'avère pas nécessaire de manière continue tout au long de l'intervention, mais pourrait jouer un rôle déterminant lors des phases chirurgicales reconnues comme complexes ou risquées. Les phases de dissection vasculaire, de délimitation de la tumeur, de résection tumorale et de reconstruction du lit tumoral sont donc concernées.

Au regard des défis et des besoins informationnels identifiés précédemment dans le tableau 2.2, la superposition du jumeau numérique sur la vue peropératoire par RA, pourrait permettre d'anticiper et de guider le geste chirurgical. Par exemple, avant la dissection rénale, la RA peut faciliter la localisation de la tumeur et la planification de la voie d'abord compte tenu de la perfusion vasculaire et de la position

de la tumeur. Durant la dissection, elle peut offrir une visualisation des structures sous-jacentes non visibles directement comme le pédicule rénal ou les vaisseaux intra-parenchymateux. Au moment de la résection tumorale, une augmentation fidèle peut fournir une visualisation précise de la tumeur et de ses marges. Enfin, pendant la phase de reconstruction, bien que la RA soit moins sollicitée, un recalage peut rester utile pour contrôler la zone réséquée par rapport au modèle initial et également aider à la fermeture des voies urinaires (calices) le cas échéant.

La mise en œuvre d'un tel guidage peropératoire par RA sans marqueurs, repose sur l'identification de primitives anatomiques de recalage. Les discussions menées avec les équipes chirurgicales ont permis de recenser plusieurs repères visuels potentiels, dont la pertinence dépend fortement de la localisation de la tumeur et de la phase chirurgicale. Un condensé de ces primitives est donné dans le tableau 4.1.

Phase chirurgicale	Structures observables dans la vue endoscopique	Primitives anatomiques potentielles pour le recalage
Accès au rein (dissection initiale)	Tissu péri-rénal (graisse, fascia) puis surface rénale	Contours flous, forme approximative du rein
Exposition du rein	Surface partielle du rein	Volume du parenchyme rénal, silhouette partielle du rein
Dissection pédiculaire (optionnelle, si tumeur hilaire)	Vaisseaux principaux (artère, veine) Hile rénal et lèvres rénales	Points de bifurcation vasculaires, formes tubulaires, points d'insertion vasculaire
Identification de la tumeur	Surface du rein, surface de la tumeur exophytique	Volume du parenchyme rénal et de la tumeur exophytique
Résection tumorale	Parenchyme rénal ouvert, tumeur exposée	Marges tumorales partielles, contours internes
Reconstruction du lit tumoral	Zone réséquée, calices exposés, branches vasculaires	Contours irréguliers, points d' émergence des calices ou vaisseaux dans le lit tumoral

TABLE 4.1 – Primitives anatomiques potentielles pour le recalage en fonction de la phase chirurgicale et des structures observables dans la vue endoscopique peropératoire. Source : auteur.

L'exploitation automatique des primitives anatomiques référencées dans le tableau 4.1 est rendue difficile par leur faible netteté et leur variabilité. Contrairement à d'autres spécialités chirurgicales ou organes, il n'existe pas de points d'ancrage géométriques fiables et constants tout au long de l'intervention de NP. La surface rénale constitue une exception partielle, mais n'impose que peu de contraintes géométriques en translation ou rotation, ce qui limite son utilité pour un recalage robuste. Ainsi, les approches de recalage 3D-2D fondées sur des correspondances géométriques explicites atteignent rapidement leurs limites.

Un exemple illustratif de cette difficulté est donné en figure 4.11. Dans l'image endoscopique les contours de l'organe sont assez facilement reconnaissables mais leur association avec des contours tridimensionnels sur le jumeau numérique est une réelle problématique.

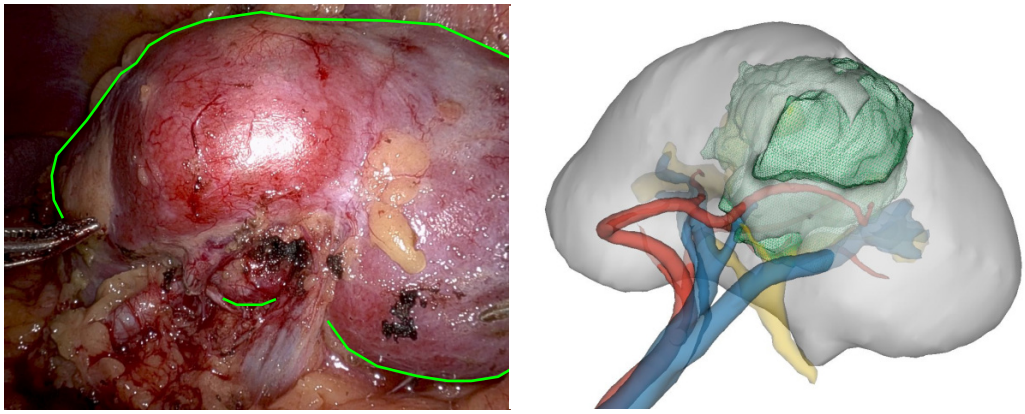


FIGURE 4.11 – Illustration de la difficulté à appareiller un contour bidimensionnel détecté (en vert) dans la vue endoscopique avec un contour tridimensionnel sur le jumeau numérique. Source : auteur.

Dans le but de mieux comprendre les mécanismes cognitifs mis en œuvre par les opérateurs humains lors du recalage 3D-2D en NP et in fine systématiser une méthodologie reproductible et transférable à un système informatique, nous avons poussé les investigations et observé l'humain.

Une application intéractive en langage Python nommée Registration Process Understanding (RPU) en Anglais a été conçue pour analyser de manière structurée la stratégie de recalage suivie par l'opérateur, de l'interprétation de l'image jusqu'à la mobilisation du jumeau numérique. Au travers de cette interface Qt5, illustrée en figure 4.12, l'opérateur chirurgien peut à partir d'une séquence vidéo étalonnée, effectuer une analyse rétrospective d'un cas réel de NP. Il dispose par l'intermédiaire de l'application d'un lecteur vidéo augmenté qui permet de visionner l'enregistrement vidéo et de mobiliser le jumeau numérique visible en superposition. Après avoir choisi une image de référence ou keyframe sur laquelle un recalage serait pertinent pour initialiser la RA, l'opérateur peut documenter sa démarche d'analyse à l'aide de commentaires textuels, oraux ou de dessins sur la vue peropératoire. Chaque événement est enrichi d'informations temporelles et spatiales, notamment les manipulations appliquées au jumeau numérique et les modifications d'apparence du maillage.

Une version plus récente de cette application intègre un mécanisme de détection automatique d'événements, déclenché par la manipulation du jumeau numérique. L'opérateur est alors invité à commenter son action. Un dictionnaire de mots-clés est également proposé afin de standardiser les explications et d'en faciliter l'analyse.

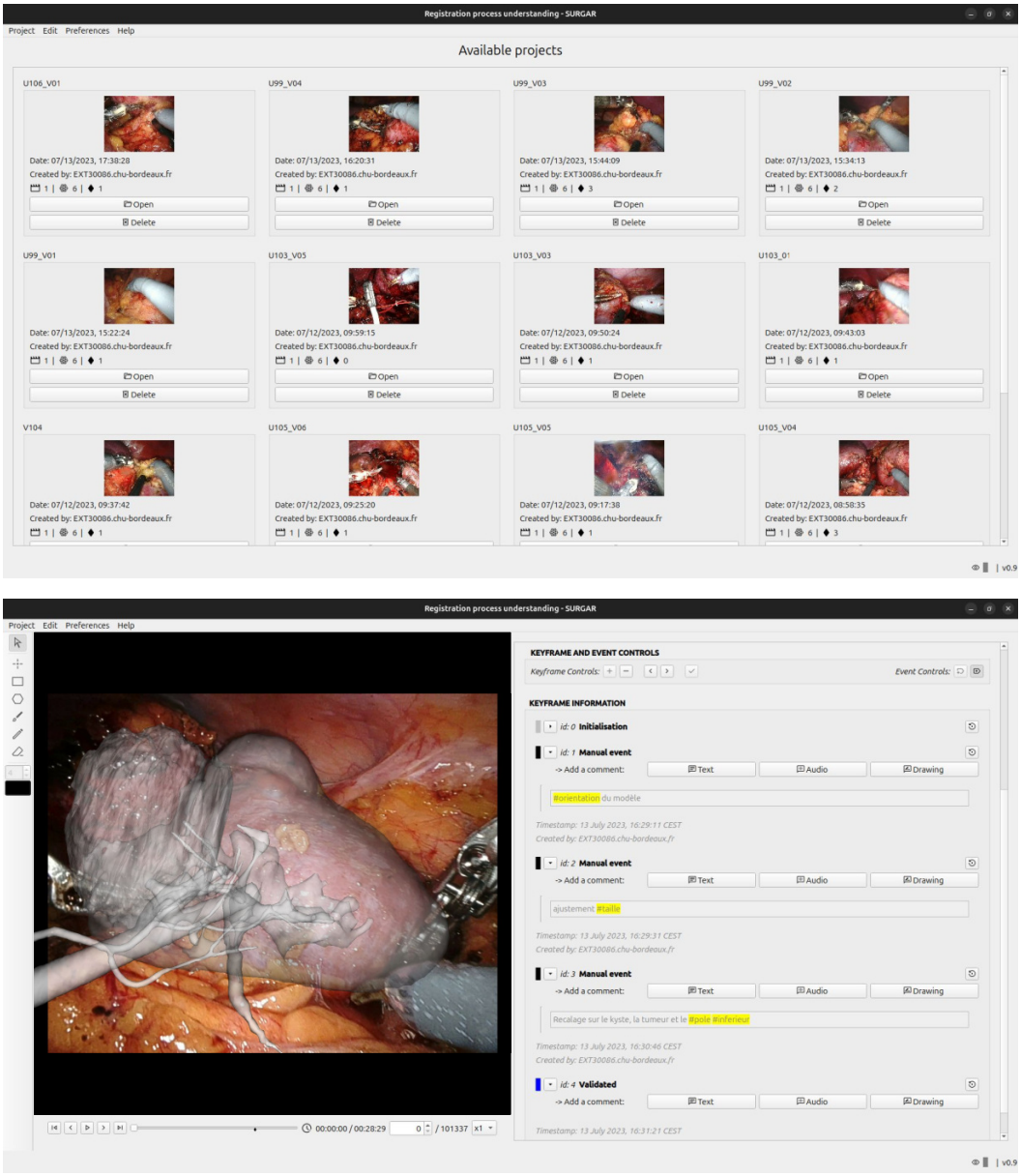


FIGURE 4.12 – Application Python proposée pour l'analyse de la méthodologie de recalage 3D-2D du jumeau numérique sur la vue peropératoire. Source : auteur.

En complément de cette approche déclarative, un module de suivi du regard (gaze tracking en Anglais) a été ajouté pour capturer les zones de focalisation du regard sur l'écran. Par hypothèse, ces zones sont constituées d'éléments visuels utiles à la prise de décision pour la réalisation du recalage 3D-2D. Le pipeline technique de suivi du regard caméra-écran est illustré en figure 4.13. Il repose sur la détection du visage de l'opérateur, filmé via une caméra étalonnée, par exemple la webcam intégrée à un ordinateur portable. Cette détection est réalisée en temps réel grâce à trois Réseaux de Neurones Artificiels (RNA) connectés en cascade. Un premier pour détecter le visage via un réseau de type BlazeFace [Bazarevsky 2019], un second pour extraire de la zone détectée 468 primitives faciales [Kartynnik 2019] et un dernier pour afiner les positions des primitives au niveau des lèvres, des yeux et des iris [Grishchenko 2020]. La position du regard est estimée à partir de ces données via un quatrième RNA de type VGG16 [Simonyan 2014b, Gudi 2020].

Un étalonnage caméra-écran est réalisé pour chaque opérateur, afin de projeter la direction du regard dans le repère écran en tenant compte de la distance, de la position de la caméra et de la taille de l'écran [Takahashi 2016]. Des cartes de chaleur peuvent alors être construites afin de matérialiser les zones de focalisation du regard au cours du processus de recalage 3D-2D.

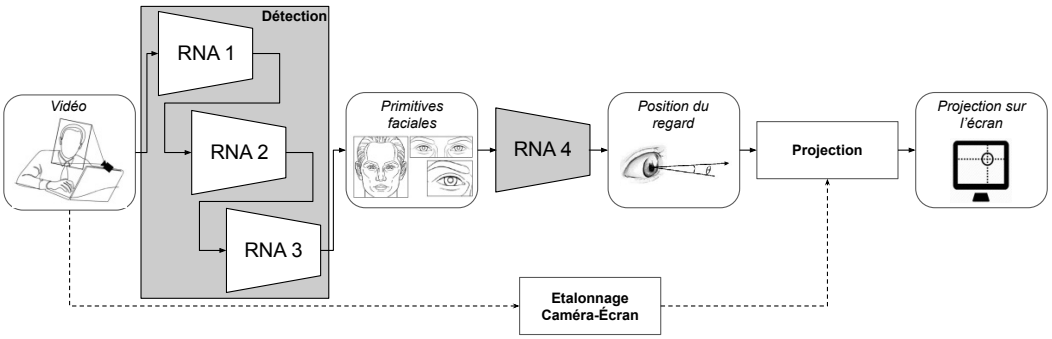


FIGURE 4.13 – Représentation schématique du pipeline technique de suivi du regard de l'opérateur. Source : généré par l'auteur à partir de [Gudi 2020].

Les premières expérimentations menées avec cet outil ont impliqué un unique opérateur chirurgien et huit cas rétrospectifs de NP. Les premières tendances confirment que la NP se caractérise par une pénurie de primitives anatomiques robustes pour résoudre le problème de recalage 3D-2D. Le cas des tumeurs hilaires fait cependant exception, car la zone péri-tumorale est très riche d'éléments anatomiques discriminants comme l'insertion des branches vasculaires dans le parenchyme au travers du hile rénal. En dehors de ce cas, soit environ 65 % du temps, aucune primitive fiable de recalage n'a pu être identifiée. Les opérateurs expérimentés s'appuient principalement sur les bords du rein, avec une prévalence pour le bord externe et les contours des tumeurs exophytiques. Le processus de recalage 3D-2D nécessite

entre trois et dix minutes afin d'être réalisé de manière cliniquement satisfaisante et dans la perspective d'initialiser un système de RA.

Au cours de ces investigations, l'impact du dispositif d'affichage sur la qualité et le temps de recalage 3D-2D peropératoire a également été exploré. Les premières observations issues de l'utilisation des différentes versions du prototype de RA lors des essais de faisabilité (cf. section 2.2.3) suggèrent qu'un affichage stéréoscopique de bonne facture pourrait améliorer la précision et la rapidité du recalage comparativement à un écran standard. Ces résultats restent à confirmer sur un nombre de cas et d'utilisateurs plus important et fait actuellement l'objet d'une étude rétrospective sur données.

4.5 Approches de recalage 3D-2D basées sur des primitives

Comme mentionné précédemment en section 4.3, les approches de recalage 3D-2D basées sur la détection et la mise en correspondance de primitives visuelles peuvent être regroupées en trois catégories : les méthodes manuelles, semi-automatiques et automatiques.

Dans la suite de cette section, plusieurs méthodes de recalage 3D-2D, qu'elles soient issues de la littérature ou développées spécifiquement dans le cadre de ce travail, sont mises en œuvre et analysées dans le contexte applicatif de la NPRA guidée par RA.

4.5.1 Méthodes de recalage 3D-2D manuelles

Les méthodes de recalage 3D-2D manuelles reposent sur l'utilisation d'un système de rendu graphique intéractif qui fait appel à des interfaces bas niveau (Application Programming Interface (API) en Anglais) comme OpenGL ou Vulkan de manière directe ou indirecte par l'intermédiaire de bibliothèques logicielles spécialisées comme VTK (Visualization Toolkit, Kitware) ou Open3D [Zhou 2018].

La génération d'une image augmentée résulte de deux étapes principales : d'une part la configuration de la scène par positionnement des divers objets (maillages, caméras et sources de lumière) et d'autre part le processus de rendu pixel par pixel dans le plan image aussi appelé viewport en Anglais.

Dans un premier temps, le maillage tridimensionnel du jumeau numérique, défini par un ensemble de sommets et de faces, est rendu dans le repère monde de la scène. Par composition des transformations modèle, vue et projection, ce modèle est projeté dans le repère image, comme illustré en figure 4.14. Les matrices modèle et vue traduisent respectivement les transformations appliquées au jumeau numérique et à la caméra, sous le contrôle de l'utilisateur via des interactions directes. La matrice de projection intègre quant à elle les paramètres intrinsèques de la caméra endoscopique réelle (cf. section 2.2.6) afin de modéliser une projection perspective cohérente. A l'issue de la rastérisation, un shader d'apparence (fragment shader

en Anglais) est appliqué afin d'attribuer à chaque pixel une couleur en fonction des propriétés visuelles souhaitées (couleur, transparence, réflexion, etc.), comme illustré en figure 4.15.

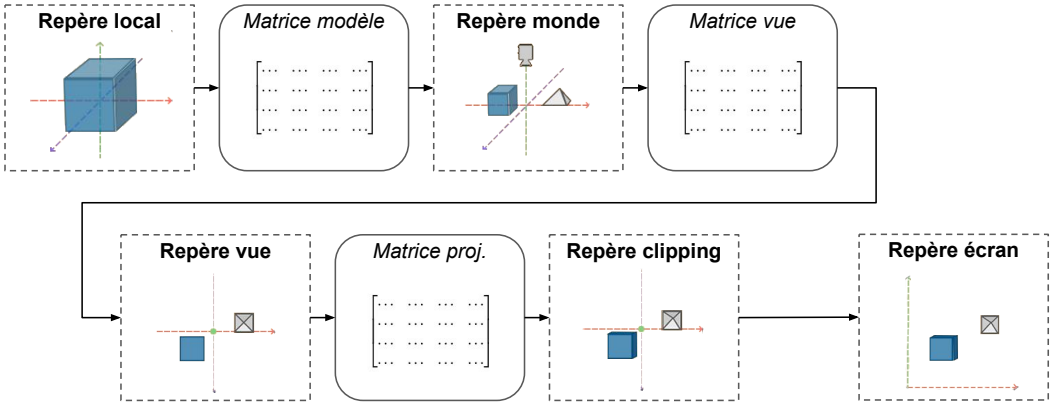


FIGURE 4.14 – Représentation schématique de la composition des matrices modèle, vue et projection qui permet de rendre une image de la scène tridimensionnelle dans le repère de l'écran. Source : générée par l'auteur à partir de [Vries 2014].

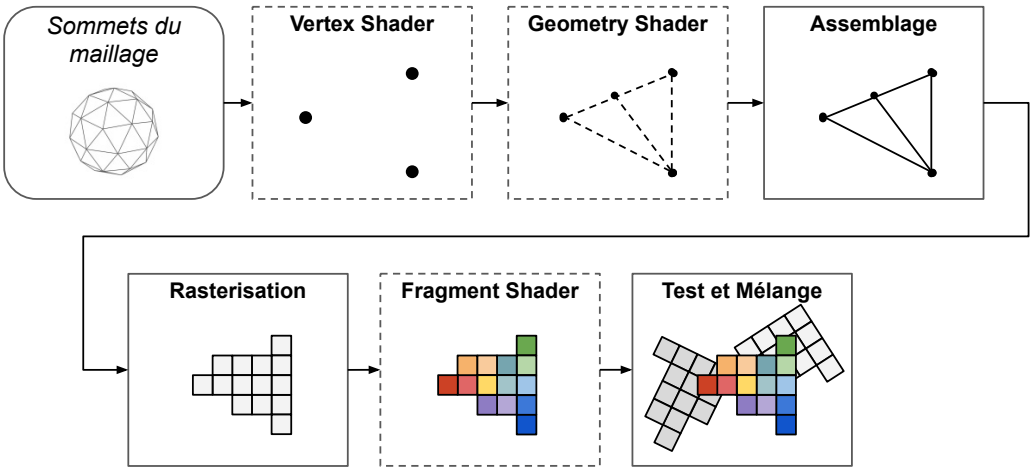


FIGURE 4.15 – Représentation schématique du pipeline de rendu graphique, qui permet à partir des sommets d'un maillage d'obtenir une projection de la géométrie dans le plan image. Pour chaque pixel résultant est attribuée une couleur en fonction des propriétés de surface de l'objet et de ses interactions avec les autres éléments de la scène. Source : générée par l'auteur à partir de [Vries 2014].

Un second rendu d'un quad texturé avec l'image provenant du flux vidéo est réalisé. La composition finale des deux rendus (arrière plan et maillage) est effectuée par mélange de couleur afin d'obtenir l'image augmentée affichée à l'utilisateur.

Le recalage 3D-2D manuel repose ainsi sur la capacité de l'utilisateur à interagir avec une interface graphique capable de retranscrire virtuellement les interactions physiques numérisées de l'opérateur via un périphérique informatique (souris, écran tactile, souris 3D, etc.) afin de manipuler la scène tridimensionnelle.

Les méthodes de recalage 3D-2D manuelles se distinguent en deux sous-catégories en fonction du nombre de paramètres laissés libre à l'utilisateur. Les méthodes dites de recalage rigide nécessitent un ajustement manuel des trois rotations, des trois translations et de l'échelle du jumeau numérique affiché sur la vue peropératoire. Les nouvelles valeurs sont alors assignées par simplicité à la matrice modèle. Ce processus, bien que relativement simple conceptuellement, demande une bonne compréhension de la scène chirurgicale et un budget temporel de trois à dix minutes en fonction de la qualité de la vue peropératoire. Plus la vue est large, plus le recalage est aisé, grâce à une plus grande richesse d'indices visuels.

A l'opposé, les méthodes manuelles dites non-rigides ou déformables, introduisent en plus des paramètres de transformation rigide, un ensemble de paramètres de déformation linéaire ou non-linéaire du jumeau numérique. En théorie, chaque sommet du maillage pourrait être déplacé indépendamment pour s'aligner parfaitement avec l'image peropératoire. Dans la pratique, cette approche est irréaliste en raison de sa complexité et du temps nécessaire. Un compromis est ainsi adopté, limitant les interactions a un sous-ensemble de paramètres directement contrôlable par l'utilisateur.

Une première approche exploitée repose sur la déformation libre (Free Form Deformation (FFD) en Anglais) [Sederberg 1986], illustrée en figure 4.16. Le maillage tridimensionnel du jumeau numérique est englobé par une grille de contrôle qui prend généralement la forme d'un parallélépipède rectangle ou d'une représentation sous-échantillonnée du maillage initial. Chaque sommet de la grille est appelé point de contrôle et est manipulable manuellement en translation dans les trois directions de l'espace. Les coordonnées des sommets du maillage sont alors recalculées à l'aide de polynômes de Bernstein.

Un polynôme de Bernstein de degré d dans une base canonique est défini comme suit :

$$f(x) = \sum_{i=0}^{d} c_i x^i \tag{4.1}$$

Avec c_i les coefficients.

Dans la base de Bernstein, l'expression est la suivante :

$$f(x) = \sum_{i=0}^{d} b_i B_i^d(x) \tag{4.2}$$

Avec :

$$B_i^d(x) = \binom{d}{i} x^i (1-x)^{d-i} \tag{4.3}$$

Les polynômes de Bernstein sont définis dans l'intervalle $x \in [1, 0]$ et forment une combinaison linéaire convexe des points $(i/d, bi)$. Ainsi, les points (x, y) se trouvent à l'intérieur de la coque convexe formée par les points de contrôle $(i/d, bi)$.

La méthode FFD pour une grille de contrôle parallélépipédique, exploite les propriétés énoncées précédemment dans l'espace tridimensionnel et consiste en trois étapes différentes :

1. Pour tout sommet du maillage x dans cette grille de hauteur H, de largeur W et de longueur L, ses coordonnées (s, t, u) sont :

$$\overrightarrow{x} = \overrightarrow{x_0} + s\overrightarrow{H} + t\overrightarrow{W} + u\overrightarrow{L} \ni 0 < s < 1, 0 < t < 1, 0 < u < 1 \tag{4.4}$$

   Ou $\overrightarrow{x_0}$ sont les coordonnées de l'origine de la cage.

2. La grille de points de contrôle $\overrightarrow{p_{ijk}}$ forme un treillis. Ceux-ci forment $l + 1 <$ plans dans la direction $\overrightarrow{H}$, $m + 1$ plans dans la direction $\overrightarrow{W}$ et $n + 1$ plans dans la direction $\overrightarrow{L}$.

$$\overrightarrow{p_{ijk}} = \overrightarrow{x_0} + \frac{i}{l}\overrightarrow{H} + \frac{i}{m}\overrightarrow{W} + \frac{k}{n}\overrightarrow{L} \ni i \in \{0, ..., l\}, j \in \{0, ..., m\}, k \in \{0, ..., n\} \tag{4.5}$$

3. Les déformations sont spécifiées en déplaçant les points de contrôle aux positions $\overrightarrow{p_{ijk}}'$. La position déformée $\widehat{\overrightarrow{x}}$ d'un sommet du maillage est trouvée en évaluant le polynôme de Bernstein trivarié. Les points de contrôle $\overrightarrow{p_{ijk}}'$ sont en fait les coefficients du polynôme de Bernstein :

$$\widehat{\overrightarrow{x}} = \sum_{i=0}^{l} \binom{l}{i} s^i (1-s)^{l-i} \left(\sum_{j=0}^{m} \binom{m}{j} t^j (1-t)^{m-j} \left(\sum_{k=0}^{n} \binom{n}{k} u^k (1-u)^{n-k} \overrightarrow{p_{ijk}}' \right) \right) \tag{4.6}$$

Malgré son potentiel, l'utilisation de la FFD en peropératoire, évaluée au cours des essais de faisabilité (cf. section 2.2.3) s'est révélée peu concluante. D'une part, l'amplitude des déplacements à appliquer aux points de contrôle pour arriver au résultat mentalisé est difficile à estimer pour l'utilisateur. D'autre part, les vues endoscopiques partielles limitent la visibilité de la grille de contrôle, complexifiant la tâche.

Pour pallier ces limitations, des modes de déformation inspirés de l'anatomie sont explorés, ils facilitent la compréhension et l'interprétation des manipulations par l'opérateur, en s'appuyant sur les axes anatomiques définis dans le modèle canonique du rein (cf. section 4.6.4). Des déformations dirigées en torsion, flexion, traction et compression peuvent alors être implémentées selon les équations de Barr [Barr 1984], comme illustré en figure 4.17. Cette approche permet de simuler des transformations complexes avec une faible charge computationnelle. Cette méthode semble plus convaincante, il reste cependant à effectuer des essais en conditions réelles.

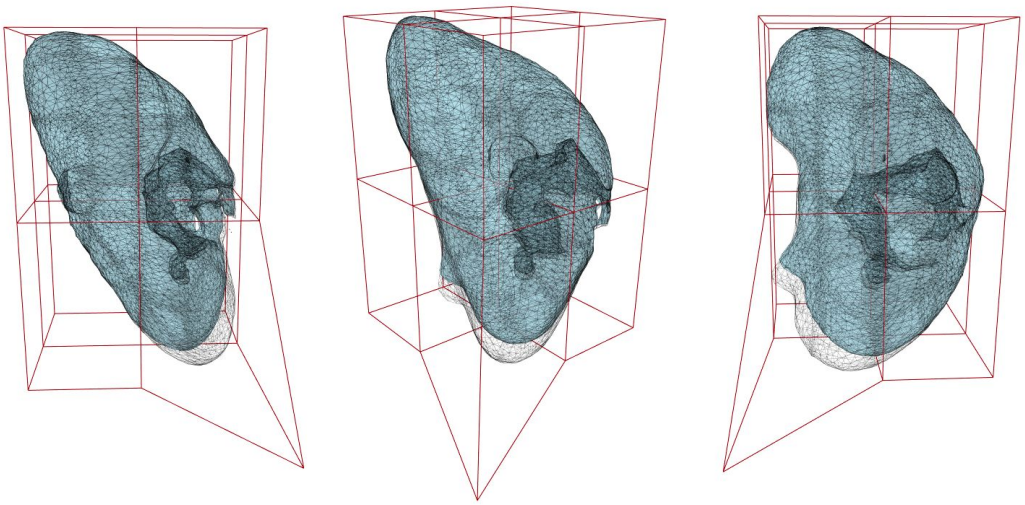


FIGURE 4.16 – Illustration sous trois angles de vue d'une déformation libre (FFD - Free Form Deformation en Anglais) d'un rein avec une grille de contrôle parallélépipédique $3 \times 3 \times 3$ en rouge. Le jumeau numérique non déformé du rein est représenté en bleu, tandis que la version déformée apparaît en gris. Source : auteur.

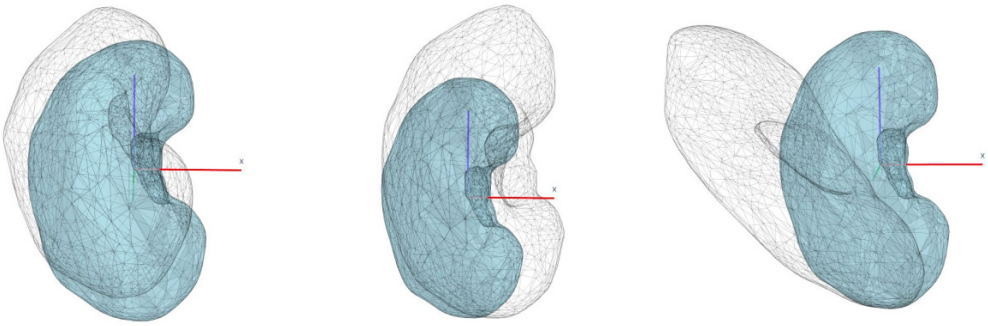


FIGURE 4.17 – Illustration de déformations anatomiques via un modèle élastique mobilisable en torsion, flexion, traction et compression régi par les équations de Barr. Le jumeau numérique non déformé du rein est représenté en bleu, tandis que la version déformée apparaît en gris. Source : auteur.

Enfin, quelle que soit la méthode de recalage 3D-2D manuelle considérée, la qualité de l'alignement final ainsi que le temps nécessaire à son obtention dépendent étroitement des capacités de l'utilisateur, en particulier sa compréhension clinique de l'image peropératoire et sa capacité à visualiser les objets en trois dimensions. Dès lors, il apparaît peu réaliste d'envisager une généralisation de cette méthode comme solution unique en pratique clinique. Néanmoins, le recalage 3D-2D manuel constitue une solution de secours en cas de défaillance ou d'indisponibilité d'un système de recalage automatique et participe ainsi à la maîtrise des risques résiduels (ISO 14971 :2019).

4.5.2 Méthodes de recalage 3D-2D semi-automatiques

Les méthodes de recalage 3D-2D semi-automatiques constituent en NP une solution rapide pour initier un premier recalage grossier entre le jumeau numérique et l'image peropératoire. Ce recalage initial est fréquemment suivi d'un ajustement manuel, destiné à affiner la localisation du jumeau numérique. Ces méthodes nécessitent des données d'entrée généralement fournies par l'utilisateur.

Pour le cas applicatif de la NP, des données d'initialisation comme des correspondances 3D-2D entre le jumeau numérique et la vue caméra peuvent être utilisées pour estimer la pose de la caméra par PnP. Toutefois, seuls certains cas se prêtent à l'utilisation de cette méthode, comme les tumeurs hilaires où des points saillants facilement repérables sont présents (points d'insertions ou bifurcations vasculaires).

Il est également possible de pré-positionner grossièrement le jumeau numérique en traçant au moins deux axes anatomiques principaux (cf. modèle canonique du rein présenté en section 4.6.4) directement sur le plan image. Un solveur détermine alors la transformation homogène qui aligne les segments tracés dans l'image avec le repère canonique : l'origine ainsi que les axes X et Y du repère estimé sont contraints à coïncider avec les plans de rétroprojection π_x et π_y correspondant aux segments tracées. Cette méthode est illustrée en figure 4.18. L'utilisateur doit cependant disposer d'une vue suffisamment globale de l'organe afin de réaliser les tracés.

Les méthodes de recalage 3D-2D semi-automatiques permettent de réduire le nombre d'interactions manuelles avec le jumeau numérique. Cependant elles nécessitent des primitives de qualité parfaitement appareillées et restent contraintes à des cas spécifiques d'utilisation. Par conséquent, elles ne peuvent donc à elles seules résoudre la problématique de recalage en NP.

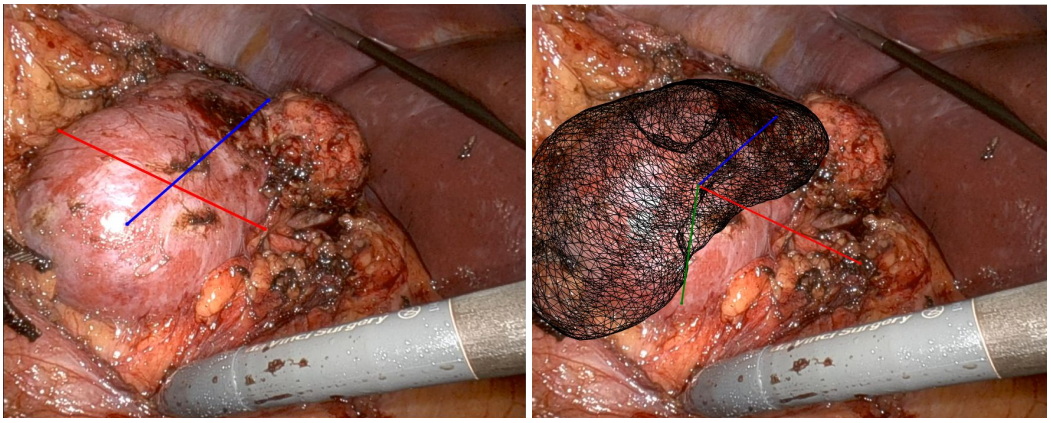


FIGURE 4.18 – Méthode de recalage 3D-2D semi-automatique utilisant l'alignement entre un repère tracé sur l'image peropératoire et le repère canonique du rein. Source : auteur.

4.5.3 Méthodes de recalage 3D-2D automatiques

Les méthodes de recalage 3D-2D automatiques basées sur l'exploitation de primitives sont difficilement implémentables en NP par manque de repères robustes et systématiquement observables au cours de la chirurgie. Ce constat est détaillé en section 4.4.

La surface du parenchyme rénal semble être la seule primitive potentiellement exploitable pour une méthode de recalage automatique. Cette structure est généralement visible pendant les phases chirurgicales ou la RA présente un intérêt clinique. Plusieurs techniques de vision par ordinateur peuvent être utilisées via trois étapes successives.

Une première étape consiste à segmenter cette surface dans les images peropératoires issues d'un flux vidéo étalonné à l'aide d'un réseau de neurones. Ici, nous utilisons un réseau de type UNet spécifiquement entraîné à cette tâche, tel que présenté en section 5.4, qui à partir d'une unique image d'entrée couleur infère un masque binaire de localisation du parenchyme rénal.

Dans une seconde étape, une reconstruction tridimensionnelle dense de la surface rénale est réalisée à partir d'images peropératoires. Un nouveau modèle tridimensionnel est alors généré et servira de proxy pour le recalage. Plusieurs techniques de vision par ordinateur passives peuvent être utilisées à cet effet : des approches géométriques comme le SfM, le SLAM [Tosi 2024, Zhou 2019a], le Gaussian Splatting [Xu 2024b, Xie 2024a, Liu 2024a, Zheng 2025] ou la stéréovision via SGBM [Hirschmuller 2005, Hirschmuller 2007]. A cela s'ajoute des approches par apprentissage profond comme Dense and Unconstrained Stereo 3D Reconstruction (DUSt3R) en Anglais [Wang 2024b], Matching And Stereo 3D Reconstruction (MASt3R) en Anglais [Leroy 2024, Duisterhof 2024, Murai 2025] ou Visual Geometry Grounded Transformer (VGGT) en Anglais [Wang 2025] en imagerie monocu-

laire ou FoundationStereo [Wen 2025] en imagerie stéréoscopique. L'exploitation du flux stéréoscopique, souvent disponible en chirurgie robotique (cf. section 1.2.4.3), impose toutefois une parfaite synchronisation temporelle entre les vues gauche et droite, comme discuté dans le chapitre 3.

Les techniques de reconstruction actives qui utilisent des dispositifs tiers comme une source de lumière structurée, un capteur de temps de vol (Time Of Flight (TOF) en Anglais) ou un balayage de faisceau laser [Richter 2024] sont exclues car non disponibles en routine clinique.

Quelques exemples de modèles tridimensionnels peropératoires en chirurgie rénale sont donnés en figure 4.19. La surface rénale peut alors être extraite de la reconstruction via l'utilisation des données de segmentation. Pour les approches basées sur l'apprentissage, la sélection de la zone d'intérêt doit impérativement être réalisée après l'inférence afin de ne pas compromettre significativement les résultats.

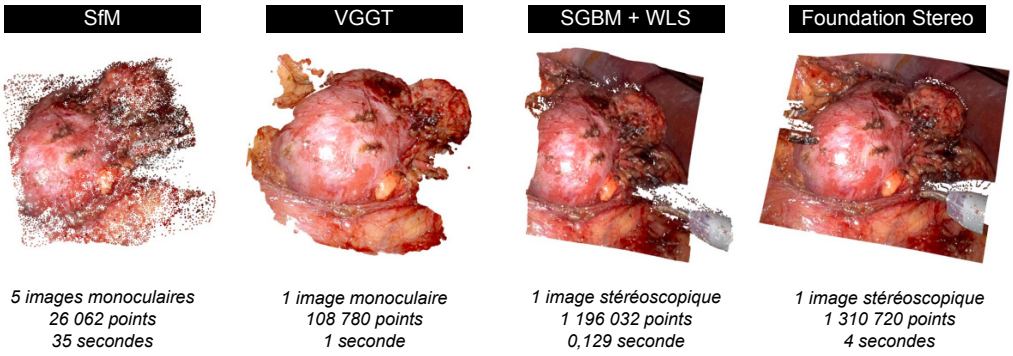


FIGURE 4.19 – Exemples de reconstructions tridimensionnelles en chirurgie rénale obtenues à l'aide de différentes techniques de vision par ordinateur, réalisées sur un PC Ubuntu équipé d'un processeur Intel Core™ i7-13620H et d'une carte graphique Nvidia RTX 4070M. (de gauche à droite) Structure-from-Motion (SfM) à partir de cinq images monoculaires ; Inférence du réseau Visual Geometry Grounded Transformer (VGGT) à partir d'une image monoculaire ; Semi-Global Block Matching (SGBM) à partir d'une image stéréoscopique ; Inférence du réseau Foundation Stereo à partir d'une image stéréoscopique. Source : auteur.

Dans le contexte spécifique de la chirurgie robotique et plus particulièrement de la NPRA, les méthodes de reconstruction tridimensionnelle qui nécessitent uniquement une image s'avèrent plus adaptées que celles qui requièrent une séquence d'images, comme le SfM. En effet, les contraintes peropératoires comme la limitation des mouvements de la caméra endoscopique dans le trocart, la proximité entre l'optique et l'organe ainsi que la faible possibilité de mobilisation de l'organe sans introduire de déformation ou d'occlusion importantes avec l'instrumentation, rendent difficile l'acquisition de séquences exploitables pour les approches multi-vues.

Enfin, une troisième étape vise à réaliser un recalage 3D-3D entre le modèle peropératoire reconstruit et le jumeau numérique. Des algorithmes de recalage 3D-3D rigides comme l'ICP ou déformables comme le CPD ou le NR-ICP peuvent être utilisés à ces fins. Le recalage est souvent réalisé suivant une approche hiérarchique coarse-to-fine [Zhang 2019b], ce qui permet d'améliorer la convergence et la précision du processus global d'alignement. Quelques exemples de recalages 3D-3D rigides en chirurgie rénale sont donnés en figure 4.20, réalisés à partir des reconstructions tridimensionnelles peropératoires illustrées en figure 4.19.

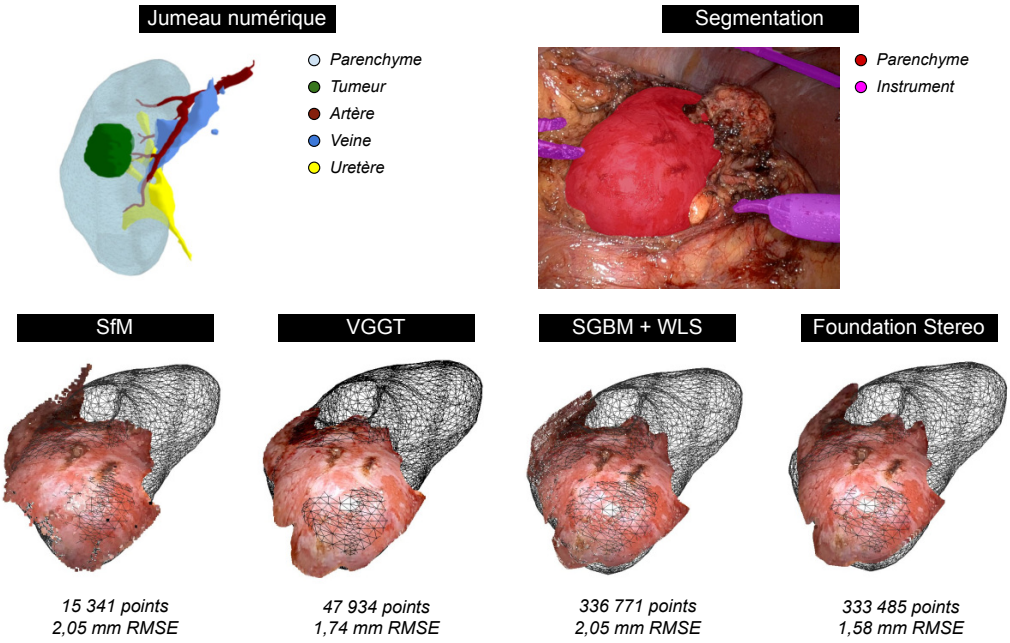


FIGURE 4.20 – Exemples de recalages 3D-3D rigides via Iterative Closest Point (ICP) en chirurgie rénale entre le jumeau numérique et la surface rénale segmentée, réalisés sur un PC Ubuntu équipé d'un processeur Intel Core™ i7-13620H et d'une carte graphique Nvidia RTX 4070M. (de gauche à droite sur la deuxième ligne) Modèles peropératoires obtenus par Structure-from-Motion (SfM), Visual Geometry Grounded Transformer (VGGT), Semi-Global Block Matching (SGBM) et Foundation Stereo. Source : auteur.

Il est important de noter que le recalage 3D-3D sans initialisation préalable présente un risque élevé d'échec, notamment avec des méthodes comme l'ICP. L'étendue de la surface rénale reconstruite tridimensionnellement et la richesse de sa topologie jouent un rôle déterminant pour contraindre suffisamment le problème de recalage et assurer la convergence vers une solution fiable.

4.6 Approche de recalage 3D-2D basée sur l'apprentissage automatique

L'ensemble des limitations identifiées au sujet des méthodes de recalage 3D-2D existantes applicables à la NP, ont guidé notre choix vers une méthodologie fondée sur l'apprentissage automatique. Celle-ci vise à extraire des représentations visuelles robustes et adaptatives selon la phase chirurgicale.

L'analyse méthodologique préliminaire présentée en section 4.4 met en évidence que l'efficacité de la RA en NP repose sur la capacité à réaliser un recalage fiable du jumeau numérique sur la vue peropératoire. Ce recalage doit être robuste, multi-échelle et sensible à la phase chirurgicale.

4.6.1 Généralités

Suivant la tendance dominante en matière de RA, nous supposons que l'information virtuelle est contenue dans le Jumeau du Patient (JP), qui est défini comme un jumeau numérique spécifique. Il représente l'anatomie unique d'un patient, conçu pour des applications médicales ou chirurgicales personnalisées. Nous utilisons par la suite le terme de JP pour désigner le modèle tridimensionnel préopératoire reconstruit à partir de la TDM ou de l'IRM (cf. sections 1.3.1.1 et 1.3.1.2). Le JP inclut les maillages tridimensionnels surfaciques représentant la géométrie parenchymateuse du rein et les structures internes d'intérêt pour la RA. Nous supposons que le JP est disponible avant l'intervention chirurgicale, ce qui est déjà courant dans les cas complexes de NP.

Nous introduisons également la notion de Prototype de Forme (PF) pour désigner un modèle tridimensionnel générique de la forme du parenchyme rénal. L'utilisation du PF permet de généraliser l'approche proposée à un ensemble de patients.

La méthode de recalage 3D-2D proposée fonctionne en deux étapes. Dans un premier temps, le JP est recalé avec précision sur le PF lors de la phase de préparation des données, réalisée en amont de l'intervention chirurgicale. En peropératoire, la seconde étape consiste à estimer la pose du PF sur l'image endoscopique. La combinaison des transformations issues de ces deux étapes permet d'obtenir la transformation 3D-2D globale du JP vers la vue peropératoire, nécessaire à l'affichage de la RA.

L'étape peropératoire repose sur un RNA end-to-end qui établit une correspondance entre l'image et la pose du PF recalé. Ce modèle est entraîné de manière supervisée sur des données réelles labellisées.

Pour cette première version de notre système d'estimation automatique de la pose du rein par apprentissage profond, nous avons ciblé trois des cinq phases clés de la NP (cf. section 2.1.4), à savoir l'exposition du rein, la dissection vasculaire et l'identification tumorale. Ces phases précèdent l'incision parenchymateuse et constituent les moments où la RA à le plus de potentiel pour le guidage peropératoire.

4.6.2 Collecte des données

L'un des principaux défis de l'apprentissage supervisé en contexte chirurgical réside dans la collecte et la labellisation rigoureuse des données nécessaires à l'entraînement et l'évaluation des réseaux.

Dans ce travail, un ensemble complet de données pour l'estimation de la pose rénale est constitué. Il s'appuie sur sept vidéos peropératoires haute résolution (1280 × 1024 pixels, 60 Hz), enregistrées lors de NPRA réalisées à l'aide du système de télémanipulation chirurgicale Da Vinci Xi (Intuitive Surgical), dans le cadre des sessions d'essais de faisabilité décrites en section 2.2.3. Les séquences vidéo spécifiques destinées à l'étalonnage géométrique (cf. section 2.2.6) de la caméra endoscopique à partir d'une mire ont également été collectées.

Les jumeaux numériques correspondants ont été reconstruits en amont de chaque intervention, par segmentation de l'examen d'imagerie tomodensitométrique avec injection de produit de contraste. Cette étape a été assurée par l'équipe clinique impliquée dans l'étude, comprenant un chirurgien principal avec plus de huit années d'expertise dans le domaine, assisté par un chirurgien avec au moins deux années d'expérience en NPRA. Chaque segmentation inclut au minimum cinq structures anatomiques d'intérêt pour la RA : le parenchyme rénal, la veine rénale, l'artère rénale, les voies urinaires et la ou les tumeurs.

Les caractéristiques cliniques des sept patients inclus sont présentées dans le tableau 4.2, tandis qu'une visualisation des JP correspondants est fournie en figure 4.21. Cet ensemble de données a été intégralement collecté, labellisé et exploité pour l'entraînement ainsi que l'évaluation du RNA d'estimation automatique de la pose du rein.

ID Patient	Côté de la NP	Type de tumeur	Localisation tumeur	Taille de tumeur (mm)	Score RENAL
1	Gauche	Exo. < 50 %	Sup.	15	05a
2	Gauche	Exo. < 50 %	Equa.	38	09ph
3	Droite	Exo. > 50 %	Equa.	25	10ah
4	Droite	Endo.	Inf.	35	08a
5	Gauche	Exo. > 50 %	Sup.	42	07a
6	Gauche	Exo. < 50 %	Inf.	55	09p
7	Gauche	Endo.	Inf.	27	08a

TABLE 4.2 – Caractéristiques cliniques des sept patients opérés pour Néphrectomie Partielle (NP). En cas de tumeur multifocale, seules les caractéristiques de la lésion de plus grand volume sont prises en compte. Source : auteur.

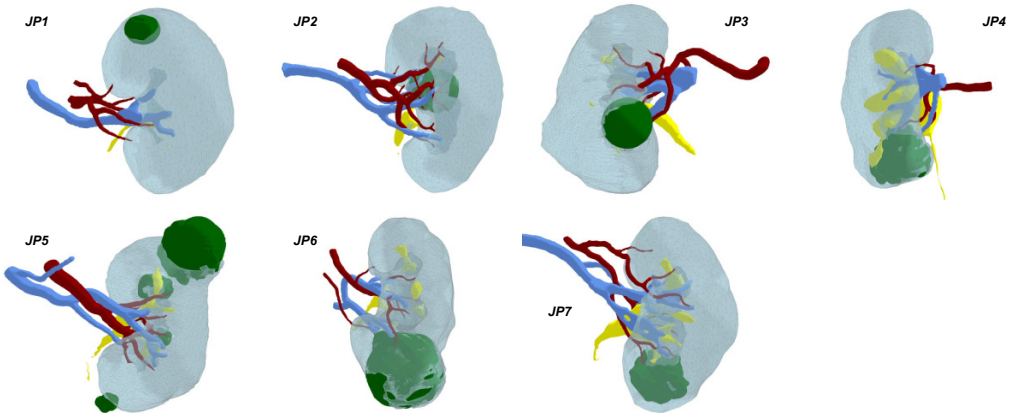


FIGURE 4.21 – Visualisation des sept Jumeaux numériques Patients (JP) associés aux cas cliniques collectés, labellisés et utilisés pour l'entraînement et l'évaluation du réseau de neurones d'estimation automatique de la pose du rein. Source : auteur.

4.6.3 Labellisation des données

La labellisation des données peropératoires a été réalisée de manière rétrospective par une équipe de chirurgiens urologues d'un centre expert à haut volume (CHU de Bordeaux, France), ci-après désignée comme "experts en NPRA". Ces praticiens sont familiers des outils de guidage peropératoire comme l'EP (cf. section 1.3.2.1) et la RV (cf. section 1.3.2.3) dans le cadre de la NP guidée par l'image.

La procédure de labellisation a été conduite par un chirurgien disposant de deux années d'expérience en NPRA, sous la supervision d'un chirurgien senior comptant plus de huit années d'expérience dans le domaine. Une application hors ligne dédiée, illustrée en figure 4.22, a été utilisée pour effectuer le recalage 3D-2D manuel du JP sur des Images Clés (IC), sélectionnées selon leur pertinence clinique. Une IC correspond à une image issue du flux vidéo peropératoire qui répond aux critères suivants : l'organe cible est visible partiellement ou totalement, l'image est nette et le champ de vue est dégagé, peu d'occlusions entre l'organe cible et la caméra sont présentes.

Le processus de labellisation n'était pas limité dans le temps, ce qui a permis à l'annotateur une navigation libre au sein des séquences vidéo. Cette souplesse a favorisé une sélection rigoureuse des IC, en s'appuyant sur l'examen des images adjacentes afin de garantir la clarté anatomique et la cohérence temporelle des données labellisées. L'objectif premier était de constituer un ensemble de données de qualité pour l'entraînement et l'évaluation de la méthode proposée.

En complément du recalage du JP sur l'image, deux méta-informations ont été annotées pour chaque élément du jeu de données : le niveau de confiance estimé par l'expert dans le recalage ainsi que la phase chirurgicale correspondante à la vue peropératoire.

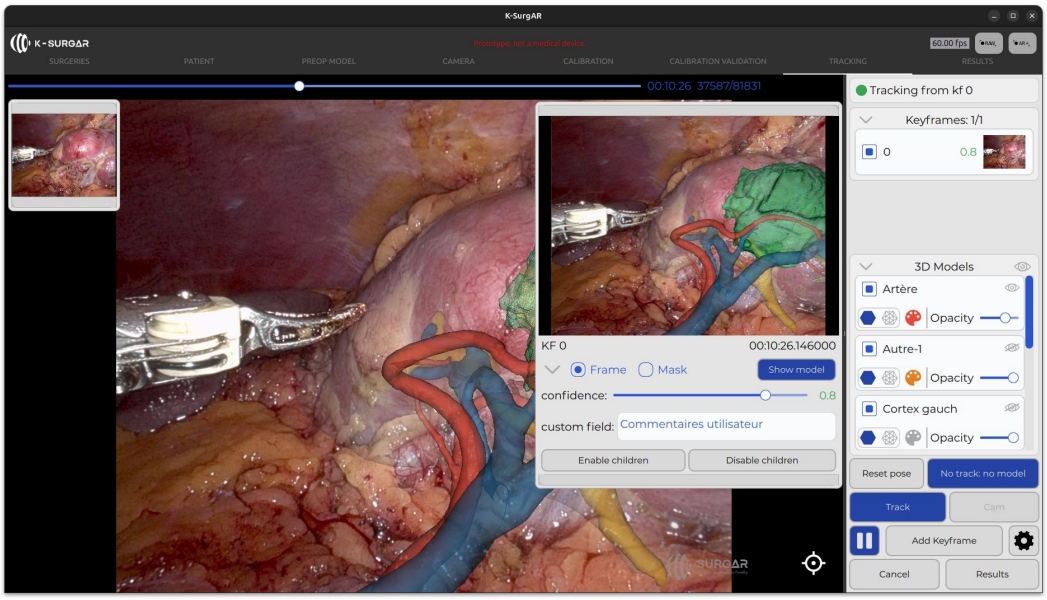


FIGURE 4.22 – Application hors ligne dédiée à la labellisation des données pour l'entraînement et l'évaluation du réseau de neurones d'estimation automatique de la pose du rein. Des recalages 3D-2D manuels sur Iimages Clés (IC) sélectionnées dans le flux vidéo peuvent être réalisés et commentés via l'ajout de méta-informations. Source : auteur.

Bien que la labellisation peropératoire en temps réel n'ait pas été possible, le caractère rétrospectif de la labellisation a permis à l'équipe chirurgicale ayant réalisé l'intervention, de s'appuyer sur sa connaissance précise de l'anatomie spécifique du patient et du déroulement opératoire.

Une fois le recalage 3D-2D manuel effectué pour chaque IC, un algorithme de suivi rigide a été utilisé pour interpoler spatialement et temporellement la position du JP sur l'ensemble de la séquence vidéo. Cet algorithme s'appuie sur la mise en correspondance automatique de caractéristiques (features en Anglais), selon une méthode proche de celle proposée par [Collins 2020, Chandelon 2022] (cf. chapitre 5). Cette dernière permet de générer des séquences vidéo augmentées, associant une pose tridimensionnelle du JP à chaque image.

Cette méthodologie est issue des développements et observations empiriques portant sur plus de trente cas réalisés lors des essais de faisabilité (cf. section 2.2.3). L'utilisation répétée du système de RA (cf. section 2.2.2) par les équipes cliniques a permis de mettre en évidence qu'un recalage 3D-2D rigide localisé, appliqué à une portion de la surface rénale, répond aux exigences cliniques en matière de localisation des tumeurs au cours de la NPRA. Bien que la prise en compte des déformations puisse améliorer la visualisation des éléments extra-rénaux comme les vaisseaux, elle introduit actuellement de fortes contraintes techniques notamment en termes de temps et de complexité de recalage.

À partir de 228 IC labellisées manuellement, ce qui représente environ 40 heures de travail expert, il a été possible de générer automatiquement sept séquences augmentées, totalisant 299 659 images. Cette stratégie a significativement amélioré l'efficacité de la labellisation en s'affranchissant de l'annotation image-par-image. Les poses du jumeau numérique ainsi obtenues constituent une semi-vérité terrain, en l'absence d'une vérité absolue mesurable. Ces données seront utilisées pour l'entraînement et l'évaluation du réseau de neurones d'estimation de pose.

Un contrôle qualité semi-automatique a été mis en place afin de valider la cohérence spatio-temporelle des poses générées. Le système de contrôle est capable d'émettre des alertes en cas de variations anormales de position du JP entre images successives ou de dépassement des limites du volume de travail défini. Cette validation a permis de renforcer la fiabilité de l'ensemble de données et de minimiser les erreurs de suivi automatique.

4.6.4 Préparation des données

Les données issues de phase labellisation sont ensuite soumises à un processus de préparation destiné à les rendre compatibles pour l'entraînement du réseau de neurones proposé. Cette étape repose sur deux opérations fondamentales : 1) les JP sont recalés avec le PF et 2) les paramètres intrinsèques des caméras endoscopiques sont adaptés à des paramètres génériques. Ces opérations ont pour objectif de rendre la méthode proposée utilisable pour tous les patients sans réentraînement.

Le PF utilisé à l'étape 1), est un modèle anatomique générique en forme de haricot avec une bordure latérale convexe et une bordure médiane concave. Il mesure onze centimètres de long, six centimètres de large et trois centimètres d'épaisseur. Bien qu'il ait été modélisé pour le rein droit, il convient également pour le rein gauche, en raison de la similarité morphologique des deux organes, les principales différences résident dans leur positionnement anatomique (cf. section 2.1.1.1). Le maillage du PF a été extrait de l'atlas numérique tridimensionnel Z-Anatomy, illustré en figure 4.23.

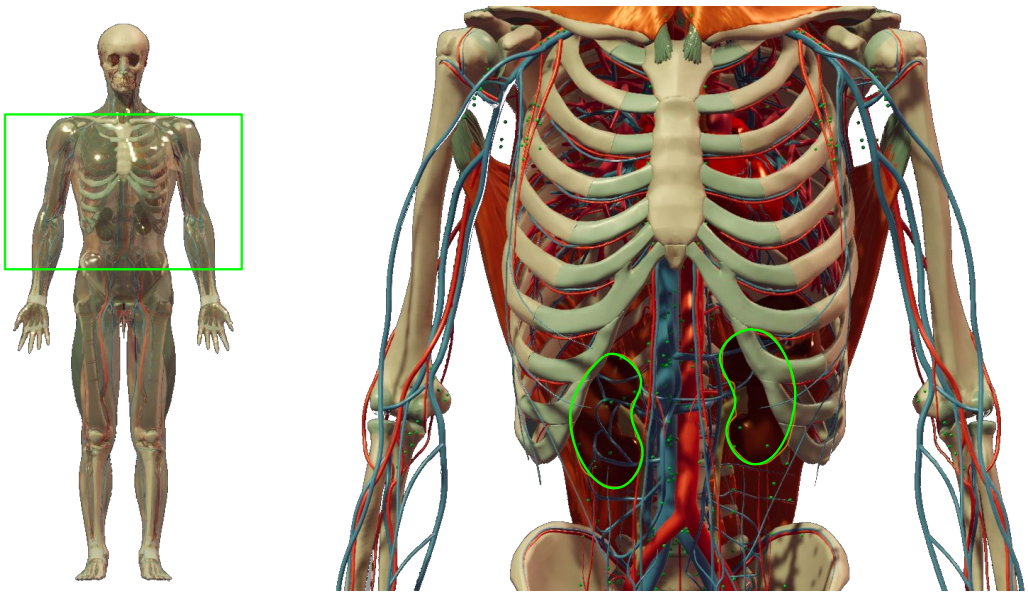


FIGURE 4.23 – Visualisation sous Blender de l'atlas numérique tridimensionnel Z-Anatomy. Source : auteur.

Le recalage préopératoire est réalisé afin d'estimer la transformation T_{jp2pf} entre le parenchyme du JP et le PF. Ce processus se décline en trois actions successives, illustrées en figure 4.24. La première action consiste à initialiser automatiquement un repère canonique fixé au centre de masse du JP. La direction de l'axe Z est orientée vers le pôle supérieur en fonction des vecteurs propres du tenseur d'inertie, tandis que l'axe X est aligné sur le hile rénal. La détection automatique du hile rénal repose sur l'analyse des sommets concaves du maillage à proximité du centre de masse, en tenant compte de la courbure locale, de la direction des normales et de la position relative des points candidats. Le pseudo-code de détection automatique du repère canonique du rein est donné dans l'algorithme 4.

Algorithme 4 Détection automatique du repère canonique d'un rein

Précondition : Maillage du rein M_{kidney}, maillage vasculaire M_{vessel}

Postcondition : Repère canonique $\{X_c,\, Y_c,\, Z_c\}$, drapeau *Right / Left*

$C_k \leftarrow \text{centre_masse}(M_{\text{kidney}})$

$C_v \leftarrow \text{centre_masse}(M_{\text{vessel}})$

$N \leftarrow \text{normales_sommets}(M_{\text{kidney}})$

$K \leftarrow \text{courbure_gaussienne}(M_{\text{kidney}})$

$V_{\text{conc}} \leftarrow \{\, p \in M_{\text{kidney}} \mid K(p) < P_{50}(K) \,\}$

$V_{\text{prox}} \leftarrow \{\, p \in V_{\text{conc}} \mid \|p - C_k\| < P_{40}(\|V_{\text{conc}} - C_k\|) \,\}$

$\vec{a} \leftarrow C_v - C_k$ %*normale « antérieure »*

$V_{\text{ant}} \leftarrow \text{garder_côté_antérieur}(V_{\text{prox}}, \vec{a})$

$V_{\text{norm}} \leftarrow \text{garder_normales_orientées}(V_{\text{ant}}, N, \vec{a}, 45^\circ)$

$H \leftarrow \text{barycentre}(V_{\text{norm}})$ %*position estimée du hile*

$I \leftarrow \text{tenseur_inertie}(M_{\text{kidney}})$

$(\lambda_1 \geq \lambda_2 \geq \lambda_3,\ E = [e_1, e_2, e_3]) \leftarrow \text{eig}(I)$

$Z_i \leftarrow e_3$ %*plus petit λ ; orienté vers le $+Z$ global*

$\vec{h} \leftarrow \dfrac{H - C_k}{\|H - C_k\|}$ %*direction centre $\rightarrow$ hile*

$X_i \leftarrow \text{proj}_{\perp Z_i}(\vec{h})\,;\ X_i \leftarrow \dfrac{X_i}{\|X_i\|}$

$Y_i \leftarrow Z_i \times X_i$

si $X_i^x > 0$ **alors**

  RightKidney $\leftarrow$ true

sinon

  $X_i \leftarrow -X_i,\ Y_i \leftarrow -Y_i$ %*assure un repère droit*

  RightKidney $\leftarrow$ false

retour $\{X_c \leftarrow X_i,\ Y_c \leftarrow Y_i,\ Z_c \leftarrow Z_i\}$, **RightKidney**

Dans une seconde phase, un recalage 3D-3D grossier est opéré par superposition des centres de masse et des axes des repères. Cette transformation initiale est ensuite affinée à l'aide d'un algorithme itératif de recalage de type ICP appliqué en avant et en arrière, suivi d'une passe supplémentaire de recalage non-rigide NR-ICP pour optimiser la correspondance de surface. Cette chaîne de traitement assure une cohérence géométrique entre les modèles JP et PF, comme illustrée sur une série de trente cas cliniques en annexe M.

En pratique, cette procédure préopératoire est entièrement automatisée et son impact temporel reste négligeable, avec un temps d'exécution total moyen de $0,335 \pm 0,623$ seconde. En peropératoire, la méthode n'introduit qu'un simple changement de repère pour revenir au référentiel du JP, dont le temps de calcul moyen sur CPU est de $3,16 \pm 5,79$ millisecondes et devient négligeable sur GPU.

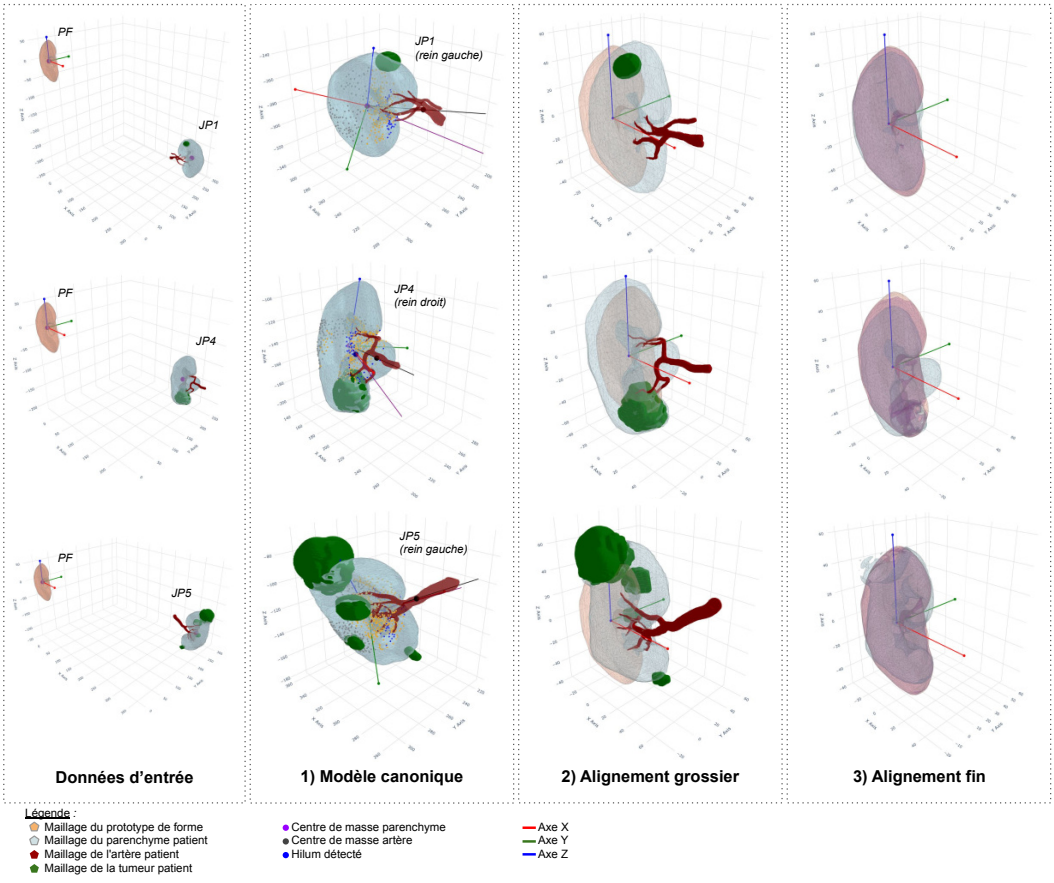


FIGURE 4.24 – Recalage 3D-3D du Jumeau Patient (JP) sur le Prototype de Forme (PF) en trois étapes : 1) initialisation du modèle canonique, 2) alignement grossier via le repère canonique et 3) réglage fin avec Iterative Closest Point (ICP). Source : auteur.

En parallèle, les paramètres intrinsèques propres à chaque caméras endoscopiques K_{spe}, sont systématiquement remplacés par des paramètres génériques K_{gen}, pour garantir l'applicabilité de la méthode à de nouveaux patients, sans nécessité de réentraînement. Les paramètres génériques choisis sont $fx = fy = 1050$, $cx = 640$ et $cy = 512$ pixels. L'adaptation des intrinsèques est obtenue par une transformation affine $A = K_{\text{gen}} K_{\text{spe}}^{-1}$, appliquée aux images labellisées, selon la méthode dite du K inverse [Fuentes-Jimenez 2022]. Elle permet d'aligner les espaces projectifs entre la configuration expérimentale et celle apprise implicitement par le réseau. Cette étape est essentielle pour garantir la robustesse du réseau vis-à-vis des variations du couple caméra-optique entre les interventions.

4.6.5 Entraînement du réseau de neurones

La méthode proposée de recalage 3D-2D automatique n'utilise pas de primitive explicite et est spécifiquement conçue pour la NP. L'approche est basée sur l'apprentissage supervisé de données expertes, comme le montre la figure 4.25.

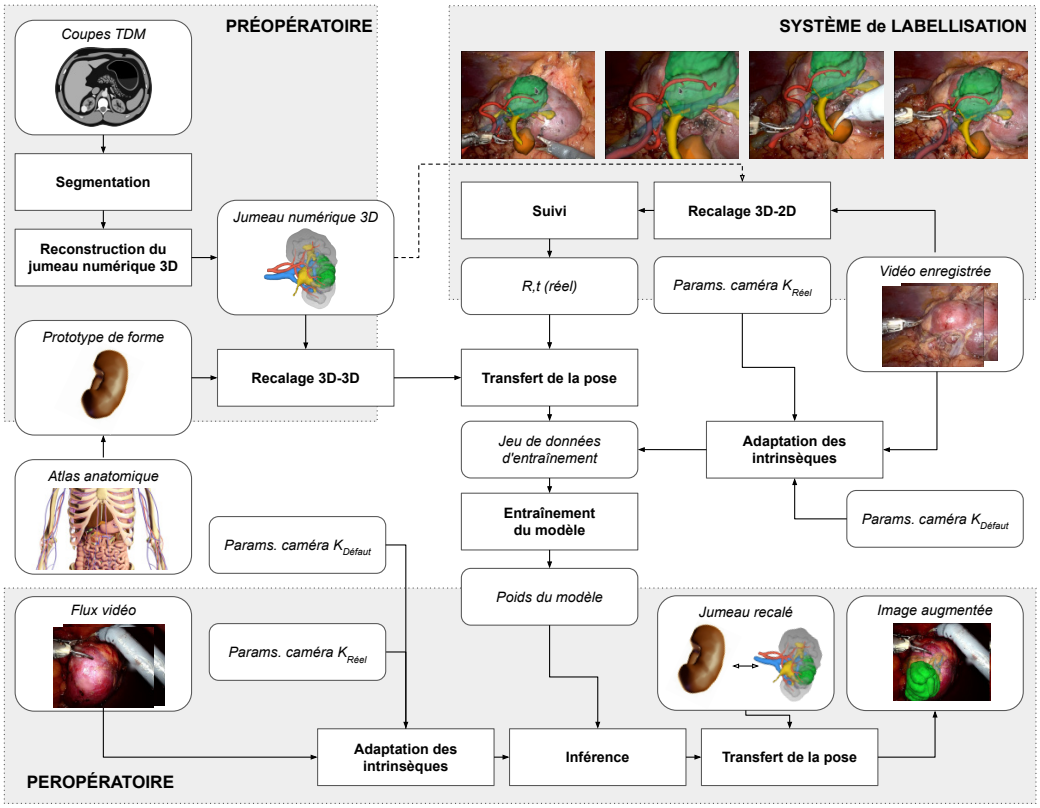


FIGURE 4.25 – Pipeline de la méthode proposée de recalage 3D-2D automatique sans repère, illustrant les étapes préopératoires, peropératoires et de labellisation. Source : auteur.

Le RNA développé pour la tâche de recalage 3D-2D intègre un encodeur d'images de type DINO-BASE-V2 [Oquab 2023] pré-entraîné sur le jeu de données ImageNet [Deng 2009]. Cet encodeur permet d'extraire des cartes de caractéristiques de dimensions $(16, 16, 768)$ à partir d'images endoscopiques standardisées RVB de taille 224×224 pixels. Chaque image d'entrée est associée à une pose labellisée $P = (R, T)$ du JP, où $R \in SO(3)$ représente la rotation et $T \in \mathbb{R}^3$ la translation. La sortie du réseau est un tenseur de dimensions $(1, 12)$ codant l'ensemble des paramètres de pose. Le choix d'un encodage stable et différentiable pour les rotations est essentiel, comme l'ont démontré de récents travaux [Geist 2024].

Les cartes de caractéristiques sont ensuite concaténées avec les paramètres de pose courants et introduites dans un module de régression constitué de trois couches entièrement connectées (fully connected layers en Anglais), utilisant la fonction d'activation LeakyReLU [Maas 2013] et une fonction d'oubli (dropout en Anglais) pour améliorer le gradient et empêcher l'ajustement excessif (overfitting en Anglais). Les paramètres de pose sont affinés de manière itérative ($i = 3$), à l'aide d'un mécanisme de retour d'erreur itératif (IEF) [Carreira 2016, Oberweger 2015], ce qui permet d'aligner le PF sur l'image d'entrée.

La fonction de perte (loss en Anglais) utilisée pour l'apprentissage est définie par :

$$\mathcal{L} = \lambda \mathcal{L}_{\text{pose}} + (1 - \lambda)\mathcal{L}_{\text{mesh}} \tag{4.7}$$

où $\lambda \in [0, 1]$ est un hyperparamètre contrôlant l'équilibre entre deux termes : une fonction de perte sur le maillage tridimensionnel $\mathcal{L}_{\text{mesh}}$, définie comme l'erreur absolue moyenne (Mean Absolute Error (MAE) en Anglais) des distances euclidiennes entre les sommets du maillage prédits et ceux labellisés manuellement ; et la fonction de perte de pose $\mathcal{L}_{\text{pose}}$, définie comme l'erreur absolue moyenne (MAE) entre les coefficients de pose normalisés prédits et labellisés manuellement, exprimés sous la forme de 12 coefficients.

Le réseau de neurones est entraîné sur 100 itérations (epochs en Anglais) avec une taille de lot de 16 (batch size en Anglais). L'optimisation stochastique est assurée par l'algorithme AdamW [Loshchilov 2017], une variante de l'optimiseur Adam, qui découple la décroissance du poids de la mise à jour du gradient. Le taux d'apprentissage (learning rate en Anglais) initial est fixé à 10^{-5} et est ajusté au cours de l'entraînement par un planificateur dynamique de type StepLR.

Afin de faciliter l'entraînement, une stratégie par curriculum learning [Soviany 2022] permet d'augmenter progressivement la complexité de la tâche, reflétant ainsi les évolutions morphologiques du rein au fil de l'intervention chirurgicale. Cette approche exploite les méta-informations de phase chirurgicale annotées lors du processus de labellisation. L'augmentation des images d'entrée est également assurée par des transformations visuelles standards comme le flou gaussien, l'ajout de bruit et la compression artificielle.

4.6.6 Inférence

Lors de l'inférence, l'ensemble des 92 678 376 paramètres du réseau de neurones est gelé, c'est-à-dire que le réseau opère en mode strictement déterministe, sans mise à jour des poids. La prédiction d'une pose P_{pf} du PF est alors effectuée à partir d'une image endoscopique prétraitée, notamment via l'adaptation des paramètres intrinsèques (cf. section 4.6.4).

Cette pose estimée P_{pf}, exprimée dans le repère du PF est convertie dans le repère propre au JP par application de la transformation inverse $T_{\text{pf2jp}} = T_{\text{jp2pf}}^{-1}$, obtenue lors de la phase préopératoire de préparation des données. Ainsi la pose du JP, P_{jp}, s'écrit :

$$P_{\text{jp}} = P_{\text{pf}} T_{\text{jp2pf}} \tag{4.8}$$

Une fois la pose P_{jp} calculée, une visualisation augmentée de l'image peut être générée en temps réel et affichée à l'utilisateur.

4.7 Expériences

L'évaluation de la méthode proposée a été conduite sur l'ensemble du jeu de données constitué à partir de sept patients (cf. section 4.6.2). Pour cela, nous avons adopté une approche Leave-One-Patient-Out (LOPO) en Anglais, qui pour chaque itération, utilise six patients pour entraîner le réseau de neurones et l'unique restant pour l'évaluation. Cette approche croisée permet de reproduire un contexte réaliste d'inférence sur un nouveau cas clinique, tout en maximisant l'utilisation des données disponibles pour l'entraînement.

4.7.1 Évaluation de la performance brute

Les performances brutes de la méthode proposée sont mesurées en comparant la pose prédite, exprimée dans le repère du JP, à la pose labellisée manuellement. Deux types d'erreurs sont considérées : d'une part l'erreur de pose, qui comprend les composantes de translation et de rotation et d'autre part, l'erreur absolue moyenne (MAE) entre les sommets des maillages positionnés respectivement selon la pose prédite et la semi-vérité terrain.

Le réseau de neurones proposé a été entraîné sur un PC Ubuntu équipé d'un processeur Intel Core™ i7-13620H et d'une carte graphique Nvidia RTX 4070M.

Les résultats présentés dans le tableau 4.3 permettent d'évaluer la précision de la méthode de recalage 3D-2D proposée. De faibles erreurs moyennes de translation et de rotation, aux côtés d'une MAE minimale des sommets, indiquent une forte concordance entre les données labellisées manuellement par des experts et les données prédites. Malgré la variabilité de l'anatomie et des mouvements, la méthode a obtenu des performances brutes constantes. Cela souligne sa capacité à aligner de manière autonome le JP avec l'image de la caméra endoscopique, ce qui en fait un atout précieux pour les applications de RA peropératoires.

ID Patient pour l'éval.	Nb d'images	Erreur de translation moyenne (mm)	Erreur de rotation moyenne (°)	Erreur moyenne aux sommets (mm)
1	38 267	13,21	0,54	14,56
2	35 077	12,63	0,62	14,02
3	29 449	16,17	0,45	16,84
4	15 952	11,50	0,43	12,32
5	66 577	14,36	0,94	15,93
6	60 237	13,44	0,72	15,14
7	54 100	12,16	0,32	12,67
MOYENNE		13,55	0,57	14,50
ECART-TYPE		1,55	0,21	1,65

TABLE 4.3 – Évaluation des performances brutes de la méthode de recalage 3D-2D proposée par rapport aux données labellisées manuellement par des experts. Source : auteur.

4.7.2 Évaluation dans le contexte clinique

Une évaluation rétrospective préliminaire de la pertinence clinique de la méthode proposée a été conduite, avec pour objectif principal d'en mesurer l'efficacité et la précision dans des conditions représentatives de la NPRA. Nous avons ciblé spécifiquement l'erreur de recalage de la tumeur, paramètre critique pour la fiabilité du guidage peropératoire et la précision du geste chirurgical.

Pour faciliter l'évaluation quantitative, nous avons sélectionné quatre cas de tumeurs exophytiques (ID 1, 3, 5 et 6), qui présentent par définition une exposition visuelle de la masse tumorale. La comparaison est alors directe entre le contour projeté de la tumeur (issu du modèle augmenté) et le contour réel de la tumeur observé dans l'image peropératoire, comme le montre la figure 4.26.

Le système de RA permet à l'utilisateur final de customiser l'apparence des structures augmentées, notamment en ajustant le niveau de transparence afin de faciliter l'analyse visuelle. À titre de référence, les erreurs d'alignement manuel observées dans des conditions non contraintes sont généralement comprises entre 3 et 9 millimètres. Elles atteignent 20 millimètres lorsque la durée de l'interaction de recalage 3D-2D est limitée à 10 secondes. Dans nos expériences, nous avons observé que le recalage manuel devient significativement plus compliqué pour les opérateurs tiers, qui n'appartiennent pas à l'équipe clinique ayant réalisé le geste et qui ne possèdent donc pas le contexte opératoire spécifique au patient.

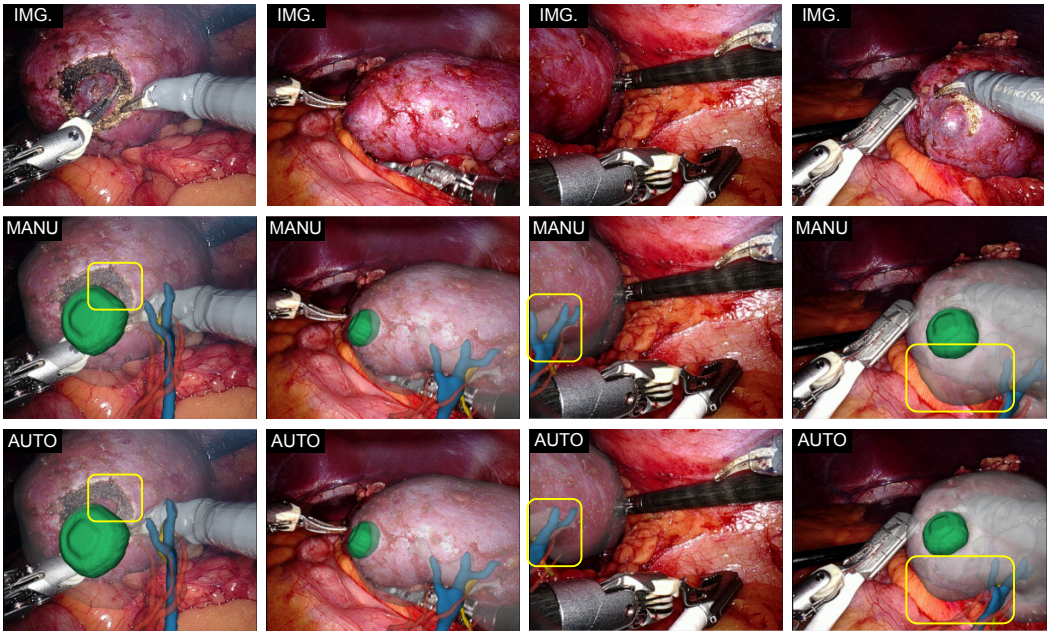


FIGURE 4.26 – Résultats de Réalité Augmentée (RA) générés via la pose du Jumeau du Patient (JP) labellisé manuellement (deuxième ligne) et la pose du Jumeau du Patient (JP) prédite avec notre méthode de recalage 3D-2D automatique (troisième ligne). Source : auteur.

De plus, une incertitude moyenne de modélisation d'environ 14 millimètres a été identifiée entre la reconstruction du jumeau numérique et l'imagerie en coupe. Cette erreur est probablement attribuable à la combinaison d'erreurs systématiques et aléatoires, tant humaines que algorithmiques.

En termes de performances, la méthode proposée atteint un temps d'exécution moyen de $9,4 \pm 0,2$ millisecondes (soit une fréquence de 106 Hz) sous PyTorch 2.5.1. Ce délai comprend à la fois les opérations de prétraitement (standardisation de l'entrée et transfert GPU) et de post-traitement (le transfert de la sortie vers le JP). L'erreur moyenne d'alignement du contour de la tumeur observée sur les quatre cas évalués est de $7,3 \pm 4,1$ millimètres.

À notre connaissance, il s'agit de la seule approche de recalage 3D-2D entièrement automatique, sans repère ni interaction utilisateur, qui atteint une précision compatible avec les exigences de la NPRA. À titre de comparaison, les méthodes de l'état de l'art en chirurgie hépatique font état d'erreurs moyennes de recalage de 30 millimètres à 36 millimètres pour les modèles spécifiques au patient et génériques respectivement [Labrunie 2024]. Notre approche permet de réduire significativement l'erreur sans nécessité d'interaction utilisateur, ni entraînement spécifique au patient, ce qui est prometteur pour sa fiabilité clinique et son déploiement pratique.

4.7.3 Application

La méthode de recalage 3D-2D automatique développée dans ce travail s'intègre dans le cadre applicatif de la génération d'images augmentées peropératoires fixes et dynamiques lors de la NPRA. Les vues peropératoires sont enrichies de structures anatomiques cachées transférées à partir du jumeau tridimensionnel préopératoire du patient.

Deux modes principaux d'intégration sont proposés et illustrés en figure 4.27 : (1) l'inférence directe en temps réel, la méthode de recalage est appliquée à chaque image de la séquence vidéo, fournissant une estimation indépendante de la pose du JP ; (2) le recalage automatique proposé est réalisé sur des IC (cf. section 4.6.3), puis prolongé temporellement à l'ensemble de la séquence à l'aide d'un algorithme de suivi rigide d'organe [Collins 2020, Chandelon 2022].

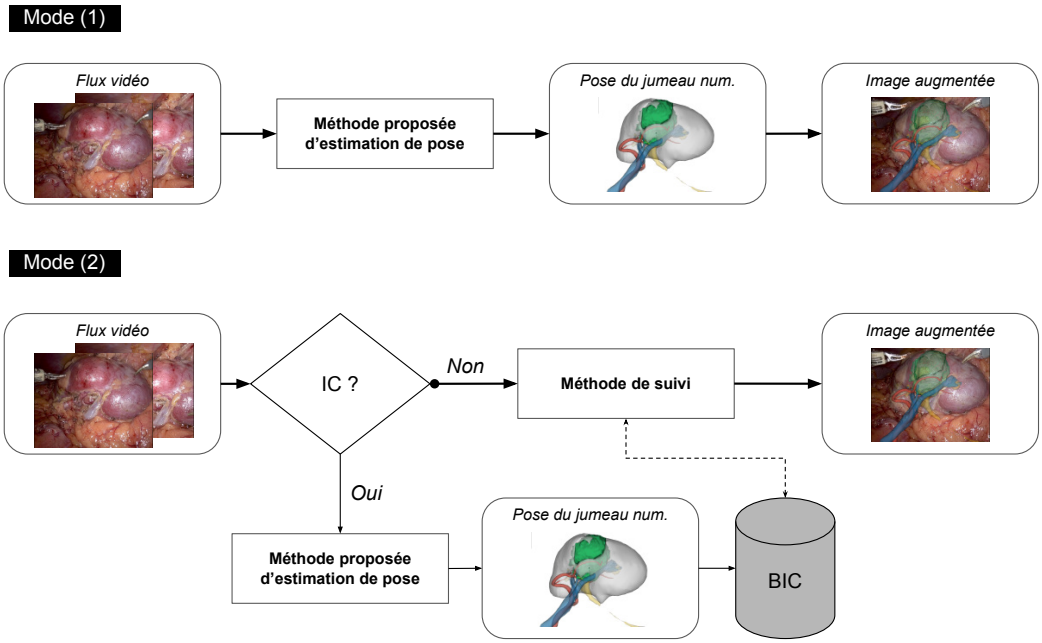


FIGURE 4.27 – Modes principaux d'utilisation de la méthode de recalage 3D-2D proposée : (1) inférence directe sur flux vidéo, (2) recalage ponctuel combiné à un suivi rigide inter-images basé sur une Base d'Images Clés (BIC). Source : auteur.

4.8 Conclusion

Nous avons proposé une méthode de recalage 3D-2D automatique sans primitive, conçue pour être appliquée à la NP. Elle repose sur la régression directe des paramètres de pose à partir d'une image endoscopique peropératoire, via un réseau de neurones entraîné à partir de données labellisées manuellement par des experts chirurgiens. L'approche proposée est bien adaptée aux contraintes temporelles et humaines typiques de la pratique clinique. Aujourd'hui, elle représente la seule solution réalisable capable d'atteindre une précision de recalage de $7,3 \pm 4,1$ millimètres (mesurée sur quatre patients) sans dépendre de l'expertise du chirurgien opérant.

Néanmoins, compte tenu du stade précoce de l'évaluation de faisabilité et bien que la cohérence avec les données labellisées par des experts ait été observée, les affirmations de fiabilité clinique restent préliminaires et doivent être validées par rapport aux résultats chirurgicaux. Les expériences actuelles sont basées sur une petite cohorte (n=7), avec une analyse des erreurs tumorales limitée à quatre cas, ce qui restreint la puissance statistique et la généralisation. Afin d'atténuer le surajustement et de favoriser la généralisation initiale, nous avons utilisé une stratégie de validation croisée LOPO. Notre méthode actuelle ne concerne que le recalage rigide, mais reste pertinente pour le guidage chirurgical dans le volume rénal restreint entourant la tumeur, où les hypothèses de rigidité sont localement valables. Cependant, l'exposition limitée du parenchyme rénal reste une limite du système actuel et n'est pas encore complètement caractérisée pour la voie rétropéritonéale en raison d'un manque de données.

Les travaux futurs porteront sur plusieurs axes complémentaires. D'un côté, l'élargissement de l'ensemble de données pour inclure des cas collectés rétrospectivement (n=30) afin d'améliorer la généralisation via une plus grande diversité anatomique et opératoire. Une extension aux autres phases chirurgicales est également envisagée pour couvrir un spectre plus large de configurations peropératoires.

D'un autre côté, l'intégration d'une composante de déformation au sein du prototype de forme permettra de tenir compte des variations topologiques inter et intra patients, en particulier au niveau des structures vasculaires. L'exploitation du flux vidéo stéréoscopique disponible en NPRA est une source précieuse d'information à explorer. Deux stratégies sont à considérer : injecter les paires stéréoscopiques dans un réseau à un ou deux encodeurs, ou bien apprendre une représentation intermédiaire comme la disparité ou la profondeur. Ce flux stéréoscopique peut également être mobilisé pour renforcer la qualité de la labellisation, en améliorant la reproductibilité et la justesse. Des opérations complémentaires comme l'ajustement de faisceaux peuvent également être envisagées pour accroître la cohérence des données.

Sur le plan méthodologique, l'adaptation dynamique de la fonction de perte via des pondérations apprises constitue une piste prometteuse pour équilibrer automatiquement les termes de pose et de géométrie au cours de l'entraînement. De plus, les résultats récents sur la tokenisation adaptative des images [Dwivedi 2024, Goel 2023], appliquée avec succès à l'estimation de la pose humaine, suggèrent que

des stratégies similaires pourraient être transférées à notre domaine afin d'améliorer la représentation des structures anatomiques saillantes.

Sur le plan perceptif, la collecte d'informations périphériques comme le suivi oculaire (eye tracking) pourrait ouvrir de nouvelles perspectives pour comprendre et modéliser les stratégies attentionnelles des experts. Ces données pourraient être exploitées pour guider dynamiquement l'attention du réseau de neurones en simulant des mécanismes de focalisation proches de ceux des chirurgiens.

Enfin, sur la base de ce module technique de recalage automatique 3D-2D, des caractéristiques cliniques telles que la profondeur d'infiltration de la tumeur, la proximité du système collecteur, les marges chirurgicales et les territoires de perfusion seront explorées pour faciliter la prise de décision chirurgicale. À court terme, cette méthode est destinée à servir d'outil complémentaire à l'échographie peropératoire, avec le potentiel de surmonter ses limites actuelles au fil du temps.

Chapitre 5

Suivi visuel d'organe en temps réel par points clés utilisant une base d'images clés dynamique

---

Dans ce chapitre, nous nous intéressons spécifiquement au suivi d'organes en temps réel dans un système de RA en contexte chirurgical. La méthode de suivi par détection implémentée utilise la mise en correspondance de points clés entre une nouvelle image peropératoire issue du flux vidéo de la caméra endoscopique et une Base d'Images Clés (BIC) dynamique sur lesquelles le jumeau numérique du patient a été préalablement recalé via une méthode de recalage 3D-2D. La méthode de suivi existante est dite statique, car elle est basée sur une BIC construite en début de chirurgie. L'extension proposée dite dynamique, permet de prendre en compte l'évolution de l'environnement peropératoire, les variations d'apparence des tissus au cours de l'intervention et les occultations, avec un organe cible qui se dévoile petit à petit. Un gestionnaire de BIC, responsable du maintien de la cohérence de la base et d'un temps d'exécution compatible avec les contraintes temps réelles, a été ajouté ainsi qu'un sélecteur automatique d'IC afin d'enrichir au fil de l'eau la BIC et d'améliorer la robustesse du suivi de l'organe. La méthode dynamique est déclinée en plusieurs versions capables d'exploiter des flux vidéo endoscopiques monoculaires et stéréoscopiques, avec des applications en chirurgie cœlioscopique de l'utérus et du rein. Les expériences menées montrent une amélioration significative du nombre d'images suivies par rapport à l'approche statique, en particulier dans des contextes de forte déformation tissulaire ou de variations de point de vue. L'intégration de la stéréovision et de la segmentation sémantique renforce la robustesse du suivi, tandis que l'appariement contraint par la géométrie épipolaire permet d'augmenter le taux d'images correctement suivies. La Correction de la Perspective (CP) s'avère déterminante pour maintenir la qualité des appariements en présence de changements angulaires importants entre l'image clé et l'image courante. Enfin, l'évaluation en conditions peropératoires sur plus de vingt-cinq cas cliniques de NPRA a permis de confirmer la faisabilité et la pertinence clinique du suivi visuel dynamique. Les résultats montrent des erreurs de pose et de recalage tumoral compatibles avec une utilisation chirurgicale et mettent en évidence l'acceptabilité du système par les utilisateurs ainsi que son intégration fluide dans le flux opératoire.

Sommaire

5.1 Introduction . . . 213
5.2 Contexte technique général . . . 214
5.3 État de l'art . . . 216
5.3.1 Mise en correspondance d'images . . . 216
5.3.2 Méthodes de suivi visuel en chirurgie assistée par ordinateur 219
5.3.3 Méthodes de réduction de la dérive en suivi visuel . . . 228
5.3.4 Correction des distorsions de la perspective . . . 230
5.4 Segmentation automatique de l'organe dans les images peropératoires . . . 231
5.4.1 Généralités . . . 231
5.4.2 Collecte des données . . . 232
5.4.3 Labellisation des données . . . 235
5.4.4 Entraînement du réseau de neurones . . . 240
5.4.5 Inférence . . . 242
5.5 Suivi monoculaire via une base d'images clés statique . . . 243
5.6 Suivi monoculaire via une base d'images clés dynamique . 247
5.7 Méthode étendue à la stéréovision . . . 250
5.8 Méthode avec correction automatique de la perspective . . 251
5.9 Méthode spécialisée à la néphrectomie partielle robot assistée 254
5.10 Méthode de suivi basée sur l'apprentissage profond . . . 257
5.11 Expériences . . . 259
5.11.1 Segmentation automatique du parenchyme rénal . . . 259
5.11.2 Apport de la méthode dynamique . . . 266
5.11.3 Apport de la stéréovision . . . 269
5.11.4 Apport de la correction de la perspective . . . 270
5.11.5 Évaluation en conditions peropératoires . . . 273
5.12 Conclusion . . . 276

5.1 Introduction

Dans le contexte des systèmes de CAO, et plus spécifiquement en CGI (cf. section 1.2.4.4), le module de suivi visuel peropératoire en temps réel est un composant fondamental. Ce module devient indispensable dans les situations où le recalage 3D-2D ne peut être réalisé à chaque image du flux endoscopique en raison de contraintes temporelles strictes. En effet, les méthodes classiques de recalage 3D-2D (cf. section 4.5) reposent sur des intéractions humaines, qui sont inadaptées à un contexte peropératoire où le temps entre deux images successives est de l'ordre de 16,7 millisecondes pour une fréquence d'acquisition de 60 Hz. Une méthode de recalage 3D-2D automatique fiable et temps réel est alors une première solution (cf. section 4.6). Le module de suivi visuel constitue une deuxième solution, qui nécessite une initialisation préalable qui consiste à l'alignement précis du jumeau numérique sur la vue peropératoire.

En RA (cf. section 2.2) sans marqueur, ce module permet d'estimer le recalage 3D-2D à chaque nouvelle image du flux vidéo, par analyse de son contenu visuel et en exploitant une base de connaissances initiales nommée BIC. Cette BIC se compose d'un ensemble d'IC, pour lesquelles la pose du jumeau numérique tridimensionnel du patient, construit lors de la planification chirurgicale (cf. section 1.3.1), est connue de manière fiable par recalage 3D-2D (cf. section 4.5). A l'aide de la BIC, le module de suivi visuel peut générer en temps réel une séquence d'images augmentées, affichées au chirurgien afin de guider le geste opératoire.

Pour des interventions de CMI comme l'hystérectomie ou la NP, le module de suivi visuel joue un rôle pivot. Il permet de passer d'une vue augmentée statique, issue du recalage 3D-2D sur une image unique, à une séquence augmentée sur un ensemble d'images successives. Concrètement, le chirurgien conserve ainsi la liberté de manipuler la caméra ou l'organe sans perdre l'affichage de la RA. Cliniquement, ce module doit être en mesure de maintenir une qualité d'alignement du jumeau numérique sur la vue peropératoire suffisante pour la réalisation du geste chirurgical.

La BIC constitue donc un condensé de mémoire visuelle pour le module de suivi, valable aux instants où les IC qui la composent ont été capturées et recalées. Si la BIC reste statique tout au long de la chirurgie, ou en d'autres termes qu'elle n'est pas enrichie par de nouvelles IC, elle devient rapidement obsolète dès que l'apparence de l'organe évolue. Des zones initialement invisibles peuvent apparaître, tandis que d'autres se transforment de manière radicale. Dans cette configuration statique, le suivi visuel perd progressivement en efficacité, jusqu'à une perte complète du suivi visuel et donc une interruption de la RA.

Pour faire face à ces limites et garantir un suivi visuel d'organe robuste, qui maintient un niveau de performance cliniquement satisfaisant lors des phases où la RA peut avoir un impact clinique significatif, notamment lors de la localisation, la délimitation et l'exérèse de la tumeur, il est nécessaire de développer des mécanismes adaptatifs capables de tenir compte de l'évolution continu de la scène observée. Nous proposons ainsi un module de suivi visuel fondé sur une BIC dynamique, où les IC

peuvent être mise à jour automatiquement en fonction des besoins, exploitable aussi bien en chirurgie endoscopique monoculaire que stéréoscopique.

La tâche soulève plusieurs défis méthodologiques. Le processus de sélection des IC est intrinsèquement multicritère et critique. L'ajout d'une IC inappropriée peut entraîner une dérive du suivi et compromettre la cohérence de l'alignement du jumeau numérique sur la vue peropératoire. Les contraintes matérielles des systèmes informatiques sont également à prendre en compte. La BIC ne peut pas avoir une taille infinie car cela entraînerait des coûts computationnels importants lors de l'exploitation de la base. Un mécanisme de gestion mémoire doit ainsi permettre l'oubli sélectif des IC devenues obsolètes car dépassées temporellement, afin de préserver les performances du suivi en temps réel.

Les contributions techniques de ce travail sont de cinq ordres : 1) le développement d'une méthode de suivi visuel par détection basée sur une BIC dynamique. Ce système repose sur deux composants : un sélecteur d'IC qui combine un ensemble de critères de sélection et un gestionnaire de BIC capable de manager la base afin d'en assurer la cohérence spatio-temporelle tout en respectant les contraintes de performances ; 2) l'extension de la méthode de suivi visuel basée sur une BIC dynamique au cas de la stéréovision en exploitant la géométrie épipolaire entre les vues gauche et droite en NPRA ; 3) la sélection précise de la zone de recherche afin de réduire le temps de calcul et d'améliorer les performances du suivi par détection ; 4) la contribution au développement d'une méthode de CP utilisant la stéréovision afin de renforcer la détection et la description de points clés en chirurgie sous de fortes variations de points de vue ; et 5) l'utilisation d'un réseau de neurones d'inférence de trajectoires visant à remplacer la chaîne conventionnelle de détection, description et mise en correspondance de points clés.

5.2 Contexte technique général

La génération de contenu en RA peropératoire s'appuie sur deux étapes distinctes, représentées schématiquement en figure 5.1. La première, préopératoire, également désignée comme la planification chirurgicale (cf. section 1.3.1), se déroule hors ligne, c'est-à-dire en dehors des contraintes de temps réel. Elle inclut la réalisation d'examens d'imagerie, la segmentation des structures anatomiques d'intérêt, puis la génération des maillages surfaciques tridimensionnels aboutissant au jumeau numérique du patient.

La seconde étape, peropératoire, s'inscrit dans un cadre temps réel et repose sur l'exploitation du flux vidéo endoscopique. Après l'acquisition des images, quatre étapes se succèdent : l'étalonnage géométrique de la caméra endoscopique monoculaire ou stéréoscopique (cf. section 2.2.6), le recalage 3D-2D du jumeau numérique sur la vue peropératoire (cf. section 4.5), le suivi visuel de l'organe et le rendu graphique des éléments augmentés (cf. section 1.3.3), en vue de leur affichage au chirurgien.

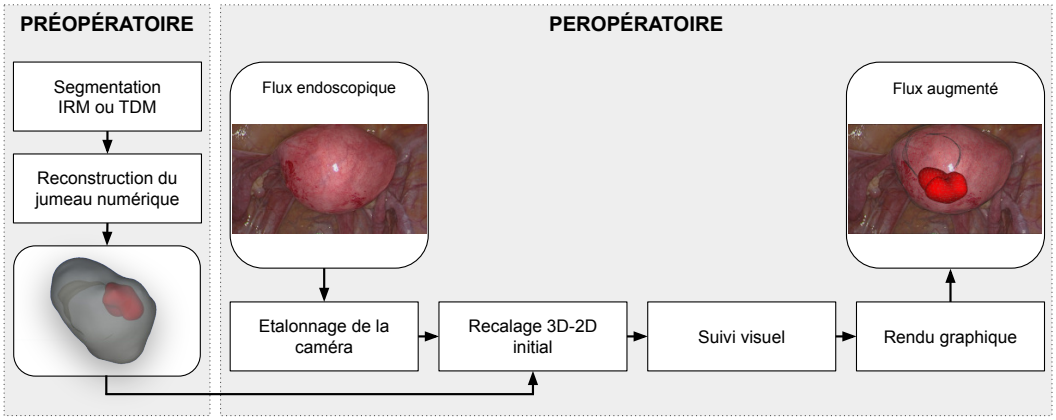


FIGURE 5.1 – Représentation schématique des étapes préopératoire et peropératoire indispensables à la réalisation de la Réalité Augmentée (RA) en chirurgie endoscopique. Source : auteur.

Du point de vue technique, deux grandes familles de méthodes permettent le suivi d'objets dans une séquence d'images issue d'une caméra observant une scène tridimensionnelle.

La première famille de méthodes, dite de suivi par appariement, suppose que la cinématique de l'objet suivi soit relativement régulière et que les variations inter-images restent modérées. A partir d'un gabarit, aussi appelé template, le système recherche sa correspondance dans l'image suivante à l'aide de critères de similarité locale. Ces méthodes présentent l'avantage d'un temps de calcul réduit car la recherche est locale et ne nécessite pas de détecteur. Elles sont en mesure de produire des estimations de position fines, potentiellement sous-pixelliques. Toutefois, elles sont particulièrement sensibles aux erreurs cumulatives, qui peuvent induire une dérive du suivi, ainsi qu'aux occultations, aux changements d'échelle et aux déformations. Leur extension au suivi multi-objets requiert souvent l'ajout d'heuristiques spécifiques [Forsyth 2002].

La seconde famille de méthodes, dite de suivi par détection (tracking by detection en Anglais) repose sur la capacité à ré-identifier l'objet indépendamment sur chaque image. Pour cela, un détecteur d'objets est appliqué image par image. Les réponses du détecteur sont ensuite associées entre images afin de reconstruire des trajectoires temporelles. Ces méthodes offrent une plus grande robustesse aux changements brusques d'apparence ou de pose. Elles s'adaptent naturellement aux configurations multi-objets et s'accordent parfaitement avec les détecteurs d'objets modernes. Cependant, elles restent sensibles aux faux négatifs, qui peuvent interrompre les trajectoires. Elles sont généralement plus coûteuses en temps de calcul, en raison de l'appel répété au détecteur [Forsyth 2002].

Contrairement aux méthodes de suivi visuel par appariement basées sur la localisation et reconstruction simultanée (SLAM), qui supposent une reconstruction incrémentale de la scène, le suivi par détection se révèle particulièrement adapté à des environnements peropératoires complexes et dynamiques. Dans le contexte spécifique de la chirurgie cœlioscopique, le SLAM reste un défi en raison de la nature répétitive de la texture du tissu, des mouvements rapides et imprévisibles de la caméra, du flou de mouvement et des changements d'apparence liées à la présence de sang, de fumée ou de marques de coagulation [Chauvet 2018].

Le suivi par détection présente en autre avantage majeur, sa capacité à reprendre trivialement le suivi après des périodes de non-visibilité de l'organe, par exemple après un retrait temporaire de l'endoscope ou un nettoyage de l'optique. Il permet ainsi d'assurer une continuité fonctionnelle de la RA sur une durée prolongée.

La suite de nos travaux s'appuie sur un module de suivi visuel par détection initialement conçu pour la CMI de l'utérus, tel que proposé par Collins et al. [Collins 2020], et présenté en section 5.5. Cette approche sera qualifiée de méthode de suivi visuel statique. L'ensemble des contributions décrites dans ce chapitre visent à étendre ou adapter la méthode statique afin d'en améliorer la robustesse, la durabilité et l'adaptabilité aux contraintes cliniques réelles.

5.3 État de l'art

Dans cette section, nous nous intéresserons pour commencer aux travaux antérieurs pertinents relatifs à la mise en correspondance d'images et plus particulièrement aux méthodes de détection, description et d'appariement de points clés. Nous poursuivons avec l'étude des approches de suivi visuel adaptées au contexte de la RA appliquée à la chirurgie abdomino-pelvienne. Nous abordons ensuite les méthodes de réduction de la dérive en suivi visuel, avant de traiter des techniques de correction des distorsions liées à la perspective, au service de l'amélioration des performances des méthodes de suivi visuel.

5.3.1 Mise en correspondance d'images

La mise en correspondance d'images constitue un problème central et récurrent en vision par ordinateur. Cette sous-tâche, souvent critique, intervient dans de nombreuses applications telles que la reconstruction tridimensionnelle (SfM ou SLAM par exemple), la stabilisation vidéo [Wang 2023] ou encore la recherche d'images (image retrieval en Anglais) [Datta 2008]. Elle demeure encore aujourd'hui un défi majeur en raison des multiples transformations que subissent les images, qu'il s'agisse de changements de point de vue, de conditions d'illumination, d'occultation ou d'hétérogénéité des capteurs.

Deux catégories de méthodes se distinguent selon les éléments à appareiller : d'une part, celles basées sur des régions qui utilisent des stratégies de comparaison directe comme la similarité entre pixels et d'autre part, celles basées sur des caractéristiques qui offrent plus de stabilité et reposent sur la détection de points clés, suivie de la construction de descripteurs [Huang 2024].

L'extraction de caractéristiques est une étape importante, réalisable via des méthodes traditionnelles par analyse des variations d'intensité (coins, contours, blobs) ou par interprétation des cartes de caractéristiques (feature map en Anglais) inférées par des RNA [Huang 2024].

Parmi les méthodes standards, il est possible de retrouver FAST [Rosten 2006], HL [Mikolajczyk 2004], KAZE [Alcantarilla 2012] ou encore SIFT [Lowe 2004] et SURF [Bay 2006], qui présentent l'avantage d'être invariants aux changements d'échelle et de rotation.

Les méthodes de détection à base d'apprentissage profond comme DetNet [Lenc 2016] ou MagicPoint [DeTone 2018], attribuent à chaque pixel de l'image un score correspondant à la probabilité que le pixel soit une caractéristique saillante.

Une fois les points clés détectés, des descripteurs sont calculés en fonction de leur localisation et des relations avec leur voisinage. Ils sont généralement représentés par des vecteurs qui encapsulent l'information relative à la structure locale de l'image au point considéré (keypoint en Anglais). Dans cet espace de description, des points différents sont idéalement séparés tandis que des points similaires se retrouvent rapprochés [Huang 2024].

Deux familles de descripteurs dominent la littérature : ceux fondés sur la statistique des gradients comme SIFT [Lowe 2004], qui encode l'orientation et l'amplitude des gradients dans une région locale et ceux basés sur la comparaison d'intensités locales comme BRIEF [Calonder 2010].

Les travaux récents, comme SuperGlue [Sarlin 2020] ou LoFTR [Sun 2021], exploitent l'apprentissage profond et les mécanismes d'attention pour produire des descripteurs discriminants et directement optimisés pour l'appariement. Certaines méthodes bout en bout intègrent simultanément détection et description via des réseaux de neurones comme LIFT [Yi 2016], D2-Net [Dusmanu 2019], R2D2 [Revaud 2019], ASLFeat [Luo 2020b] ou DALF [Potje 2023].

Une liste non exhaustive des méthodes de détection et de description est donnée en annexe L.

L'appariement consiste ensuite à rechercher, pour chaque descripteur, son plus proche voisin dans l'espace de description via des distances euclidiennes ou de similarités cosinus. En pratique, une stratégie en deux étapes est généralement adoptée : calcul des correspondances, suivi d'un filtrage fondé sur des contraintes de consistance géométrique [Huang 2024].

Le test proposé par Lowe [Lowe 2004], communément désigné sous le nom de TRL, repose sur le rapport entre la distance du premier et celle du deuxième plus proche voisin. L'appariement est jugé non ambigu lorsque ce rapport est inférieur à un seuil donné, typiquement 0,80.

Afin d'éliminer les fausses correspondances (outliers en Anglais), le filtrage peut être assuré par deux grandes catégories de méthodes : celles qui définissent manuellement des contraintes explicites et celles qui exploitent l'apprentissage. Parmi les approches manuelles, il est possible de distinguer trois sous-catégories.

Les méthodes basées sur le rééchantillonnage qui s'inspirent de RANSAC [Fischler 1981] et de ses nombreuses variantes comme Extreme Value SAmple Consensus (EVSAC) en Anglais [Fragoso 2013], N Adjacent Points SAmple Consensus (NAPSAC) en Anglais [Torr 2002], PROgressive SAmple Consensus (PROSAC) en Anglais [Chum 2005], GroupSAC [Ni 2009], Maximum Likelihood Estimation SAmple Consensus (MLESAC) en Anglais [Torr 2000], Universal SAmple Consensus (USAC) en Anglais [Raguram 2012] et bien d'autres [Chum 2003, Lebeda 2012, Barath 2019, Jegou 2008, Avrithis 2014, Li 2015, Schönberger 2017].

D'autres méthodes non-paramétriques ont émergé afin de traiter des transformations d'images complexes et des mouvements non-rigides d'objets. Elles s'appuient sur la relaxation spectrale, qui regroupe les correspondances en ensembles optimaux et construit une matrice d'adjacence intra-groupe plutôt qu'un appariement individuel. Le problème se formule ainsi en une recherche de la grappe qui maximise un score d'appariement [Leordeanu 2005, Liu 2010, Lin 2014, Lin 2017b].

La théorie de corrélation des mouvements peut également être utilisée via le consensus de champ vectoriel (Vector Field Consensus (VFC) en Anglais) [Ma 2014]. D'autres méthodes comme Gaussian Field Criterion (GFC) en Anglais [Ma 2018] ou Identifying point Correspondence Function (ICF) en Anglais [Li 2010] ont également été développées. Les méthodes non-paramétriques demeurent coûteuses en calcul et sensibles au bruit, contrairement aux méthodes inspirées de RANSAC.

Certaines méthodes exploitent quant à elles la relaxation de contraintes géométriques, en vérifiant la cohérence des correspondances à partir du voisinage. Parmi elles figurent GMS [Bian 2017], Locality-Preserving Matching (LPM) en Anglais [Ma 2019b], Spatially Consistent RAndoM SAmple Consensus (SCRAMSAC) en Anglais [Sattler 2009], Feature-Based Deformable Surface Detection (FBDSD) en Anglais [Pizarro 2012], COherence based DEcision boundaries (CODE) en Anglais [Lin 2017b], TopKRP [Jiang 2019a], Neighborhood Manifold Representation Consensus (NMRC) en Anglais [Ma 2022], Neighborhood Affine Subspace Embedding (NASE) en Anglais [Li 2022c], Density Based Spatial Clustering method of Applications with Noise (DBSCAN) en Anglais [Jiang 2019b], Guided Local Outlier Factor (GLOF) en Anglais [Wang 2021a], Adaptive Locally-Affine Matching (AdaLAM) en Anglais [Cavalli 2020] ou LOcality-guided Global-preserving Optimisation (LOGO) en Anglais [Xia 2022].

La méthode AdaLAM [Cavalli 2020] illustre une approche hybride qui combine contraintes locales et robustesse statistique. Elle sélectionne d'abord des graines (seeds en Anglais), c'est-à-dire des correspondances fiables et spatialement bien réparties, qui définissent chacune une région locale. Dans chaque région, les correspondances proches en termes de position, d'orientation et d'échelle sont regroupées, puis une vérification affine est réalisée via RANSAC. Les erreurs de transformation

permettent ensuite de calculer une mesure de confiance statistique issue d'un test d'hypothèse nulle. Les inliers sont alors retenus selon un seuil de confiance adaptatif.

La méthode Globally-and-Locally-Affine Matching (GaLAM) en Anglais [Lu 2025] propose une amélioration du schéma précédent. Dans une première étape, les correspondances locales sont affinées à l'aide de contraintes géométriques explicites (distances, rotation, échelle) et des transformations affines avec un seuil fixe, afin de réduire le coût de calcul par rapport à AdaLAM. Une seconde étape applique PROSAC [Chum 2005] afin de valider la cohérence des graines vis-à-vis d'une matrice fondamentale globale, ce qui permet d'éliminer des correspondances incohérentes qui échappent au filtrage local. Expérimentalement, GaLAM surpasse AdaLAM ainsi que d'autres méthodes comme GMS, LPM ou SuperGlue, en termes de précision et de robustesse sur plusieurs jeux de données (HPatches, SUN3D, Middlebury).

Enfin, les méthodes basées sur l'apprentissage profond utilisent l'information contextuelle pour apprendre à distinguer les bons appariements des mauvais [Chen 2021b, Rocco 2018, Zhao 2019]. Parmi elles, il est possible de retrouver Point net with Context Normalisation (PointCN) en Anglais [Yi 2018], Neural Nearest Neighbors Networks (N3Net) en Anglais [Plötz 2018], Order-Aware Network (OANet) en Anglais [Zhang 2019a], Learning Mismatch Removal (LMRe) en Anglais [Ma 2019a], Local Structure Visualisation-Attention Network (LSV-ANet) en Anglais [Chen 2021a], Graph Attention Network (GANet) en Anglais [Jiang 2022b], ainsi que des architectures qui reposent sur des réseaux de neurones graphiques (Graph Neural Network (GNN) en Anglais) comme SuperGlue [Sarlin 2020], ClusterGNN [Shi 2022] ou LightGlue [Lindenberger 2023].

Actuellement, les méthodes manuelles restent globalement plus performantes que les approches par apprentissage profond en termes de complexité de calcul. Cependant, elles présentent encore des limites dans le traitement d'images soumises à des transformations géométriques inconnues ou issues de modalités hétérogènes.

5.3.2 Méthodes de suivi visuel en chirurgie assistée par ordinateur

Le suivi visuel repose sur l'analyse du contenu des images acquises par une caméra afin d'estimer les déplacements et la position des objets observés. En chirurgie, cela inclut à la fois les déplacements de la caméra, des instruments et des tissus [Schmidt 2024a].

L'environnement peropératoire présente toutefois des défis spécifiques comme le manque de texture des tissus [Widya 2019] ; la présence de réflexions spéculaires à la surface des tissus humides [Makki 2023], qui nécessitent souvent un traitement préalable [Zhao 2022a] ; ainsi que les déformations d'organes, parfois partiellement en dehors du champ de la caméra, ce qui complexifie davantage l'analyse [Azagra 2023]. Dans de telles conditions, il est donc nécessaire de formuler des hypothèses afin de modéliser cet environnement complexe [Schüle 2022]. A ces difficultés s'ajoutent la présence de sang et de fluides qui peuvent troubler l'optique de la caméra endoscopique [Richter 2021], ou encore la fumée qui perturbe la visibilité du champ

opératoire [Liu 2023]. La robustesse d'un système de suivi visuel peropératoire dépend donc étroitement de sa capacité à gérer ces contraintes [Schmidt 2024a].

D'un point de vue macroscopique, un module de suivi visuel repose sur l'intégration de diverses techniques de vision par ordinateur, regroupées en six grands groupes illustrés en figure 5.2 et détaillés dans la suite de cette section.

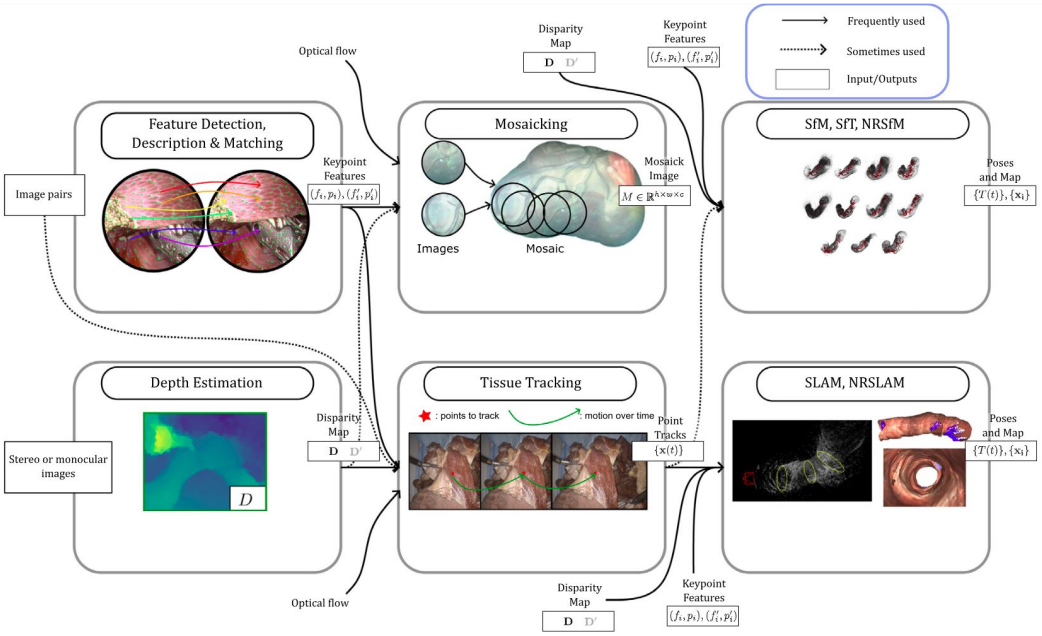


FIGURE 5.2 – Représentation schématique du champ des possibles pour l'assemblage de diverses techniques de vision par ordinateur dans un module de suivi visuel. Source : générée par l'auteur à partir de [Schmidt 2024a].

Appariement de caractéristiques Les premières approches de suivi visuel reposent sur la détection, la description et l'appariement de caractéristiques (cf. section 5.3.1). Ces méthodes, encore employées en neurochirurgie [Jiang 2015, Jiang 2016] et en endoscopie [Marmol Velez 2017], exploitent majoritairement des algorithmes standards [Schmidt 2024a] ou des réseaux de neurones comme SuperPoint [Barbed 2023, DeTone 2018]. Dans le cas de ces derniers, des stratégies d'adaptation de domaine ont été proposées, notamment à partir de données chirurgicales semi-synthétiques reconstruites via COLMAP [Schonberger 2016] ou au travers de réseaux développés spécifiquement pour cet environnement comme Real-Time Rotated descriptor (ReTRo) en Anglais [Schmidt 2021].

Photomosaïque La photomosaïque (Mosaicking en Anglais) vise à générer une image composite à partir d'une séquence d'images. Pour cela, une première étape consiste à mettre en correspondance les caractéristiques issues d'un détecteur-descripteur afin d'estimer les transformations affines qui alignent les images [Atasoy 2008]. Une seconde étape de correction des couleurs est généralement effectuée [Burt 1983]. L'image composite permet ainsi de créer un champ de vue élargi de la scène observée [Seshamani 2006].

Les premières applications rapportées concernent notamment le rein et la rate [Burt 1983, Miranda-Luna 2008]. Plus récemment, des méthodes de photomosaïque basées sur l'apprentissage profond utilisent des réseaux de neurones convolutifs pour l'estimation des homographies [Bano 2019]. Dans certains cas cliniques, la photomosaïque peut être fondée sur la segmentation et l'appariement de structures vasculaires [Bano 2020, Bano 2023]. Toutefois, en contexte chirurgical, la photomosaïque reste limitée par les artéfacts visuels et la présence de sang [Schmidt 2024a].

Éstimation de profondeur Lorsque le suivi implique des points tridimensionnels, il peut être important de disposer d'une estimation de la profondeur de la scène observée. En endoscopie, plusieurs techniques existent pour estimer la disparité ou la profondeur (cf. section 2.2.6) selon que les données soient monoculaires ou stéréoscopiques. Dans un contexte chirurgical, un prétraitement des images, tel que l'ajustement de la luminosité, est souvent requis. De plus, la prise en compte d'informations temporelles peut renforcer la cohérence des estimations [Li 2023].

Suivi tissulaire Le suivi des tissus, qui vise à estimer les déplacements des surfaces et des organes, peut être réalisé au moyen du flux optique (optical flow en Anglais) ou de l'appariement de caractéristiques temporelles, par exemple via le traqueur Kanade-Lucas-Tomasi (KLT) [Lucas 1981, Tomasi 1991].

En chirurgie, il est notamment utilisé pour compenser les mouvements comme les battements cardiaques, par combinaison de la détection STAR [Agrawal 2008], de la description BRIEF [Calonder 2010] et des modèles de déformation non rigide tels que Thin-Plate Splines (TPS) en Anglais [Richa 2008, Yip 2012]. Dans le contexte de la chirurgie endoscopique, le détecteur STAR, le descripteur BRIEF et les détecteurs/descripteurs SIFT et SURF ont démontré leur efficacité dans les scènes chirurgicales dynamiques [Schmidt 2024a, Yip 2012].

L'intégration d'un filtre de Kalman étendu (Extended Kalman Filter (EKF) en Anglais) constitue par ailleurs une stratégie courante pour accroître la robustesse du suivi face au bruit [Giannarou 2012]. Dans les zones faiblement texturées, le recours au flux optique dense améliore également les performances [Du 2015].

Reconstruction tridimensionnelle : SfM, NRSfM et SfT Certaines approches s'écartent de l'utilisation directe du flux optique ou de l'appariement de caractéristiques éparses et consistent à estimer la déformation d'un modèle texturé projeté dans l'image, principe proche du Shape-from-Template (SfT) en Anglais [Collins 2016].

Les approches par apprentissage profond incluent par exemple l'inférence haute fréquence du flux optique dense avec FlowNet [Ihler 2020]. Des variantes qui exploitent des représentations éparses [Schmidt 2022a], des architectures récurrentes [Schmidt 2022b] ou encore l'analyse tridimensionnelle [Schmidt 2023] ont été proposées.

En 2024, le challenge SurgT [Cartucho 2024], dédié à l'évaluation des méthodes de suivi tissulaire, a récompensé en troisième position le traqueur CSRT [Lukezic 2017], en deuxième position une variante améliorée de CSRT intégrant la gestion des occultations et l'ajustement d'échelle. Enfin, en première position, un traqueur basé sur le flux optique dense combinant ARFlow [Liu 2020] et PWC-Net [Sun 2018].

Les techniques de reconstruction tridimensionnelle hors-ligne comme le SfM, le Non-Rigid Structure-from-Motion (NRSfM) en Anglais et le SfT constituent des composants fréquemment utilisés dans les modules de suivi.

Le SfM permet de reconstruire un environnement rigide à partir d'images multiples et d'estimer les poses de caméra correspondantes. Bien qu'effectué hors-ligne, il peut permettre d'élargir le champ de vue [Hu 2007, Hu 2012, Tretschk 2023].

Le NRSfM lève l'hypothèse de rigidité, ce qui rend l'optimisation plus complexe et coûteuse en calcul. Des applications cliniques existent notamment en chirurgie cardiaque, via l'utilisation de Low Rank Shape Model (LRSM) en Angalis [Torresani 2008] pour la reconstruction d'un cœur battant [Hu 2009]. Ou encore en endoscopie digestive pour modéliser des structures tubulaires sous l'hypothèse d'isométrie [Sengupta 2021]. Afin de rendre le problème tractable, des simplifications sont généralement introduites.

Le SfT débute par la définition ou la construction d'un template canonique, qui est ensuite aligné sur les images successives. Cette étape peut être réalisée à partir d'un modèle SfM rigide [Malti 2012], puis enrichie par SfM déformable ou Shape-from-Shading (SfS) en Anglais, basé sur des modèles de lumière comme Lambert ou Cook-Torrance avec micro-facettes de Beckmann [Malti 2014]. L'intégration de contraintes issues du jumeau numérique a également été étudiée [Cheema 2018, Zhang 2021c].

SLAM et variantes non rigides Les méthodes de type SLAM et en particulier Visual Simultaneous Localisation and Mapping (VSLAM) en Anglais, visent à construire une carte de l'environnement (souvent représentée par un nuage de points) et à estimer la localisation de la caméra dans ce même espace. L'optimisation s'appuie généralement sur une méthode de moindres carrés non linéaire, minimisant un ensemble de termes d'erreur définis dans un graphe de poses. Cette cartographie fournit une mémoire persistante de l'environnement.

Le SLAM rigide est largement utilisé dans différentes applications cliniques comme la chirurgie des sinus [Burschka 2004]. Parmi les solutions les plus répandues figure ORB-SLAM [Mur-Artal 2015], qui exploite des caractéristiques éparses ORB [Rublee 2011]. La relocalisation est assurée par un sac de mots (Bag of Words (BoW) en Anglais) appelé DBoW2 [Gálvez-López 2012], suivi d'un raffi-

nement de la pose par ajustement de faisceaux local et global. Un filtre EKF est également utilisé. La carte initiale éparse peut être densifiée par l'exploitation du flux optique [Mahmoud 2016, Mahmoud 2018].

Une autre approche, Direct Sparse Odometry (DSO) en Anglais [Engel 2017, Ma 2019c, Huo 2023], repose directement sur l'estimation de la profondeur, obtenue soit géométriquement ou soit par apprentissage, plutôt que sur l'appariement de caractéristiques.

Des méthodes hybrides comme BodySLAM, combinent CycleVO (estimation non supervisée de pose monoculaire, inspirée de CycleGAN [Zhu 2017]) et InfoGAN [Chen 2016] avec Zoe [Bhat 2023], un réseau d'estimation de la profondeur, pour produire une carte tridimensionnelle cohérente de la scène chirurgicale [Manni 2024].

Des contraintes spécifiques liées aux déplacements de la caméra endoscopique dans un trocart fixe peuvent être intégrées dans le problème d'optimisation, comme dans la méthode RCM-SLAM [Vasconcelos 2019]. Une extension aux caméras endoscopiques stéréoscopiques a aussi été proposée via l'utilisation de la géométrie épipolaire [Jia 2021, Zhou 2021].

Le Non-Rigid Simultaneous Localisation and Mapping (NR-SLAM) en Anglais pose des difficultés supplémentaires, puisqu'il doit distinguer les déformations réelles des tissus du bruit ou des effets de la caméra. Par exemple, il est impossible de découpler totalement la pose de la caméra de la taille des objets en environnement non contraint. Idem, la pose de la caméra ne peut pas être découplée du mouvement des tissus sans référence fixe. Les solutions proposées reposent souvent sur des modèles éléments finis combinés à une méthode de suivi rigide comme ORB-SLAM. Elles présupposent une parfaite connaissance des forces appliquées, hypothèse rarement vérifiable en chirurgie [Schüle 2022].

D'autres approches comme MIS-SLAM [Song 2018] associent ORB-SLAM rigide à un suivi déformable basé sur Embedded Deformation (ED) en Anglais [Sumner 2007], utilisé comme modèle de mouvement. De même, Deformable Simultaneous Localisation and Mapping (DefSLAM) en Anglais [Lamarca 2020] s'appuie sur un maillage triangulaire dense dont l'optimisation correspond à la minimisation de l'erreur de reprojection des points clés détectés dans les images. La variante Semi-Direct - Deformable Simultaneous Localisation and Mapping (SD-DefSLAM) en Anglais [Gómez-Rodríguez 2021] enrichit ce schéma par du flux optique.

Certaines approches privilégient des représentations éparses afin de limiter l'impact des discontinuités de mouvement [Lamarca 2022, Rodríguez 2022]. Néanmoins, l'ensemble de ces méthodes de suivi visuel demeure sensible aux conditions particulières d'éclairage, aux limites de représentativité des modèles déformables, aux contraintes computationnelles et à la difficulté d'évaluer l'incertitude en cas de perte de suivi [Schmidt 2024a].

Méthodes basées sur l'apprentissage profond Les approches récentes considèrent le suivi visuel comme un problème de bout-en-bout, résolu par des réseaux de neurones profonds spécialisés. Ils se distinguent principalement par la dimension du domaine d'analyse (bidimensionnel ou tridimensionnel) et par la nature des données suivies (éparses ou denses).

CoTracker [Karaev 2024b] constitue une référence. Il repose sur une architecture transformer avec fenêtres temporelles glissantes et mécanismes d'attention spatio-temporelle, appliqués au suivi conjoint de points bidimensionnels. Les trajectoires sont décrites par des caractéristiques spatiales, temporelles et d'apparence extraites par un réseau de neurones convolutifs, puis raffinées de manière itérative par le transformer.

Son évolution CoTracker3 [Karaev 2024a], reprend cette architecture en l'enrichissant d'une attention croisée et d'une stratégie d'entraînement semi-supervisée basée sur des vidéos réelles pseudo labellisées par des modèles enseignants préexistants comme TAPIR ou CoTracker [Karaev 2024b]. Deux variantes existent : une version online (suivi glissant, traitement séquentiel) et une version offline (accès à toutes les images simultanément).

AllTracker [Harley 2025], pousse cette logique vers le suivi dense et de longue portée des pixels dans les vidéos haute résolution. Il utilise cependant un encodeur ConvNeXt-Tiny [Liu 2022b]. En sortie, les champs de flux, la visibilité et la confiance pour chaque pixel de l'image de requête sont inférés. Sur les évaluations réalisées par les auteurs, les performances de AllTracker dépassent celles de CoTracker3 [Karaev 2024a] en précision sur l'ensemble de données d'évaluation TAP-Vid [Doersch 2022] et en fréquence d'inférence.

Dans la gamme des méthodes d'analyse via un espace tridimensionnel, Dense Efficient Long-range 3D Tracking for Any video (DELTA) en Anglais [Ngo 2024] propose un suivi dense et longue portée des pixels des vidéos monoculaires, sous la forme de trajectoires tridimensionnelles (X, Y, Z). L'architecture est basée sur un transformer spatio-temporel inspiré de CoTracker [Karaev 2024b] et une stratégie coarse-to-fine. Le suivi est meilleur et plus rapide, cependant la courte fenêtre temporelle limite la taille de la séquence d'entrée. De plus, la performance dépend de la qualité de l'estimation de la profondeur.

SpatialTracker [Xiao 2024], est également une méthode de suivi itératif de trajectoires tridimensionnelles via projection de points bidimensionnels. La régularisation apprise suit l'hypothèse as-rigid-as-possible et chaque image est encodée en un ensemble de trois cartes de caractéristiques (triplane encoding en Anglais). Comme précédemment, la performance est limitée par la qualité de l'estimateur de profondeur.

Tracking Any Point In Persistent 3D geometry (TAPIP3D) en Anglais [Zhang 2025] utilise un espace tridimensionnel persistant. Les points bidimensionnels dans le plan image sont projetés dans l'espace tridimensionnel via l'utilisation des paramètres intrinsèques et de la profondeur. Le suivi visuel itératif est réalisé dans le référentiel monde et obtient de meilleures performances que CoTracker3 [Karaev 2024a] et DELTA [Ngo 2024].

Tracking Any Point as Next token prediction (TAPNext) en Anglais [Zholus 2025] reformule le problème comme une tâche de prédiction de jetons masqués. Pour cela, l'architecture TRecViT [Pătrăucean 2024] est utilisée, elle est composée d'une alternance de couches temporelles State Space Model (SSMod) en Anglais et spatiales Vision Transformer (ViT) en Anglais. TAPNext est l'état de l'art en avril 2025 avec des performances temps réel.

Les modèles de fondation visuels (Vision Foundation Models (VFM) en Anglais) comme Masked AutoEncoders (MAEn) en Anglais [He 2022a], knowledge DIstillation with NO labels (DINO) en Anglais [Caron 2021], DINOv2 [Oquab 2023], DINOv2-Reg [Darcet 2023], Contrastive Language Image Pre-training (CLIP) en Anglais [Radford 2021], Segment Anything Model (SAM) en Anglais [Kirillov 2023] ou Data-efficient Image Transformer (DeIT) en Anglais [Touvron 2022], offrent également des perspectives prometteuses pour le suivi de points longue durée [Aydemir 2024]. DINOv2 est aujourd'hui un bon candidat pour le suivi visuel [Song 2024, Tumanyan 2024] avec des atouts dans la compréhension des scènes chirurgicales [Khan 2025].

Des alternatives combinant flux optique et suivi long terme ont aussi été proposées comme FlowTrack [Cho 2024], qui compense les dérives et occultations en chaînant les flux optiques.

LEAP-VO [Chen 2024b] est une nouvelle approche de l'odométrie visuelle (Visual Odometry (VO) en Anglais) qui introduit un module de suivi de points à long terme appelé Long-term Effective Any Point tracking (LEAP) en Anglais. Contrairement aux méthodes classiques centrées sur des correspondances entre paires d'images (approche "two-view"), LEAP exploite le contexte temporel complet pour suivre efficacement les points, même dans des scènes dynamiques, avec occultations ou faible texture. LEAP s'appuie sur un suivi de points combinant trois types d'informations : visuelles, inter-trajectoires et temporelles. Elle introduit une estimation probabiliste pour modéliser l'incertitude des trajectoires. Sur le plan technique les caractéristiques sont extraites via un CNN et raffinées itérativement. L'attention inter-trajectoire et l'ajustement de faisceaux sont également utilisés.

Applications cliniques Le problème de suivi concerne également d'autres modalités d'imagerie médicale, notamment l'échographie. Des méthodes spécifiques comme EchoTracker [Chernyshov 2025a, Azad 2024], inspirées de Tracking Any Point with per-frame Initialisation and temporal Refinement (TAPIR) en Anglais) [Doersch 2023], utilise une stratégie de renforcement itératif proposée dans Persistent Independent Particles (PIP) en Anglais [Harley 2022] pour le suivi du myocarde. Une version simplifiée de CoTracker2 [Karaev 2024b], MyoTracker [Chernyshov 2025b], réduit drastiquement le temps d'inférence à 23 millisecondes au lieu de 244 millisecondes pour EchoTracker [Azad 2024] et 1 692 millisecondes pour CoTracker2 [Karaev 2024b].

En endoscopie, Endo-TTAP [Zhou 2025] est une méthode de suivi tissulaire qui utilise le Multi-Facet Guided Attention (MFGA) en Anglais pour la fusion de six types de caractéristiques et intègre deux têtes légères, une pour l'incertitude l'autre

pour l'occultation. L'entraînement est réalisé en deux temps : 1) supervisé sur données synthétiques et 2) pseudo labels CoTracker3 [Karaev 2024a] sur données réelles.

Plus récemment (après nos travaux), une méthode exploite CoTracker [Karaev 2024b] pour le suivi de la prostate et permet la navigation sans marqueur dans une vidéo monoculaire. La méthode est initialisée par un recalage 3D-2D manuel entre le jumeau numérique et une IC. Un dessin virtuel sur la surface de la prostate est réalisé pour placer des points qui seront suivis par CoTracker. Le recalage 3D-2D est ensuite effectué classiquement par estimation de pose PnP RANSAC ou par Rendu différentiable (Differentiable Rendering (DR) en Anglais) [Kato 2020, Zeng 2025]. L'erreur de suivi rapportée varie entre 2,46 et 5,36 pixels selon les patients, soit une erreur de recalage TRE de 6,11 millimètres [Bender 2025].

Ensembles de données La littérature recense plusieurs types d'ensembles de données pour l'évaluation des méthodes de suivi dans le domaine médical. La majorité des enregistrements sont monoculaires ou stéréoscopiques, mais rarement labellisés [Mountney 2010, Stoyanov 2005b, Giannarou 2012, Ye 2017]. Certains ensembles de données incluent des semi-vérités terrain dérivées de méthodes de reconstruction tridimensionnelles comme le SLAM, le SfM ou la stéréovision [Azagra 2023], comme illustré en figure 5.3. D'autres sont labellisés manuellement par des annotateurs via des boîtes englobantes [Ye 2016, Li 2020b, Cartucho 2024] ou l'utilisation de marqueurs visibles/infrarouges [Lin 2023c, Schmidt 2024b].

Des données d'évaluation ex-vivo peuvent être capturées via des scanners tridimensionnels [Ozyoruk 2021, Allan 2021] et/ou complétées par des capteurs tiers comme les centrales inertielles (Inertial Measurement Unit (IMU) en Anglais) [Hartwig 2022].

Les données simulées, générées par rendu graphique, simulation et/ou adaptation de domaine constituent une ressource importante [Visentini-Scarzanella 2017, Rau 2019, Zhang 2021c, Zhang 2020a, Cartucho 2021, Guy 2022, Bobrow 2023, Azagra 2023], comme illustré en figures 5.4 et 5.5. Enfin des ensembles de données issus de phantoms (modèles physiques imprimés par fabrication additive ou sculptés) sont fréquemment utilisés pour la validation [Stoyanov 2010, Pratt 2010, Edwards 2022].

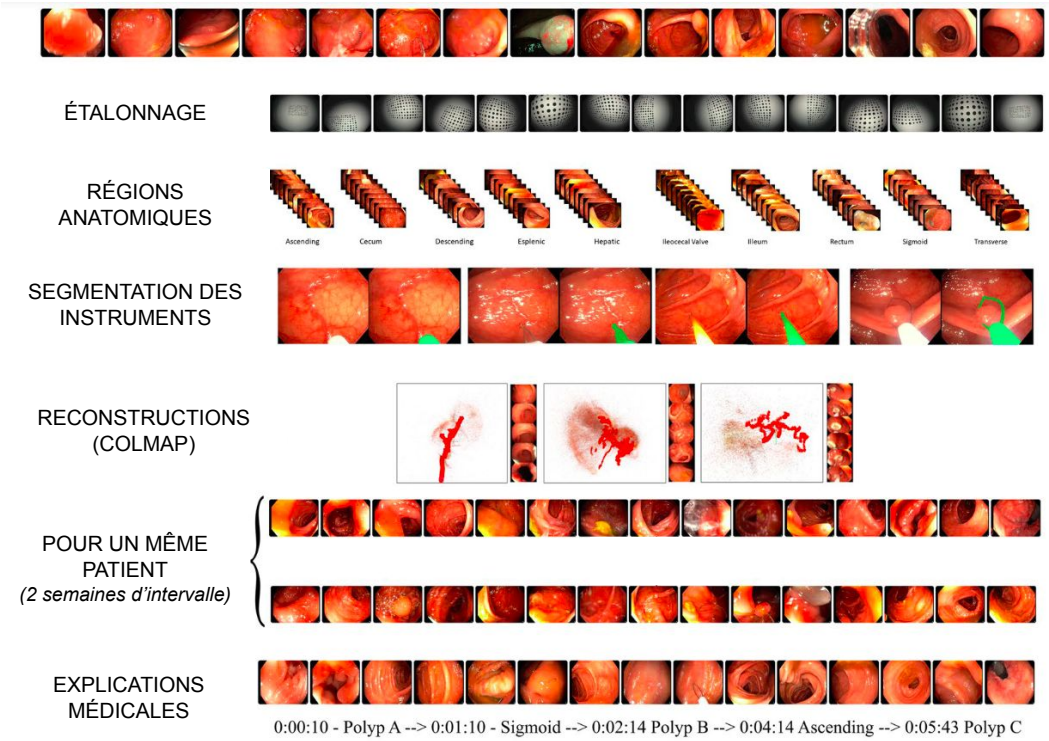


FIGURE 5.3 – Vue globale de l'ensemble de données EndoMapper. Source : générée par l'auteur à partir de [Azagra 2023].

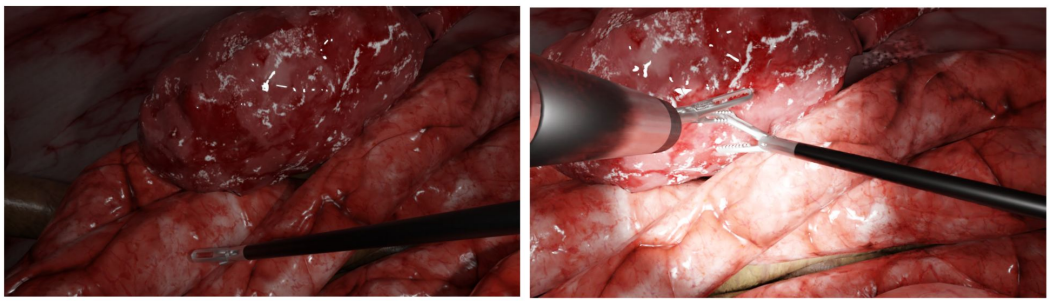


FIGURE 5.4 – Exemples d'images simulées dans un environnement virtuel conçu sous Blender. Source : générée par l'auteur à partir de [Guy 2022].

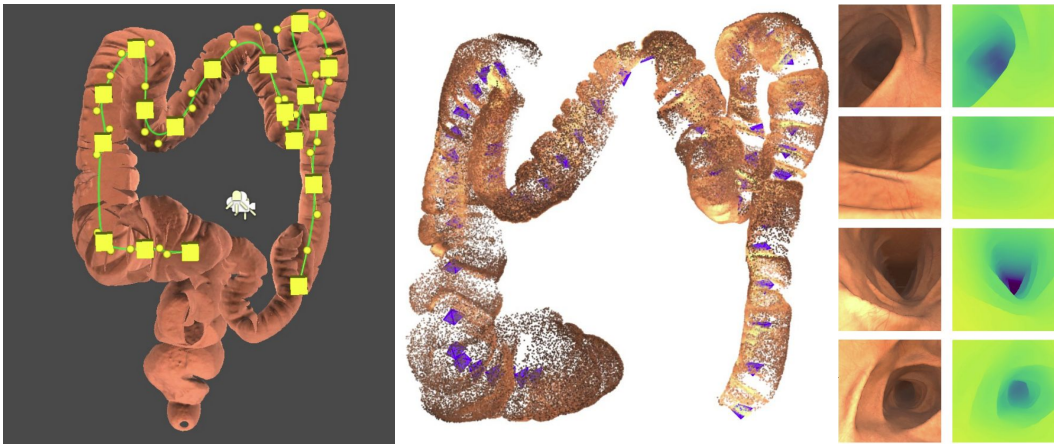


FIGURE 5.5 – Exemples d'images et de trajectoires de caméra simulées dans un environnement virtuel conçu sous Unity 3D. Source : générée par l'auteur à partir de [Rau 2023].

5.3.3 Méthodes de réduction de la dérive en suivi visuel

La dérive temporelle en suivi visuel désigne l'accumulation progressive d'erreurs locales au fil des mises à jour successives du modèle d'apparence ou de géométrie. Ces erreurs, souvent imperceptibles à court terme, conduisent à un glissement progressif de la pose estimée vers des régions erronées, qui entraîne une dégradation graduelle des performances du suivi au cours du temps. Ce phénomène, intrinsèquement lié aux approches itératives, constitue l'un des principaux défis rencontrés par les systèmes de suivi à long terme. Dans le contexte du suivi par template, il a été montré qu'une mise à jour naïve du modèle induit inévitablement une dérive cumulative [Matthews 2004]. De manière plus générale, plusieurs études expérimentales ont mis en évidence que la dérive du modèle demeure l'un des principaux facteurs d'échec dans les suivis de longue durée et en environnements variés [Smeulders 2013]. Ces constats ont conduit au développement de nombreuses stratégies destinées à limiter la dérive, via la combinaison d'un suivi local robuste, d'une correction périodique et d'un contrôle adaptatif des mises à jour.

Les approches dites de suivi à long terme articulent généralement un suivi local précis avec un mécanisme de re-détection global destiné à corriger les pertes de cible ou les décalages progressifs. L'approche Long-term Correlation Tracking (LCT) en Anglais [Ma 2015] illustre ce principe en décomposant le problème en une estimation conjointe de la translation et de l'échelle, combinant un filtre de corrélation pour le suivi à court terme et un classifieur en ligne pour la re-détection en cas de perte. Ce cadre exploite la corrélation temporelle du contexte visuel afin d'accroître la robustesse face aux occultations et aux variations d'apparence. Dans une perspective plus générale, certains auteurs proposent une structuration tripartite du suivi (suivi, vérification, re-détection), qui met en évidence le rôle central des politiques de mise

à jour dans la prévention de l'apprentissage des erreurs cumulatives [Liu 2022a]. Dans la même logique, le modèle Combining Complementary trackers in Long-term visual Tracking (CoCoLoT) en Anglais, associe plusieurs algorithmes de suivi complémentaires à un module de vérification dynamique, afin d'ajuster la stratégie de suivi en fonction de la fiabilité instantanée [Dunnhofer 2022].

D'autres travaux se concentrent sur le contrôle de la fréquence d'adaptation du modèle. Le système Universal Update-Pacing Framework (UUPF) en Anglais [Hu 2016] repose sur l'utilisation simultanée de plusieurs algorithmes de suivi opérant à des rythmes distincts. En comparant leurs trajectoires avant-arrière, le comportement le plus cohérent est sélectionné, limitant ainsi la dérive engendrée par des mises à jour prématurées ou inexactes.

Afin d'atténuer la contamination du modèle par des observations aberrantes, certaines approches introduisent la notion de "souvenirs fiables" (reliable memories en Anglais). Elle consiste à conserver dans une mémoire sélective des représentations stables extraites d'images antérieures. Ces ancrages fiables servent de référence lors des mises à jour, réduisant ainsi la propagation des erreurs cumulatives [Wang 2016a]. D'autres méthodes renforcent la stabilité du modèle à l'aide de représentations parcimonieuses. Le suivi visuel a ainsi été formulé comme un problème d'approximation parcimonieuse sous contrainte ℓ_1, avec comme représentation du candidat une combinaison linéaire de templates précédents, afin de gérer naturellement les occultations et le bruit [Mei 2011]. Cette idée a ensuite été généralisée dans un cadre multitâche structuré exploitant la corrélation entre particules afin d'éviter la dérive [Zhang 2013]. Ces approches fondées sur le codage épars partagent une même philosophie : contrôler la mise à jour du modèle en ne retenant que les composantes temporellement stables et cohérentes.

Dans les méthodes à filtres de corrélation, la dérive résulte fréquemment d'un surajustement local ou d'une dégradation progressive du filtre. Pour pallier ce problème, plusieurs travaux intègrent des contraintes de régularisation temporelle adaptative ou encore des modélisations contextuelles spatio-temporelles visant à stabiliser le modèle [Zhang 2015]. Une méthode, dénommée adaptive sparse spatial-temporal context-aware en Anglais, exploite des contraintes structurées pour réduire la dérive du modèle [Su 2021]. En complément, certains filtres de corrélation sont enrichis d'un mécanisme de rétroaction entre l'extraction de la cible et le processus de suivi, qui permet une adaptation dynamique de la taille et une correction progressive des décalages. Ce couplage itératif améliore la précision de localisation tout en maintenant un traitement en temps réel [Tunali 2018]. Par ailleurs, la stabilisation préalable des séquences vidéo peut réduire significativement la dérive de suivi en atténuant les erreurs d'estimation dues aux vibrations ou aux mouvements parasites de la caméra [Verma 2018].

La logique de re-détection hiérarchique a été prolongée par des approches qui intègrent une évaluation spatio-temporelle de la fiabilité et une stratégie de re-détection multi-régions pour corriger les pertes dans les filtres de corrélation classiques [Li 2024b]. Enfin, dans le domaine de l'odométrie visuelle, les stratégies de fusion multi-images [Badino 2013] offrent des pistes complémentaires pour atténuer la dérive en exploitant la redondance des observations.

L'émergence de réseaux neuronaux à mémoire a également favorisé le développement d'architectures capables de détecter et de compenser la dérive de manière autonome. Le réseau Drift Determination Network (DDN) en Anglais, fondé sur un double modèle statique et dynamique, évalue la fiabilité des prédictions sans recourir à des seuils explicites et combine les deux modèles afin de corriger les dérives détectées [Hou 2023]. Dans un contexte différent, notamment celui des vidéos satellitaires, un module anti-dérive basé sur des contraintes de mouvement adaptatives et sur un modèle de trajectoire appris par réseau Long Short-Term Memory (LSTM) en Anglais permet de distinguer les dérives visuelles des variations dynamiques réelles du mouvement [Fan 2024].

5.3.4 Correction des distorsions de la perspective

Plusieurs travaux ont été consacrés à la suppression des distorsions de perspective avant l'extraction des caractéristiques [Koser 2007, Cao 2009, Zeisl 2013]. Une méthode récente [Toft 2020] exploite une carte de profondeur générée par un réseau de neurones de type mono-depth afin de rétroprojeter les pixels dans l'espace tridimensionnel et regrouper les normales de surface selon trois directions dominantes, afin d'identifier les régions planes. Pour chacune des régions, une homographie est calculée afin de rectifier localement la perspective, puis les points clés sont détectés sur l'image rectifiée.

Une approche similaire a été proposée dans [Zeisl 2012] pour corriger la distorsion de perspective dans le cas d'objets aux surfaces développables. Après détection et paramétrisation des surfaces développables dans la carte de profondeur, celles-ci sont dépliées en plans et texturées. Les points clés sont alors extraits à partir des images dépliées.

Cependant, l'état de l'art en matière de correction des distorsions de perspective demeure difficilement applicable aux surfaces complexes, notamment dans le contexte chirurgical. Les structures anatomiques comme le rein présentent des formes courbes, non développables et non lisses (cf. section 2.1.1), ce qui limite fortement l'efficacité des approches existantes.

Deux stratégies principales permettent d'aborder les changements de perspective. La première repose sur l'utilisation de descripteurs invariants, efficaces face aux transformations de rotation et de mise à l'échelle mais encore insuffisants pour la compensation de la perspective. La seconde consiste à appliquer une correction géométrique préalable par déformation (warping en Anglais) des images avant la détection. Les méthodes classiques recourent généralement à des homographies, mais elles reposent sur l'hypothèse de planéité de la scène, ce qui est irréaliste dans le

cas d'organes. De plus, certaines exigences supplémentaires, comme la connaissance préalable de la géométrie ou la possibilité de déplier la surface, ne sont pas réunies en contexte chirurgical.

Dans ce cadre applicatif, la méthode Surgical Image Perspective Correction (SurgIPC) en Anglais [Sharifian 2025a] a été développée. Elle utilise en entrée l'image chirurgicale ainsi qu'une carte de profondeur afin de reconstruire un maillage tridimensionnel peropératoire de l'organe. Un warp de l'image est ensuite calculé en minimisant le rééchantillonnage final grâce à une fonction de coût tridimensionnelle conforme qui généralise le Least-Squares Conformal Mapping (LSCM) en Anglais [Lévy 2023, Mullen 2008] et via un coût de déplacement d'image. L'optimisation du coût global est réalisée par une méthode de type Linear Least Square. Cette approche, dont le temps d'exécution est de 120 millisecondes indépendamment du contenu, permet des gains significatifs en termes de correspondances établies, avec par exemple pour SuperGlue une amélioration de 66 % sur des images synthétiques.

5.4 Segmentation automatique de l'organe dans les images peropératoires

Dans le cadre du suivi visuel d'organes par détection via l'utilisation de points clé, l'évaluation de l'appartenance d'un point clé à la surface de l'organe cible est une notion essentielle à prendre en compte. Ce critère est indispensable pour fiabiliser et stabiliser le suivi automatique, puisqu'il permet de partitionner l'ensemble des points en segments, généralement associés à une entité physique comme un organe ou un instrument chirurgical. Dans nos travaux, le critère d'appartenance est estimé via les segments issus de la segmentation sémantique des images peropératoires.

5.4.1 Généralités

Au cours d'une intervention chirurgicale, la scène observée par la caméra endoscopique est dynamique. Cette variabilité est principalement induite par les phénomènes physiologiques tels que les battements cardiaques, les cycles respiratoires, la circulation sanguine dans les vaisseaux et les organes, ainsi que le flux urinaire en transite par les uretères.

Par ailleurs, les gestes chirurgicaux mobilisent les tissus, entraînant à la fois des déplacements globaux des structures entre elles et des déplacements locaux qui traduisent des déformations.

Les méthodes de suivi visuel par détection que nous avons mise en œuvre dans le cadre de la chirurgie gynécologique et urologique, utilisent l'appariement d'images via la mise en correspondance de descripteurs calculés à des points clés (cf. section 5.5). Ces correspondances servent d'entrées au calcul de la pose de la caméra par projection PnP. Comme l'illustre la figure 5.6, après l'étape de détection, les points clés se distribuent sur l'intégralité de l'image. En l'absence de filtrage spatial et selon les mouvements conjoints de la caméra et des structures anatomiques, plu-

sieurs trajectoires émergent sous la forme de faisceaux de correspondances. Ceux-ci reflètent macroscopiquement le déplacement relatif des différentes structures d'une image à l'autre. Dans ce contexte, le calcul de pose ne peut converger vers une solution stable qu'à condition de sélectionner le faisceau lié à l'organe cible.

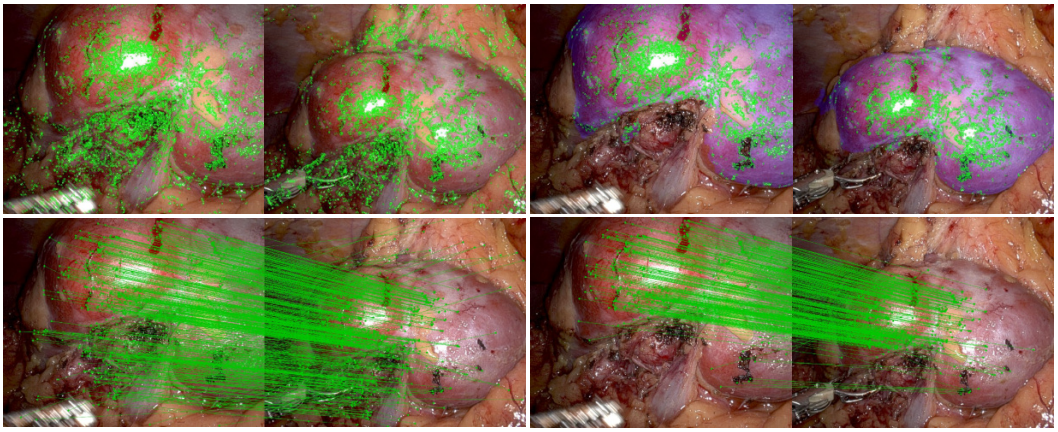


FIGURE 5.6 – Illustration en Néphrectomie Partielle Robot-Assistée (NPRA), de l'intérêt du filtrage spatial par segmentation sémantique de l'organe cible. (à gauche) Mise en correspondance de descripteurs SIFT sans segmentation. (à droite) Mise en correspondance de descripteurs SIFT avec segmentation automatique de la surface du parenchyme rénal (en bleu). Source : auteur.

La segmentation sémantique automatique de la scène peropératoire constitue alors une solution pertinente. Elle est appliquée sur chaque IC lors de son ajout à la BIC, afin de générer un masque utilisé comme filtre spatial. Ce dernier permet d'isoler les tissus ayant un mouvement global cohérent et d'éliminer les éléments perturbateurs pour le calcul de pose, comme les structures anatomiques adjacentes ou les instruments chirurgicaux, fréquemment situés au premier plan en occultation avec l'organe cible. À cet effet, un réseau de neurones dédié à la segmentation est entraîné sur des données issues de la pratique clinique.

5.4.2 Collecte des données

Dans le cadre de la NPRA, application clinique principale de ce travail, nous avons constitué en collaboration avec nos partenaires cliniques deux ensembles de données brutes : DataSet Patient Test (DSPT) et DataSet Patient (DSP). Cette collecte prospective s'est déroulée entre février 2022 et mai 2025 dans trois CHU, Bordeaux, Caen et Angers, du réseau français de recherche sur le cancer du rein, UroCCR. Les principales caractéristiques clinico-techniques sont synthétisées dans le tableau 5.1.

DSPT est issu des sessions d'essais de faisabilité de la RA (cf. section 2.2.3). Il se distingue de DSP par la disponibilité systématique du flux stéréoscopique, des données d'étalonnage et du jumeau numérique.

	DSPT	DSP
Nb. de chirurgies (NPRA)	28	203
Latéralité NPRA	13 Gauches et 15 Droites	104 Gauches et 99 Droites
Genre du patient	20 Hommes et 8 Femmes	126 Hommes et 77 Femmes
Flux vidéo	Stéréoscopique	Monoculaire
Nb. d'heures	Plus de 67	Plus de 500
Nb. d'images	Plus de 14 000 000	Plus de 100 000 000
Poids en Go	Plus de 1 400	Plus de 6 300

TABLE 5.1 – Caractéristiques clinico-techniques principales des deux ensembles de données brutes DataSet Patient Test (DSPT) et DataSet Patient (DSP), issu de l'enregistrement prospectif de Néphrectomies Partielles Robot-Assistées (NPRA) dans trois Centres Hospitaliers Universitaires (CHU) (Bordeaux, Caen et Angers) du réseau français de recherche sur le cancer du rein UroCCR. Source : auteur.

Sur le plan technique, les données des ensembles DSPT et DSP proviennent de la capture du flux brut en sortie du processeur vidéo du système de télémanipulation chirurgicale Da Vinci Xi (Intuitive Surgical) (cf. section 2.2.5). Le flux est enregistré avec une résolution de 1280×1024 pixels et une fréquence moyenne de 60 Hz. Il provient ici d'un endoscope avec une optique de 30°. Deux systèmes d'acquisition ont été utilisés : le dispositif commercial MVR Pro (Medicapture) et un dispositif personnalisé basé sur un ordinateur médical sous Ubuntu, équipé d'une carte d'acquisition vidéo HDMI Pro Capture Quad (Magewell), illustrés en figure 5.7. Ces systèmes, installés sur la colonne du télémanipulateur chirurgical, permettent également la saisie de méta-informations utiles à l'indexation des données.



FIGURE 5.7 – (à gauche) Enregistreur MVR Pro (Medicapture). (à droite) Enregistreur personnalisé basé sur un ordinateur médical sous Ubuntu équipé d'une carte d'acquisition vidéo HDMI Pro Capture Quad (Magewell). Source : auteur.

Les différences principales entre les deux systèmes concernent la compacité (Medicapture MVR Pro : $173 \times 153 \times 50$ millimètres versus personnalisé : $420 \times 169 \times 409$ millimètres), le nombre d'entrées vidéo synchrones disponibles (aucune pour le Medicapture MVR Pro contre 2 à 4 pour le système personnalisé selon la configuration) et la fréquence d'échantillonnage (VFR pour le Medicapture MVR Pro et CFR pour l'enregistreur personnalisé).

Les paramètres d'acquisition vidéo sont résumés dans le tableau 5.2. Les deux systèmes ont été configurés de manière quasi-identique, à l'exception de la fréquence d'échantillonnage, afin d'assurer la reproductibilité inter-centres. Ces paramètres sont par ailleurs identiques à ceux utilisés pour l'inférence en peropératoire (système de RA basé sur un matériel similaire à l'enregistreur personnalisé), limitant ainsi l'écart entre les domaines d'entraînement et d'inférence. Afin de faciliter la gestion, les vidéos sont enregistrées sous la forme de fragments de taille fixe.

A réception des données brutes, un contrôle systématique de l'anonymisation ou de la pseudonymisation est effectué selon le protocole en vigueur. Les données sont ensuite intégrées au registre des données et étiquetées d'un niveau de complexité visuelle (faible, moyenne ou élevée), critère subjectif qui permet d'anticiper la difficulté de la labellisation selon la clarté de la scène chirurgicale (exemples : saignements limités, organes bien délimités, etc.).

Paramètre	Valeur
Résolution	1280 x 1024 pixels
Ratio d'aspect	5:4
Fréquence	60 Hz
Conteneur	MP4
Type de balayage	Progressif
Codec vidéo	AVC / H.264 (High@L5.2, 1 Ref Frame)
Débit vidéo	20 Mb/s constant (CBR)
Norme de couleurs	BT.709
Espace colorimétrique	YUV 4:2:0 (8 bits)

TABLE 5.2 – Paramètres principaux d'acquisition vidéo des systèmes d'enregistrement utilisés. Source : auteur.

5.4.3 Labellisation des données

Les données vidéo brutes collectées sont échantillonnées en images destinées à la labellisation. L'objectif est d'optimiser la diversité de la distribution tout en minimisant la redondance. Pour cela, une méthode semi-automatique de sélection est utilisée. Elle repose sur trois critères : 1) la distance temporelle avec la dernière image extraite (supérieure à cinq secondes), 2) la netteté estimée par la variance du laplacien (supérieure à 0,07) et 3) la similarité visuelle évaluée par la distance entre les hashs de Marr-Hildreth (supérieure à 200). Le pseudo-code correspondant est donné dans l'algorithme 5. La sélection est ensuite validée manuellement.

Algorithme 5 Sélection automatique d'images à labelliser dans des vidéos chirurgicales

```
tant que parcours des images de la vidéo faire
    si écart temporel avec la dernière image sélectionnée < seuil alors
        ignorer l'image
    si netteté de l'image < seuil alors
        ignorer l'image
    si similarité avec la dernière image sélectionnée < seuil alors
        ignorer l'image
    sauvegarder l'image comme nouvelle image à labelliser
```

Sur le plan méthodologique, les données labellisées utilisées pour l'entraînement, issues de vidéos de NPRA, ont été obtenues après une série d'étapes représentées schématiquement en figure 5.8. Un des prérequis fondamentaux est la définition de règles de labellisation élaborées conjointement par les équipes clinique et technique. Celles-ci spécifient avec précision la manière de labelliser les structures d'intérêt

et d'aborder les situations complexes, à l'aide d'exemples illustrés. Ces règles sont consignées dans un guide de labellisation dont des extraits sont donnés en figure 5.9. Elles évoluent en fonction du besoin et du retour d'expérience. Dans le cas du projet sur la chirurgie du rein, cinq versions successives de règles de segmentation panoptique se sont succédées entre avril 2022 et mai 2024, pour aboutir à la définition de cinq classes labellisées : la surface du parenchyme rénal, la graisse périrénale persistante sous forme d'îlots, la surface visible des tumeurs exophytiques ou des kystes, la zone de coagulation qui apparaît au cours de l'intervention sous l'effet d'un instrument d'électrochirurgie (souvent une pince bipolaire) et l'instrumentation chirurgicale dans son ensemble (outils, compresses, clips, etc.). À ces classes s'ajoutent trois tags colorimétriques, dont l'un est destiné à qualifier l'apparence de la tumeur ou du kyste et deux relatifs à la surface du parenchyme rénal. Ces informations complémentaires permettent respectivement d'apprécier la nature de la tumeur ou du kyste et d'estimer l'état de la vascularisation du rein. Un résumé illustré de ces éléments est présenté en figure 5.10.

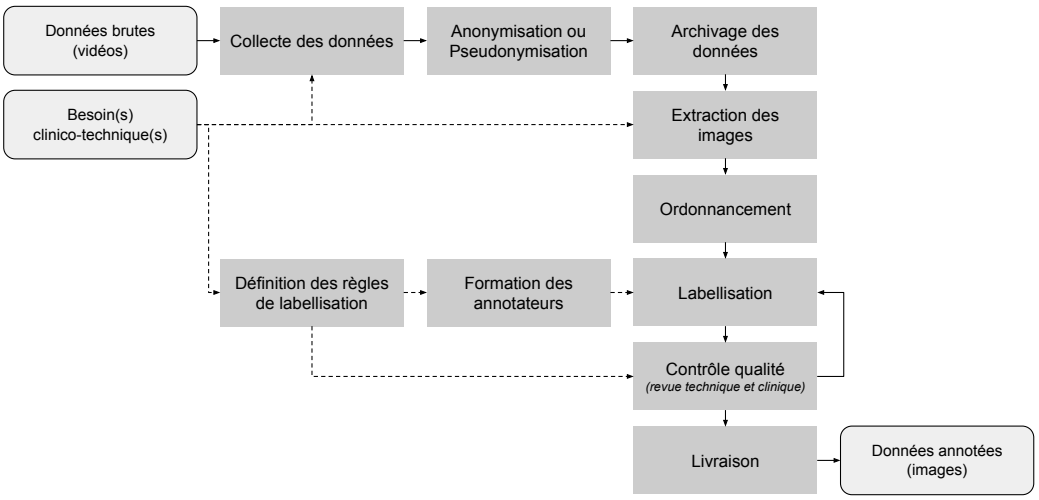


FIGURE 5.8 – Séquence d'étapes d'un processus industriel de labellisation qui permet d'obtenir à partir de données brutes des données labellisées qualitatives. Source : auteur.

Le flux de labellisation est conçu comme un processus industriel standardisé, organisé par lots et incluant des phases de montée en charge, de la présérie jusqu'à la série. Un principe d'amélioration continue est mis en œuvre afin d'affiner itérativement les méthodologies de labellisation (instructions et outils). Cette démarche vise d'une part à rendre le processus scalable à une échelle industrielle via l'utilisation de sous-traitance non experte, formée et évaluée, et d'autre part à adapter la labellisation aux capacités discriminantes du réseau de neurones. Il répond également à des impératifs de rendement et de maîtrise des coûts de production.

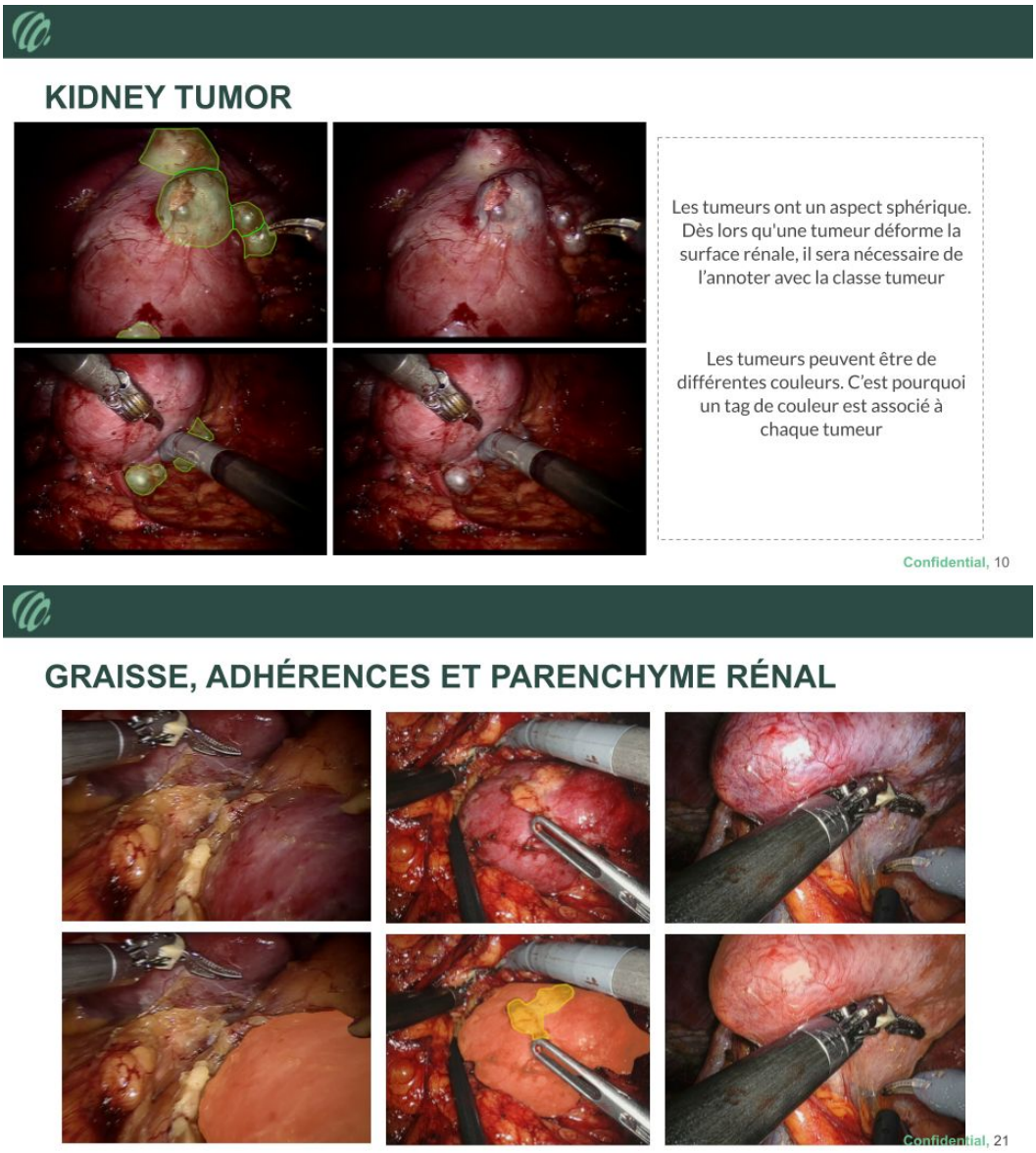


FIGURE 5.9 – Extraits du guide de labellisation élaboré pour la segmentation panoptique en Néphrectomie Partielle Robot Assistée (NPRA). Source : auteur.

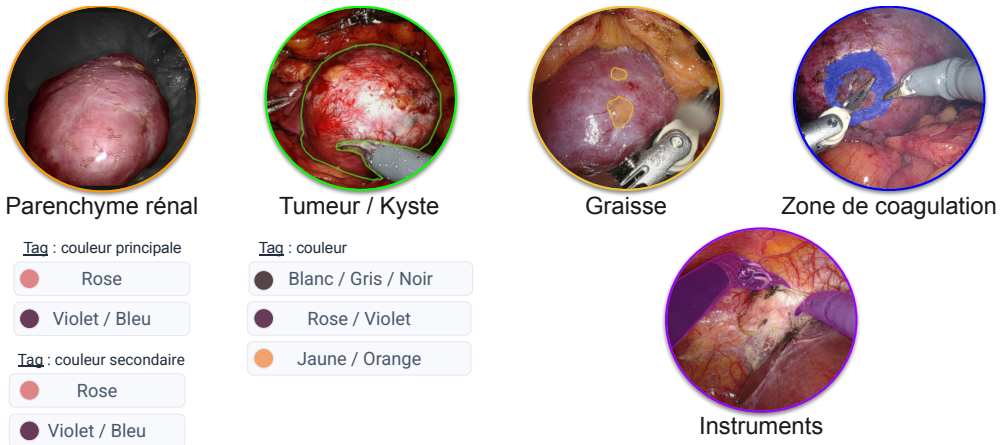


FIGURE 5.10 – Résumé schématique des éléments labellisés dans le cadre de la segmentation panoptique en Néphrectomie Partielle Robot Assistée (NPRA). Source : auteur.

D'un point de vue opérationnel, un outil collaboratif de labellisation est utilisé pour piloter le flux de production des données labellisées. Il centralise l'ensemble des fonctions de gestion, comme l'attribution des tâches de labellisation aux annotateurs, le suivi de la performance, la mise à disposition d'interfaces et d'outils de labellisation comme illustré en figure 5.11, ainsi que la gestion des différents états des données (à labelliser, labellisées, en cours de contrôle et terminées). Dans ce projet, l'application Supervisely Enterprise Edition [Supervisely Team 2025] est utilisée. L'outil de labellisation a lui-même évolué avec l'état de l'art. Avant 2023, la labellisation était réalisée manuellement ou à l'aide d'outils interactifs dédiés (Smart Tool dans Supervisely) entraînés spécifiquement pour la tâche. Les résultats n'étaient pas convaincants, une labellisation manuelle restait plus efficace en termes de temps. Après 2023, l'introduction de méthodes de segmentation interactives génériques telles que SAM [Kirillov 2023], Segment Anything Model in High Quality (SAM-HQ) en Anglais [Ke 2023] et SAM 2 [Ravi 2024] a profondément modifié cet équilibre, en permettant une accélération notable de la labellisation pour des objets aux contours bien délimités.



FIGURE 5.11 – Application Supervisely Enterprise Edition utilisée pour la labellisation des données. Source : auteur.

Tout comme un flux de production industriel, la qualité est assurée par un ensemble de contrôles réalisés tout au long de la chaîne. Des contrôles aléatoires sont effectués en ligne, complétés par des contrôles par échantillonnage ou à 100 % en bout de chaîne sur les lots produits. Trois niveaux de contrôles sont mis en place : les revues par l'équipe de production (vérification croisée inter-opérateurs et par le pilote de production), les revues techniques (vérification de l'application des règles de labellisation et de la cohérence temporelle des données labellisées, elles permettent de détecter d'éventuelles incohérences cliniques), et enfin les revues cliniques, qui valident la pertinence médicale des données à travers une vérification aléatoire. Cette stratification garantit une couverture complète des sources potentielles d'erreurs et maximise la qualité des données produites. A l'issue de ces étapes de contrôle, une décision est prise pour chaque lot : acceptation, dérogation ou rebut avec relabellisation.

À partir des 52 chirurgies sélectionnées parmi les données collectées, 9 issues de DSPT et 43 issues de DSP (cf. section 5.4.2), soit un total de 110 heures de vidéos enregistrées, 48 140 images (6 980 + 41 160) ont été extraites et mises à disposition pour la labellisation. Les étapes de préparation des données ont représenté environ une centaine d'heures de travail. L'intégralité des images a été labellisée par l'équipe de production, composée de trois annotateurs simultanés au pic d'activité. Entre janvier 2023 et juillet 2025, le temps cumulé de labellisation (temps de réflexion et de réalisation) est estimé à 11 723 heures, soit une moyenne d'environ 15 minutes par image pour l'ensemble des cinq classes et trois tags.

Les revues technique et clinique des données labellisées sont actuellement en cours. L'expérience montre qu'une double revue technique constitue un levier essentiel pour fiabiliser la qualité de la donnée labellisée, par correction des erreurs résiduelles issues des contrôles aléatoires et via l'établissement d'un consensus sur les cas complexes.

5.4.4 Entraînement du réseau de neurones

L'architecture typique des réseaux de neurones artificiels dédiés aux tâches de segmentation repose généralement sur deux modules principaux. Le premier, appelé module d'encodage, extrait à partir de l'image d'entrée des cartes de caractéristiques par un processus de sous-échantillonnage, projetant ainsi l'information dans un espace latent de plus faibles dimensions. Le second, le module de décodage, a pour objectif de réinjecter ces représentations dans l'espace initial par suréchantillonnage des caractéristiques.

Parmi les architectures de références, le UNet [Ronneberger 2015] occupe une place prépondérante. Il permet de pallier aux problèmes des opérations de suréchantillonnage éparses en exploitant, à travers les connexions résiduelles (skip connections en Anglais), les cartes de caractéristiques intermédiaires de l'encodeur de dimensions compatibles. Cette stratégie permet de préserver l'information contextuelle tout en affinant la localisation spatiale des structures segmentées.

Dans le cadre de nos travaux, l'architecture implémentée est une variante du UNet, représentée en figure 5.12. Son encodeur est un EfficientNet-B3 [Tan 2019b], pré-entraîné sur l'ensemble de données ImageNet [Deng 2009]. Ce choix offre la capacité d'extraire, à partir d'images endoscopiques standardisées RVB de résolution 512×512 pixels, des cartes de caractéristiques de taille $(16, 16, 384)$. L'intérêt d'EfficientNet réside dans sa stratégie de "scaling compound", qui met simultanément à l'échelle la profondeur, la largeur et la résolution du réseau à l'aide d'un coefficient unique, garantissant un bon compromis entre précision et efficacité computationnelle. De plus, cet encodeur repose sur des blocs Mobile inverted Bottleneck (MBConv) en Anglais [Sandler 2018, Tan 2019a] enrichis par des modules squeeze-and-excitation [Hu 2018], qui améliorent la capacité de représentation tout en limitant le nombre de paramètres. À titre d'exemple, EfficientNet-B3 atteint 81,6 % de précision top-1 sur ImageNet avec seulement 12 millions de paramètres et 1,8 billion d'opérations en virgule flottante par seconde (FLoating-point Operations Per Second (FLOPS) en Anglais), surpassant largement ResNet-34 [He 2016] (73,3 % top-1, 21 millions de paramètres, 3,6 billions de FLOPS) tout en étant presque deux fois plus compact et deux fois plus efficace en calcul [Tan 2019b]. Le module de décodage et les connexions résiduelles sont ajustés en nombre et en dimensions pour assurer leur compatibilité avec cet encodeur.

Pendant l'entraînement, les données labellisées sont exploitées sans recours à des techniques d'augmentation. Seules des opérations de prétraitement sont appliquées, incluant une standardisation selon une moyenne fixe $[0, 4850, 4560, 406]$ et un écart-type $[0, 2290, 2240, 225]$ ainsi qu'un redimensionnement à 512×512 pixels.

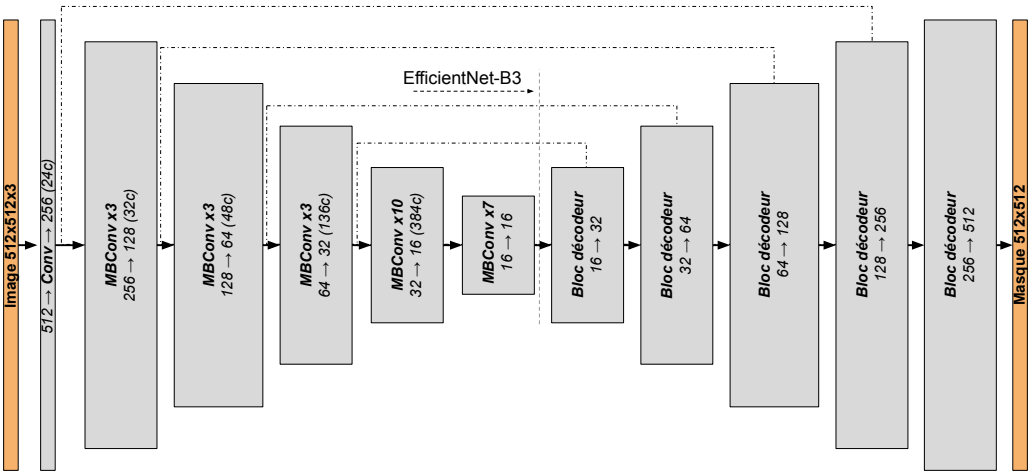


FIGURE 5.12 – Représentation schématique du réseau de neurones de segmentation de type UNet basé sur un encodeur EfficientNet-B3, qui d'une image d'entrée ($512 \times 512 \times 3$) permet d'inférer un masque (512×512). Source : auteur.

Afin d'améliorer les performances, les classes associées à la surface du parenchyme rénal ainsi qu'à la graisse périrénale persistante sous forme d'îlots sont fusionnées en une entité unique désignée sous le terme de "tissu rénal". Cette décision de nature empirique découle des observations rapportées en section 5.11.1 et s'explique également par la forte variabilité de labellisation de la graisse périrénale persistante. Par ailleurs, une classe supplémentaire intitulée "arrière-plan" est ajoutée.

La fonction de perte (loss en Anglais) utilisée pour l'apprentissage est une combinaison pondérée définie par :

$$\mathcal{L} = 0,5 \times \mathcal{L}_{CE} + \mathcal{L}_{Dice} + \mathcal{L}_{Focal} \tag{5.1}$$

Dans cette formulation, chaque terme joue un rôle spécifique dans la stabilisation et l'efficacité de l'apprentissage.

La première composante $\mathcal{L}_{CE}$, correspond à la fonction d'entropie croisée (Cross Entropy (CE) en Anglais) multi-classe, pour C classes, s'écrit :

$$\mathcal{L}_{CE} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log(p_{i,c}) \tag{5.2}$$

où $y_{i,c} \in \{0, 1\}$ correspond à l'étiquette de vérité terrain indiquant si le pixel i appartient à la classe c et $p_{i,c} \in [0, 1]$ représente la probabilité prédite par le réseau que i appartienne à cette même classe. Cette fonction mesure la divergence entre les distributions réelles et prédites, favorisant l'ajustement des probabilités estimées vers les étiquettes correctes.

La deuxième composante $\mathcal{L}_{Dice}$, repose sur l'indice de Dice couramment utilisé en segmentation médicale afin d'évaluer le chevauchement entre la prédiction et le masque de référence [Milletari 2016]. La formulation de la perte est la suivante :

$$\mathcal{L}_{Dice} = 1 - \frac{2\sum_{i=1}^{N} y_i p_i + \varepsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i + \varepsilon} \tag{5.3}$$

où ε est un terme de régularisation introduit pour éviter une division par zéro. Cette mesure est particulièrement adaptée aux données médicales où l'équilibre entre classes est rarement respecté, car elle pondère directement le recouvrement relatif des régions.

Enfin, la troisième composante $\mathcal{L}_{Focal}$, est dérivée de la fonction de perte focale [Lin 2017a], introduite pour remédier à la forte disparité de distribution entre les classes des ensembles de données déséquilibrés. Sa formulation est :

$$\mathcal{L}_{Focal} = -\frac{1}{N}\sum_{i=1}^{N} \alpha\,(1-p_i)^{\gamma}\, y_i \log(p_i) \tag{5.4}$$

où $\alpha \in [0,1]$ permet de pondérer l'importance relative de la classe minoritaire et $\gamma > 0$ est un paramètre de focalisation qui réduit la contribution des échantillons correctement classés au profit des échantillons difficiles.

Le réseau de neurones est entraîné sur 50 itérations (epochs en Anglais) avec une taille de lot (batch size en Anglais) de 8. L'optimisation est assurée par l'algorithme AdamW [Loshchilov 2017], une variante de l'optimiseur Adam, qui découple la décroissance du poids de la mise à jour du gradient. Le taux d'apprentissage (learning rate en Anglais) initial est fixé à 10^{-3}. Un mécanisme d'arrêt précoce (early stopping en Anglais) est également mis en œuvre afin de régulariser l'apprentissage et de prévenir le sur-apprentissage. La surveillance du processus d'entraînement est continue via l'estimation des métriques d'IoU et de F-score sur l'ensemble de données de validation.

5.4.5 Inférence

Lors de l'inférence, l'ensemble des 13 159 613 paramètres du réseau de neurones est gelé, ce qui implique que le réseau opère en mode strictement déterministe, sans mise à jour des poids. La prédiction du masque de segmentation est alors effectuée à partir d'une image endoscopique prétraitée, c'est-à-dire standardisée (moyenne : $[0,4850, 4560, 406]$ et écart-type : $[0,2290, 2240, 225]$) et redimensionnée au format 512×512 pixels.

En production, des stratégies d'optimisation du réseau de neurones sont employées comme l'élagage de connaissance (pruning en Anglais) et la quantisation afin de réduire la complexité du réseau, d'alléger sont empreinte mémoire et de diminuer son temps d'inférence.

Les poids des neurones d'un réseau sont classiquement représentés en précision flottante 32 bits, la quantisation consiste à convertir cette représentation numérique en format réduit 16 ou 8 bits. Cette conversion entraîne un gain direct sur

le temps d'inférence, de l'ordre de 1,5 à 4 fois plus rapide selon les cas. Deux approches principales sont couramment utilisées : la quantisation post-entraînement (Post Training Quantisation (PTQ) en Anglais), qui ne nécessite pas de réentraînement et présente un faible coût computationnel mais induit une dégradation des performances ; et la quantisation pendant l'entraînement (Quantisation-Aware Training (QAT) en Anglais), qui intègre la simulation de la quantisation pendant l'entraînement mais nécessite en contrepartie davantage de ressources de calcul [Rokh 2023, Gholami 2022, Nagel 2021, Wu 2020].

Le moteur d'inférence, chargé de l'exploitation du modèle via l'utilisation des ressources matérielles disponibles dans l'environnement d'exécution, doit être adapté au déploiement, car ce choix va influencer directement le temps d'inférence du réseau. Dans le cadre du réseau de segmentation proposé, les temps d'inférence ainsi que les performances sont donnés pour les différents moteurs, notamment Pytorch [Ansel 2024], ONNX Runtime [ONNX developers 2021] et TensorRT (Nvidia), en combinaison avec les optimisations par quantisation en FP32, FP16 et UINT8 en section 5.11.1.

5.5 Suivi monoculaire via une base d'images clés statique

La méthode de suivi visuel statique par détection constitue le point de départ de nos travaux. Elle met en œuvre le pipeline proposé par Collins et al. [Collins 2020], illustrée en figure 5.13, initialement conçu pour la chirurgie de l'utérus guidée par RA. Les principaux modules et adaptations sont détaillés dans cette section.

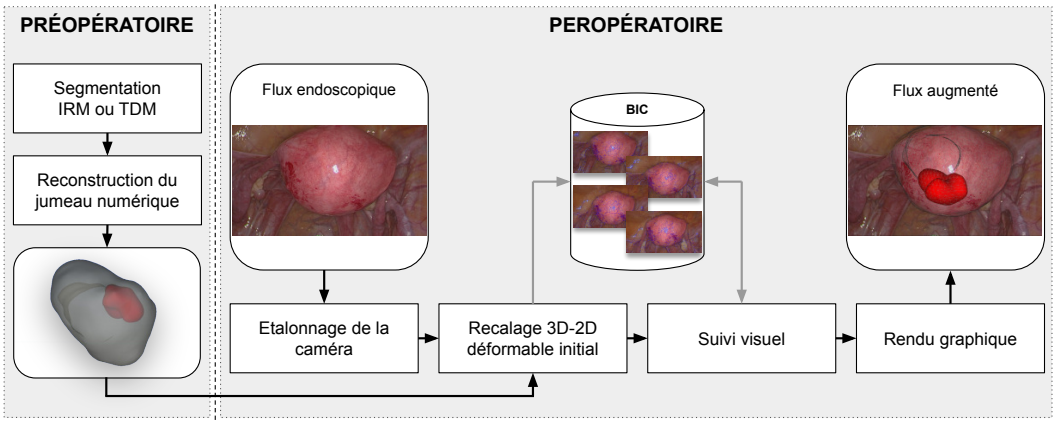


FIGURE 5.13 – Représentation schématique du pipeline proposé par Collins et al. en chirurgie de l'utérus guidée par Réalité Augmentée (RA). Il comporte trois modules peropératoires essentiels : 1) l'étalonnage de la caméra, 2) le recalage 3D-2D déformable initial et 3) le suivi visuel basé sur une Base d'Images Clés (BIC) statique. Source : auteur.

Avec cette approche, un flux vidéo monoculaire haute définition (1920 × 1080 pixels) à 60 Hz issu d'une caméra endoscopique est traité en quasi temps réel. L'étape initiale consiste en un étalonnage géométrique de l'endoscope (estimation des paramètres intrinsèques et de distorsion, cf. section 2.2.6), suivi d'un recalage 3D-2D destiné à initialiser le suivi visuel de l'utérus. Pour ce faire, une reconstruction dense de la surface de l'organe est générée à partir d'une vidéo exploratoire étalonnée de la cavité péritonéale via SfM [Griwodz 2021]. Le problème de recalage 3D-2D est alors résolu par une optimisation numérique qui combine des contraintes géométriques, comme les distances organe-reconstruction et les distances entre fragments de silhouettes ainsi que des contraintes d'un modèle d'énergie.

Pour réaliser la reconstruction du modèle peropératoire via SfM les auteurs ont développé un algorithme de prise d'IC qui présentent des déplacements spatiaux mutuels significatifs. Des points clés SIFT [Lowe 2004] sont alors extraits puis suivis sur l'intégralité du flux vidéo. Dans la version originale du papier, la sélection des nouvelles IC est guidée par des critères liés au nombre de points clés et à leur amplitude de déplacement. Cette stratégie est ensuite remplacée par l'analyse du flux optique et l'estimation de poses relatives. Le but est de sélectionner en temps réel, au cours de l'exploration de la cavité, un minimum de quinze IC nécessaires à la reconstruction SfM. Par la suite, l'intégration d'une segmentation sémantique de la surface utérine [Zadeh 2023] a permis de fiabiliser le processus, en éliminant les déplacements contradictoires générés par les tissus environnants et les instruments (cf. section 5.4.1).

Le recalage 3D-2D automatique déformable repose sur la détection des contours de la silhouette utérine, marqués initialement de façon manuelle sur les IC, puis automatisé au moyen d'un réseau de neurones d'architecture UNet spécifiquement entraîné [François 2020, Zadeh 2023], comme illustré sur la figure 5.15. Cette étape permet de contraindre en forme le recalage 3D-2D. L'initialisation est réalisée soit à partir d'une position canonique de l'utérus ou soit par une méthode de recalage 3D-2D semi-automatique de type PnP (cf. section 4.5.2).

Afin d'initialiser l'échelle de la reconstruction, le maillage du jumeau numérique recalé lors de l'initialisation est rendu graphique en tenant compte des paramètres intrinsèques de l'endoscope, pour obtenir une carte de profondeur synthétique à partir de la mémoire tampon de profondeur (z-buffer en Anglais) OpenGL, comme illustré en figure 5.14. À partir de cette dernière, une barymap est construite : elle associe à chaque pixel un quadruple $(f, \alpha, \beta, \gamma)$, où f est l'indice de la face triangulaire correspondante et (α, β, γ) sont les coordonnées barycentriques du point projeté sur cette face. Ces informations permettent de relier précisément chaque pixel de l'image à une position tridimensionnelle sur le maillage.

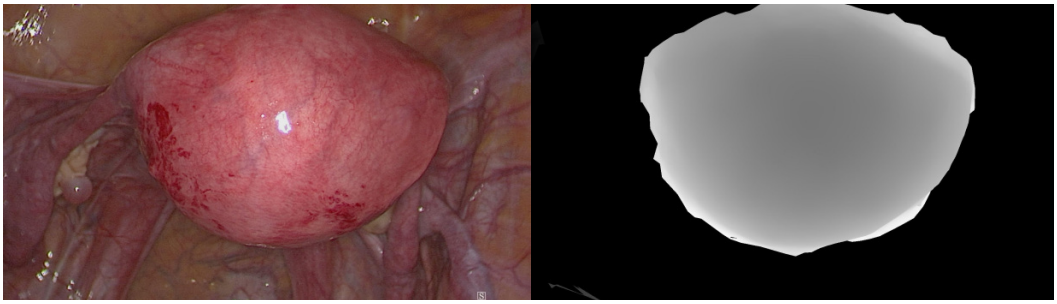


FIGURE 5.14 – Exemple d'une carte de profondeur synthétique générée à partir de la mémoire tampon de profondeur OpenGL, obtenue par rendu graphique du recalage 3D-2D d'un jumeau numérique de l'utérus sur une image peropératoire. Source : auteur.

La fonction d'énergie utilisée pour le recalage 3D-2D déformable comprend trois termes : une énergie interne du modèle, un coût lié au nuage de points, basé sur la fonction de coût de l'ICP afin de rapprocher la surface du jumeau numérique de celle reconstruite en peropératoire et un terme pour contraindre la silhouette du jumeau numérique à s'aligner sur les fragments de silhouette détectés dans les IC. L'optimisation utilise un algorithme de Gauss-Newton itératif jusqu'à la convergence ou un maximum de vingt itérations.

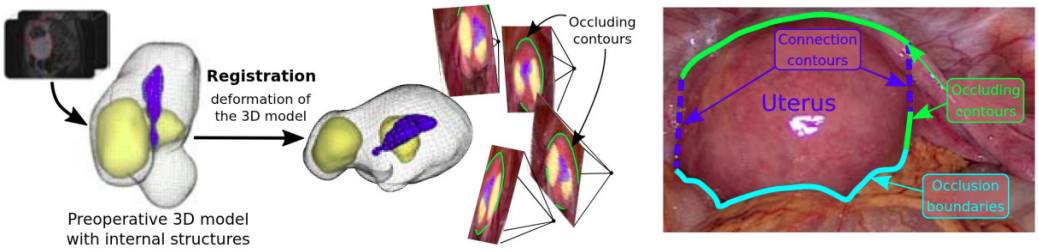


FIGURE 5.15 – Réalité Augmentée (RA) en Chirurgie Mini-Invasive (CMI) de l'utérus. (en haut) Recalage 3D-2D du jumeau numérique sur les vues peropératoires via ajustement des silhouettes. (en bas) Exemples de contours de l'utérus visibles en peropératoire. Source : générée par l'auteur à partir de [François 2020].

Une fois le recalage 3D-2D initial réalisé, le maillage du jumeau numérique est rendu avec OpenGL pour chaque IC et une carte de profondeur synthétique de la zone située sous le maillage est générée et sauvegardée. Cette donnée contribue à calculer la position tridimensionnelle de chaque point clé bidimensionnel présent dans une IC. Les points clés SIFT sont extraits via l'implémentation GPU POP-SIFT [Griwodz 2018], après un prétraitement des images incluant redimensionnement, sélection du canal vert et filtrage gaussien avec un noyau de 15×15 pixels,

afin d'accélérer l'extraction. L'ensemble des données obtenues : IC, cartes de profondeurs synthétiques, points clés bidimensionnels et leurs projections tridimensionnelles constituent la BIC statique.

Pour chaque nouvelle image issue du flux vidéo étalonné, des points clés sont extraits selon la même procédure que pour les IC. L'IC la plus proche de l'image est recherchée dans la BIC statique par mise en correspondances des descripteurs, via une approche de type force brute (BFM en Anglais) et TRL. L'IC la plus similaire est sélectionnée par une stratégie winner-takes-all, puis une estimation de la pose de la caméra est effectuée via PnP RANSAC. Les inliers permettent ensuite d'affiner la pose via une nouvelle estimation PnP, avant d'appliquer un filtre EKF pour réduire les oscillations. L'image augmentée est alors rendue et affichée.

Dans la méthode originale, les IC peuvent être ajoutées manuellement à la BIC statique afin de tenir compte des changements de points de vue et d'apparence de l'organe. Cette opération n'avait cependant pas été automatisée par les auteurs afin de limiter le risque de dérive et de dégradation du suivi [Collins 2014].

Deux évaluations du TRE ont été réalisées par les auteurs sur des modèles physiques synthétiques aussi appelés phantoms, filmés par une caméra endoscopique en conditions proches du réelle.

La première, dédiée au recalage 3D-2D initial, a montré que l'erreur de recalage diminue avec le nombre d'IC utilisé, pour atteindre 1,4 millimètre pour plus de huit IC. L'erreur augmente également logiquement avec l'éloignement de la zone contrainte par les contours [Collins 2014].

La seconde évalue le pipeline complet composé du recalage 3D-2D initial et du suivi visuel. Deux cent images sont prises du modèle afin de reconstruire un maillage surfacique utilisé comme jumeau numérique préopératoire. Des déformations synthétiques isovolumétriques et quadratiques selon l'axe principal sont simulées, avec des amplitudes maximales d'environ dix millimètres pour mimer les conditions peropératoires. Dix scénarios sont générés en augmentant progressivement l'amplitude des déformations. Pour chacun, le pipeline est exécuté en prenant quinze IC initiales. Le TRE est évalué entre 1,8 et 3,9 millimètres respectivement dans la zone proche de la surface (inférieure à 20 millimètres) et pour les déformations les plus sévères [Collins 2014].

Pour le cas applicatif principal de ce manuscrit qui est la NPRA, Teluob et al. [Teluob 2019] ont évalué qualitativement la méthode de suivi visuel proposée par Collins et al. [Collins 2020]. Le recalage 3D-2D initial n'est pas applicable au rein, car la silhouette de l'organe ne constitue pas une primitive de recalage fiable et facilement détectable (cf. section 4.4).

Nous avons donc adapté la méthode aux contraintes chirurgicales et anatomiques spécifiques au rein (cf. section 5.9). Compte tenu de la faible maturité des méthodes de recalage 3D-2D en NPRA (cf. chapitre 4), une initialisation manuelle du suivi est choisie pour sa robustesse et sa facilité de mise en œuvre, sans recours à une reconstruction peropératoire de l'organe cible.

Lors du suivi, le jumeau numérique du rein qui se compose à minima de cinq structures anatomiques (le rein, la tumeur, l'artère rénale, la veine rénale et l'uretère) est affiché en RA. Un second modèle tridimensionnel, clone du jumeau numérique est généré pour permettre la sélection dynamique, lors de la projection des points clés d'une IC, des surfaces impliquées dans l'interaction avec les rayons de projection, et donc les structures anatomiques suivies. Ce filtrage spatial des points clés détectés est ajustable aux cours de la chirurgie en fonction de la visibilité des éléments anatomiques. Généralement les vaisseaux et l'uretère sont exclus sauf indication spécifique tel que la chirurgie du thrombus de la veine cave inférieure. La segmentation automatique des images de NPRA (cf. section 5.4) est également exploitée pour restreindre la sélection aux points clés directement visibles et non masqués par les tissus péri-rénaux.

L'ajout d'IC au cours du suivi visuel permet d'augmenter significativement le nombre d'images augmentées en RA. Dans certains cas, une unique IC supplémentaire permet de doubler le nombre d'images suivies. Qualitativement, la stabilité du suivi visuel est ainsi renforcée [Collins 2014]. De cette observation, nous avons développé une méthode d'ajout automatique d'IC à la BIC, tout en maintenant une exécution temps réel (cf. section 5.6).

5.6 Suivi monoculaire via une base d'images clés dynamique

La sélection automatique d'IC constitue un composant critique mais à forte valeur ajoutée pour le processus de suivi visuel. De manière plus générale, ce problème est récurrent dans de nombreuses tâches de vision par ordinateur, qu'elles soient hors ligne ou en temps réel, comme l'estimation de pose de caméra, le SLAM ou encore l'odométrie visuelle. Les IC sélectionnées doivent assurer une couverture adéquate de l'espace visible tout en évitant la redondance. Le recours à un arbre de vocabulaire pour organiser les descripteurs de points clés et accélérer le processus d'appariement a été proposé dans la littérature [Dong 2009, Dong 2014]. Le défi majeur réside alors dans la définition de critères pour décider si l'image courante du flux vidéo peut être considérée comme formant une nouvelle IC.

Nous avons étendu la méthode de suivi visuel statique (cf. section 5.5) en introduisant une BIC dynamique, initialement peuplée d'IC provenant de la phase hors ligne de recalage 3D-2D initial, puis enrichie automatiquement en temps réel, comme illustré en figure 5.16. Cette nouvelle approche, que nous qualifions de méthode de suivi dynamique, repose sur deux modules complémentaires : la sélection automatique d'IC et la gestion de la BIC. La partie statique de la BIC demeure inchangée afin de préserver la robustesse du système sur le long terme et d'éviter toute dérive susceptible de conduire à la perte complète du suivi de l'organe.

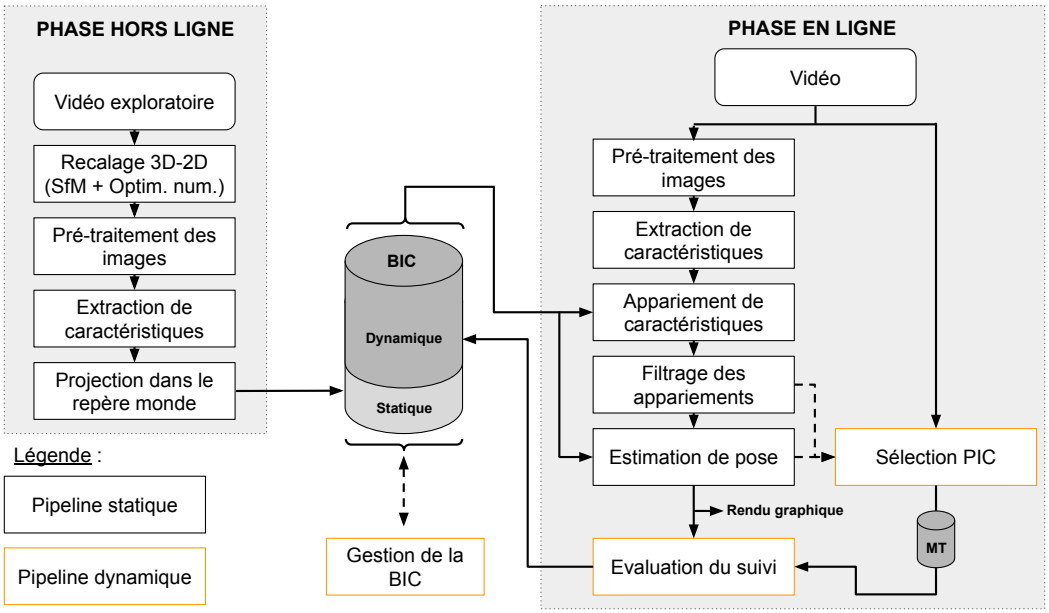


FIGURE 5.16 – Représentation schématique de la méthode de suivi visuel dynamique qui utilise une Base d'Images Clés (BIC) dynamique et un sélecteur automatique de Potentielles IC (PIC) stockées dans une Mémoire Tampon (MT). Source : auteur.

En peropératoire, un module de sélection automatique d'IC est intégré parallèlement au processus standard de suivi. Celui-ci analyse l'image courante selon plusieurs Critères de Sélection (CS), appliqués de manière cumulative, afin de déterminer si cette image doit être considérée comme une Potentielle Image Clé (PIC). Les CS doivent rester peu coûteux en temps de calcul, puisque la contrainte temporelle correspond à la période inter-images (16,7 millisecondes pour un flux à 60 Hz) mais suffisamment discriminants pour ne retenir que les images susceptibles de contribuer positivement au suivi.

A la suite d'essais expérimentaux menés en chirurgie de l'utérus, trois CS ont été retenus. Une image est considérée comme une PIC si les conditions suivantes sont satisfaites : CS1 est validé si le résidu quadratique moyen de l'erreur de reprojection lors de l'estimation de la pose est inférieur à 0,4 % de la diagonale de l'image ; CS2 est validé si l'accélération de la caméra en unités normalisées SfM est inférieure à 1 s^{-3}, une forte accélération est typiquement associée à une instabilité de pose ; CS3 est validé si la variance du Laplacien de l'image courante est inférieure à 2,9 % de l'intensité maximale de l'image, ce qui traduit un niveau de mise au point satisfaisant.

En chirurgie du rein, un quatrième critère, CS4 est ajouté afin d'évaluer la distribution spatiale des IC. Celui-ci est validé lorsque l'écart angulaire entre la PIC et l'IC la plus proche, mesuré depuis le centre géométrique du jumeau numérique, est compris entre 2° et 25°.

Chaque CS agit indépendamment sur un paramètre spécifique du module de sélection automatique d'IC et n'est pas directement corrélé aux autres. Les seuils associés ont été définis individuellement à partir d'observations empiriques. La sélection d'une PIC résulte de la validation séquentielle de tous les CS, combinés via l'opérateur logique « ET ». L'ordre de vérification influe uniquement sur l'efficacité algorithmique. Dans notre implémentation, les tests sont exécutés selon un ordre croissant de temps de calcul afin de réduire le coût global.

Lorsque tous les CS sont validés, la PIC est ajoutée à une Mémoire Temporaire (MT_PIC) de type Last In First Out (LIFO) en Anglais, de taille $n_{MTPIC} = 1$.

Un second module, absent de la méthode de suivi statique, assure la gestion de la BIC. Il a pour mission de monitorer les performances du suivi visuel et d'administrer l'ajout ou le remplacement dynamique des IC. Pour cela, deux Critères de Déclenchement (CD) sont définis : CD1 est validé si le nombre de correspondances de points clés entre l'image courante et l'IC sélectionnée est inférieur à 50 ; CD2 est validé si le nombre d'inliers obtenus lors de l'estimation de pose par PnP RANSAC est inférieur à 8.

Ces seuils ont également été établis empiriquement et de manière individuelle. Lorsque les deux CD sont vérifiés sur trois images consécutives, afin de réduire les faux positifs, le gestionnaire de BIC propose d'ajouter la dernière PIC stockée dans la MT_PIC, après confirmation par l'utilisateur. Pour cela, une fenêtre contextuelle, illustrée en figure 5.17, s'ouvre et affiche la PIC candidate, que l'utilisateur peut alors accepter, rejeter ou corriger en ajustant manuellement la pose du jumeau numérique.

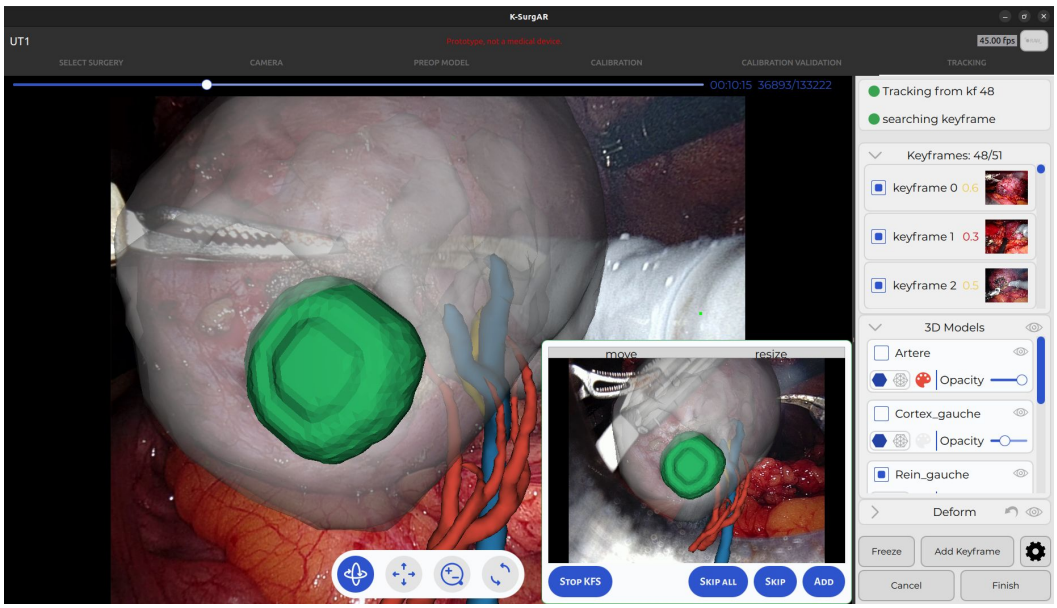


FIGURE 5.17 – Exemple d'implémentation de la fenêtre contextuelle qui laisse l'utilisateur accepter, rejeter ou corriger manuellement la pose du jumeau numérique sur la Potentielle Image Clé (PIC). Source : auteur.

Dans nos premières implémentations, les ressources matérielles disponibles limitaient la taille de la BIC : au-delà de $N = 50$, des ralentissements perceptibles par l'utilisateur apparaissaient. Pour pallier cette contrainte, nous avons enrichi le gestionnaire de BIC d'un mécanisme de remplacement des anciennes IC : si la taille de la BIC excède N, les IC les moins sollicitées (déterminées à partir du temps écoulé depuis leur dernière utilisation) sont remplacées en priorité, garantissant ainsi la pertinence du contenu de la BIC.

Enfin, l'interface utilisateur offre également des fonctionnalités de gestion manuelle de la BIC. L'utilisateur peut à tout moment ajouter de nouvelles IC ou désactivées certaines d'entre elles.

5.7 Méthode étendue à la stéréovision

Dans le contexte clinique de la NP guidée par RA, nous avons généralisé la méthode dynamique décrite en section 5.6 à un cadre de traitement de flux vidéo stéréoscopique. Rappelons que la NP est majoritairement réalisée sous assistance robotique (NPRA), ce qui permet d'accéder en temps réel aux paires d'images gauche-droite issues de la caméra endoscopique (cf. section 1.2.4.3).

La méthode proposée, étendue à la stéréovision, est illustrée en figure 5.18 et reprend les principes de base du suivi dynamique monoculaire (cf. section 5.6). L'endoscope stéréoscopique est tout d'abord étalonné géométriquement afin d'estimer les paramètres internes et de distorsion propres aux caméras gauche et droite ainsi que les paramètres externes définissant la géométrie relative du système. Un recalage 3D-2D rigide initial du jumeau numérique est réalisé manuellement sur une IC stéréoscopique. La pose déterminée sur l'une des vues de la paire est ensuite propagée à l'autre vue par exploitation des paramètres externes. Le module de suivi visuel est alors adapté afin de tirer parti des IC stéréoscopiques et de la cohérence gauche-droite induite par la géométrie épipolaire (cf. section 2.2.6). La carte de profondeur est utilisée pour renforcer la robustesse des appariement et par conséquent celle du suivi visuel.

Pour cela, l'extraction des points clés SIFT [Lowe 2004] est effectuée simultanément sur les canaux gauche et droit du flux stéréoscopique préalablement rectifié. Une première étape de mise en correspondance par force brute (BFM) est réalisée dans les deux sens (gauche-droite puis droite-gauche) afin de ne conserver que les points communs aux deux vues. Ces correspondances, sont ensuite filtrées à l'aide du TRL puis d'un test de vérification de la contrainte épipolaire. Une tolérance en termes de distance aux lignes épipolaires de 0,07 % de la diagonale de l'image, soit environ 1,5 pixel pour résolution de 1920×1080 pixels est fixée.

Comme pour la méthode monoculaire, seuls les points clés localisés à la surface du parenchyme rénal sont utilisés pour le suivi. Le réseau de neurones spécifiquement entraîné pour segmenter la zone d'intérêt (cf. section 5.4) est de nouveau mis en œuvre afin de réduire l'espace de recherche.

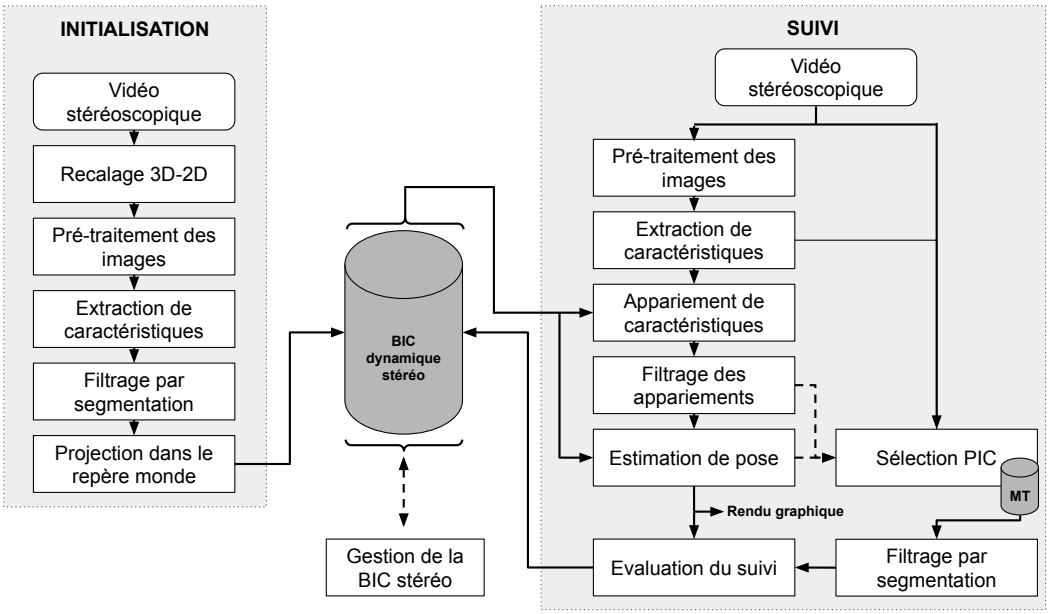


FIGURE 5.18 – Représentation schématique de la méthode de suivi visuel stéréoscopique qui utilise une Base d'Images Clés (BIC) dynamique et un sélecteur automatique de Potentielles Images Clés (PIC) stockées dans une Mémoire Tampon (MT). Source : auteur.

5.8 Méthode avec correction automatique de la perspective

Les performances des méthodes de suivi par détection sont dépendantes des détecteurs-descripteurs utilisés. Dans notre cas, SIFT [Lowe 2004] ou son implémentation GPU POPSIFT [Griwodz 2018], offrent une invariance à la rotation et aux changements d'échelle, mais demeurent sensibles aux effets de parallaxe. Ces derniers deviennent particulièrement marqués lorsque la caméra est placée à proximité de la surface d'un organe convexe, comme cela est le cas pour le rein, dont la convexité est accentuée au niveau des pôles.

Afin de limiter cette dégradation, la détection et l'appariement des points clés peuvent être enrichis par un processus que nous désignons sous le terme de CP. Notre méthode exploite conjointement la carte de profondeur stéréoscopique et l'aplatissement de surface, pour générer une transformation d'image qui compense l'effet induit par la perspective dans les images stéréoscopiques.

Ce nouveau module de CP s'intègre de manière transparente dans le système de suivi proposé en section 5.7, juste avant l'étape d'extraction des points clés, sans autre modification du processus, comme illustré en figure 5.19.

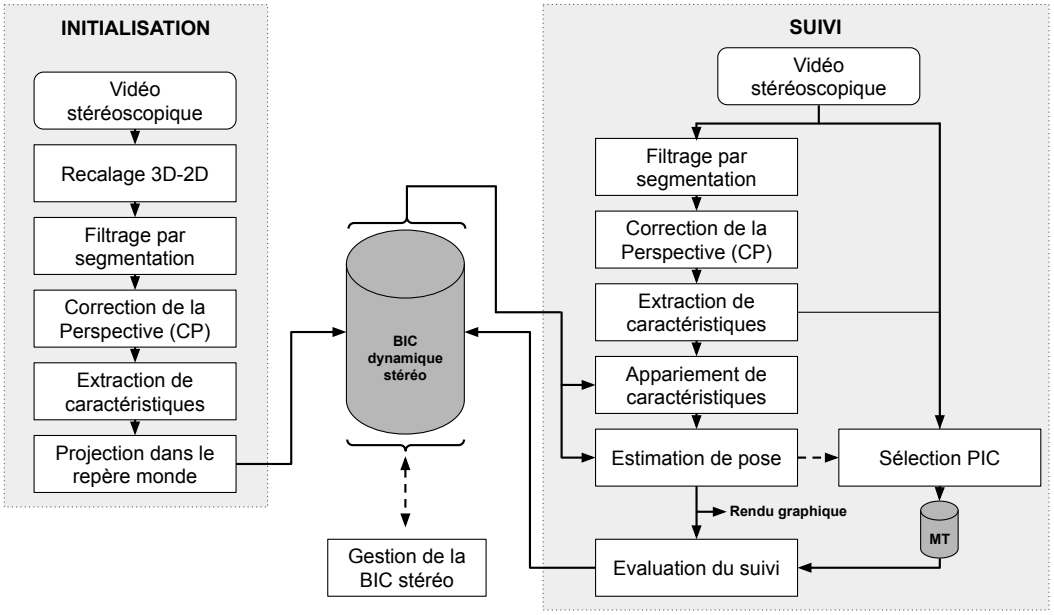


FIGURE 5.19 – Représentation schématique de la méthode de suivi visuel stéréoscopique qui utilise une Base d'Images Clés (BIC) dynamique, un sélecteur automatique de Potentielles Images Clés (PIC) stockées dans une Mémoire Tampon (MT) et un module de Correction de la Perspective (CP). Source : auteur.

Le fonctionnement du module de CP utilise la carte de profondeur calculée à partir de la paire d'images stéréoscopique. Celle-ci est d'abord rectifiée et corrigée des distorsions optiques via l'utilisation des paramètres d'étalonnage géométrique de l'endoscope. Une carte de disparité filtrée est ensuite produite au moyen de la méthode SGBM [Hirschmuller 2005, Hirschmuller 2007], complétée par un filtre WLS [Min 2014], seules techniques suffisamment fiables lors de la réalisation de ces travaux. Aujourd'hui, les méthodes basées sur l'apprentissage profond sont exploitables (cf. section 4.5.2). A partir de cette carte de disparité, un nuage de points tridimensionnel est reconstruit, comme illustré en figure 5.20.

En parallèle, la surface rénale est segmentée dans les images peropératoires via un réseau de neurones, tel que celui proposé en section 5.4. Un maillage tridimensionnel peropératoire est ensuite obtenu par triangulation du nuage de points composant le masque de la surface rénale. Ce maillage, en tant qu'approximation d'une surface différentiable, peut être projeté dans le plan en respectant le théorème de la carte conforme de Riemann, qui garantit l'existence d'une telle transformation de R^3 vers R^2 avec une distorsion angulaire minimale [Hormann 2007]. Dans cette première version du module de CP, la cartographie conforme aux moindres carrés (LSCM en Anglais) [Lévy 2023, Mullen 2008] implémentée dans la bibliothèque C++ libigl est utilisée. Elle aplatit le maillage tridimensionnel tout en préservant approximativement les angles, résolvant un problème d'optimisation qui recherche la transformation qui projette les triangles de R^3 vers R^2 avec une distorsion conforme minimale et en maintenant la connectivité des triangles.

Le maillage aplati est ensuite texturé en déformant les images originales via une transformation affine par morceaux, définie à partir des correspondances entre sommets. La faible courbure locale de la surface, comparée à la résolution du maillage, permet d'approximer la transformation de texture entre le plan de la scène et l'image par une fonction affine. Pour cette raison, les coordonnées barycentriques locales sont utilisées pour déterminer les positions correspondantes des pixels entre triangles aplatis et l'image source, ce qui assure une cartographie fidèle de la texture sur le modèle aplati. Les deux vues de la paire stéréoscopique sont traitées de manière symétrique selon ce principe.

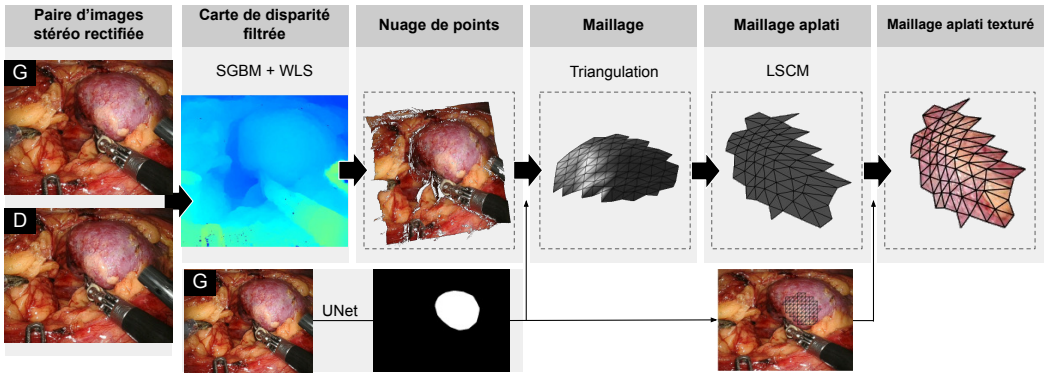


FIGURE 5.20 – Illustration du module de Correction de Perspective (CP) appliquée à l'image de gauche ; le même pipeline est également exécuté sur l'image de droite pour obtenir deux images aplaties. Source : auteur.

5.9 Méthode spécialisée à la néphrectomie partielle robot assistée

Dans le cadre spécifique de l'application clinique à la NPRA et des essais de faisabilité (cf. section 2.2.3), la stratégie retenue et évaluée en conditions réelles est représentée en figure 5.21. La planification chirurgicale, basée sur la génération du jumeau numérique, est réalisée par l'équipe chirurgicale chargée de l'intervention quelques jours avant la date prévue.

Au bloc opératoire, le matériel nécessaire à l'augmentation du flux vidéo stéréoscopique issu du processeur vidéo Da Vinci Xi (Intuitive Surgical) est installé (cf. section 2.2.5). Il comprend un écran de contrôle ainsi qu'un ordinateur Ubuntu 24.04 équipé d'un processeur Intel Core™ i9-14900, d'un GPU Nvidia RTX 4070 et d'une carte d'acquisition vidéo Magewell Pro Capture HDMI Quad. En complément des deux flux principaux (caméra gauche et caméra droite de l'endoscope stéréoscopique), deux flux auxiliaires sont intégrés afin d'utiliser l'ordinateur comme un mélangeur vidéo. Ce dernier, rend possible l'affichage, dans la console du télémanipulateur robotique, soit d'une version stéréoscopique du flux augmenté par RA, soit du flux droit augmenté couplé à un canal auxiliaire qui peut transporter par exemple le flux de l'échographie peropératoire.

Le chirurgien qui opère via la console du robot conserve une autonomie totale dans l'usage de la RA, grâce au report des commandes dans la vue du TilePro du Da Vinci Xi (Intuitive Surgical), comme illustré en figure 5.22. Après l'étalonnage de l'endoscope, réalisé en début d'intervention, le chirurgien procède manuellement, à l'aide d'une souris standard, à un recalage 3D-2D rigide du jumeau numérique sur la vue peropératoire. Ce recalage génère une première IC et initialise la méthode de suivi visuelle basée sur une BIC dynamique et un sélecteur automatique de PIC décrite précédemment (cf. sections 5.6 et 5.7).

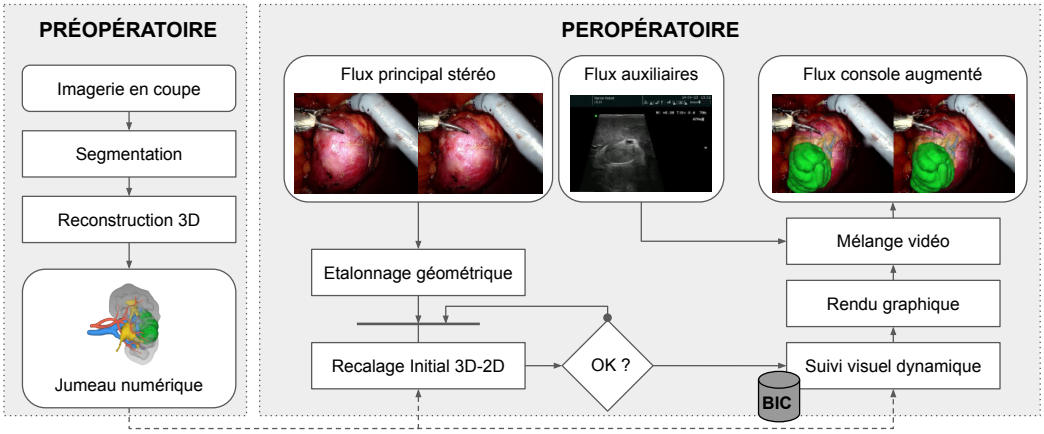


FIGURE 5.21 – Représentation schématique du pipeline proposé en Néphrectomie Partielle Robot Assistée (NPRA) guidée par Réalité Augmentée (RA), capable de traiter en temps le flux vidéo monoculaire ou stéréoscopique du télémanipulateur robotique. Source : auteur.

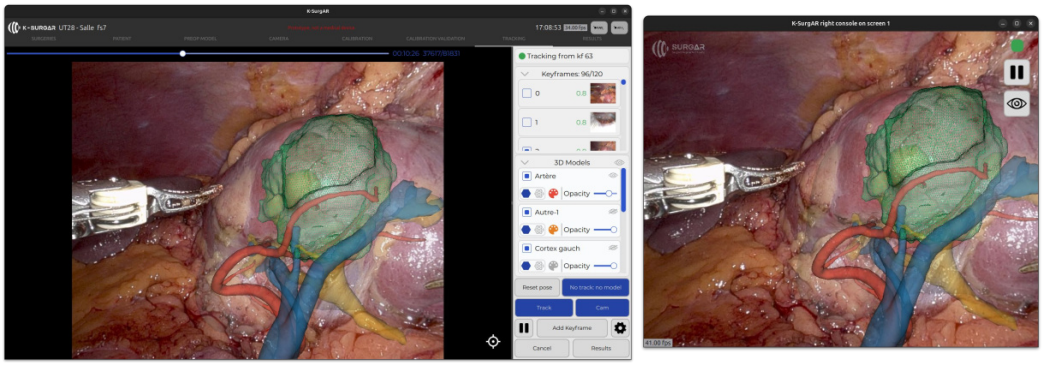


FIGURE 5.22 – Néphrectomie Partielle Robot Assistée (NPRA) guidée par Réalité Augmentée (RA). (à gauche) Vue de l'écran de contrôle avec l'ensemble des commandes. (à droite) Affichage minimaliste dans le TilePro (Intuitive Surgical). Source : auteur.

Sur le plan logiciel, plusieurs ajustements de la méthode de suivi publiée ont été introduits afin d'améliorer sa robustesse en conditions réelles et de garantir son exploitation par un unique utilisateur avec un minimum d'interactions.

Un mécanisme d'hystérésis a été ajouté au changement d'IC : une fois la transition effectuée, le retour vers l'IC précédente identique est rendu plus coûteux, favorisant ainsi une séparation franche et limitant les oscillations indésirables. La sélection de l'IC repose toujours sur un vote parmi les IC de la BIC, en fonction du nombre de correspondances filtrées avec l'image courante. La pénalisation appliquée à l'IC précédente est définie par une réduction multiplicative de son score.

Pour l'extraction des caractéristiques, une approche multi-canaux a été développée, inspirée des travaux de Selka et al. [Selka 2012]. Contrairement à la détection et description POPSIFT [Griwodz 2018] appliquées uniquement sur le canal vert, celles-ci sont désormais effectuées indépendamment et en parallèle sur les trois canaux. Une étape de fusion consolide ensuite les points détectés trop proches entre canaux et ne conserve qu'un ensemble unifié de points clés. Ce mécanisme de consensus, basé sur un seuillage du déplacement λ_d et de la dispersion inter-canaux λ_m, permet de rejeter les estimations incohérentes et d'accroître la robustesse du suivi visuel, via une augmentation d'environ 40 % du nombre de points clés suivi.

Concernant la segmentation de la surface rénale, une mesure de contournement préventive est introduite dans les cas où le rein ne serait pas ou mal détecté. Si la fraction de la surface segmentée est inférieure à 1 % de l'image totale, le masque est remplacé par le négatif du masque de segmentation des instruments, afin de prévenir une erreur critique.

Lors de l'appariement des correspondances, un filtrage GMS [Bian 2017] est appliqué en complément du TRL.

Afin d'exploiter pleinement la connaissance embarquée dans la BIC et de lisser les transitions entre IC, un ajustement de faisceaux local (bundle adjustment en Anglais) est lancé à chaque ajout d'une nouvelle IC à la BIC. Les paramètres internes et externes de l'endoscope sont fixés à partir de l'étalonnage géométrique. Les poses issues des recalages manuels sont figées et jouent le rôle d'ancres, tandis que les poses automatiques et les points tridimensionnels observés sont optimisés. L'optimisation est réalisée dans une fenêtre glissante comprenant typiquement 5 à 10 IC récentes. Le critère retenu est la somme des erreurs de reprojection avec une perte robuste de Huber ($\delta = 1$ pixel).

La résolution utilise la bibliothèque Ceres, avec les paramètres principaux suivants : stratégie trust-region de type Levenberg–Marquardt, solveur linéaire SPARSE_SCHUR et arrêt borné en nombre itérations (ici maximum de vingt) afin de préserver une latence acceptable. Ce processus est exécuté en tâche de fond afin de ne pas perturber le suivi visuel en temps réel.

Enfin, deux modes dégradés sont proposés. Le premier, un mode RV suivi permet une mobilisation automatique du jumeau numérique basée sur la méthode de suivi, mais sans affichage de l'image, réduisant ainsi la sensibilité aux erreurs de recalage. Le second, un mode RV simple, constitue une solution de dernier recours pour néanmoins bénéficier de la planification chirurgicale réalisée.

5.10 Méthode de suivi basée sur l'apprentissage profond

En complément de la méthode historique de détection, description et appariement de points clés, nous avons exploré une approche plus contemporaine basée sur l'apprentissage profond. Ces travaux, bien que prometteurs, demeurent à un stade de développement moins avancé que ceux présentés précédemment et ne sont pas complètement achevés. Nous en proposons ci-après une première version.

Dans cette approche, la BIC dynamique ainsi que le mécanisme de sélection automatique des PIC sont conservés. En revanche, les composants de détection, description et d'appariement sont remplacés par un réseau de neurones dédié au suivi des trajectoires de points, définies dans un espace bidimensionnel ou tridimensionnel. Ce composant est soumis à plusieurs défis inhérents à la nature des vidéos chirurgicales, notamment les déformations non-rigides, les auto-occultations, les changements de points de vue, les mouvements brusques de la caméra ainsi que les occultations dues aux instruments chirurgicaux. Le réseau doit donc présenter une robustesse suffisante pour s'adapter à la diversité des conditions peropératoires.

Dans le cas d'un flux monoculaire, nous retenons un réseau de type CoTracker 3 [Karaev 2024a], adapté aux contraintes temps réelles. Pour rappel, CoTracker 3 repose sur une architecture basée sur les transformers et traite la séquence d'images par fenêtres glissantes de 16 images. Un CNN extrait des caractéristiques multi-échelles pour chaque image dans la fenêtre. Autour de chaque query point, un patch de caractéristiques est prélevé puis transformé en vecteur de corrélation au moyen d'un perceptron multicouche (MLP en Anglais). Les vecteurs de corrélation, les points estimés, la confiance et la visibilité alimentent un module de raffinement itératif (six itérations dans le papier original), qui met à jour ces informations pour toutes les images contenues dans la fenêtre glissante.

Par défaut, CoTracker 3 présente un temps d'inférence de 200 millisecondes sur un GPU Nvidia RTX 4070. En réduisant la mémoire d'entrée à une unique image et en évitant la redondance des calculs, notamment par stockage en mémoire des résultats du MLP, ce temps peut être réduit drastiquement à environ 30 millisecondes.

Pour le traitement de flux stéréoscopiques, l'utilisation de réseaux de neurones qui travaillent directement dans l'espace tridimensionnel apparaît plus pertinent, compte tenu de leurs performances accrues et sachant que nous disposons des cartes de profondeur issue de la stéréovision. Leur emploi en monoculaire est en revanche limité par la nécessité d'un prédicteur de type monodepth, dont la justesse est très variable. Nous avons retenu TAPIP3D [Zhang 2025], qui constitue l'état de l'art actuel. Toutefois, son principal inconvénient réside dans un temps d'inférence d'environ 100 millisecondes, insuffisant pour le traitement d'un flux vidéo à 60 Hz. De plus, la génération des cartes de profondeur, indispensable à la projection des points image dans le repère monde, requiert elle aussi près de 100 millisecondes via SGBM et WLS, ce qui limite la fréquence de mise à jour de la pose du jumeau a environ 5 Hz, cadence encore trop faible pour une application clinique.

Les IC offrent la possibilité de sauvegarder des états de suivi pertinents et de réinitialiser, si nécessaire, la mémoire du réseau en cas de dérive excessive.

Nous avons choisi de suivre un sous-ensemble de sommets du maillage correspondant aux faces visibles lors du recalage 3D-2D sur l'IC initiale, plutôt qu'une grille bidimensionnelle régulière, comme illustré en figure 5.23. Cette stratégie vise à exploiter, dans une perspective future, les déplacements relatifs des points appariés au jumeau numérique afin d'estimer localement les déformations. A cette fin, le réseau de segmentation (cf. section 5.4) permet, comme dans les approches plus conventionnelles, de filtrer spatialement les sommets du maillage correspondant à la surface du parenchyme rénale. L'orientation des normales est exploitée pour exclure les sommets situés sur la face cachée du rein. L'ensemble des points retenus, limité à la centaine pour des raisons de performance, est suivi d'image en image par des correspondances quasi-exactes évaluées par un score de confiance. Les points jugés incertains sont remplacés dynamiquement par de nouveaux au fil du mouvement.

A ce stade, cette méthode n'a pas encore atteint un degré de maturité suffisant pour permettre une évaluation rigoureuse. Les premières observations, tant qualitatives que quantitatives, montrent une bonne cohérence visuelle de l'augmentation avec la vue peropératoire, bien qu'une dérive temporelle subsiste.

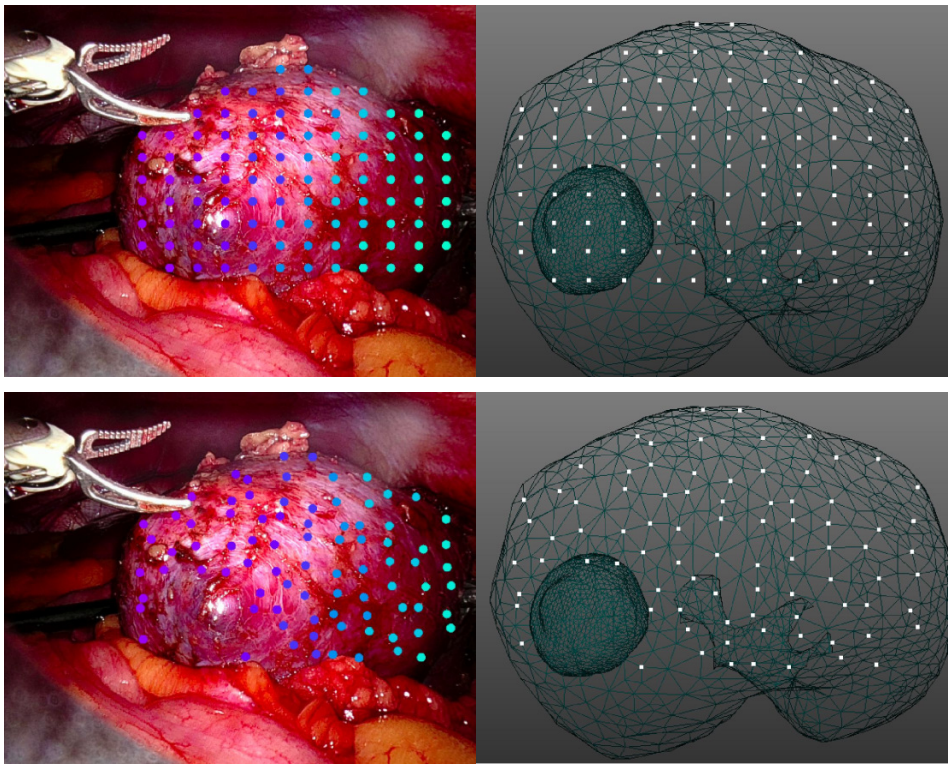


FIGURE 5.23 – Illustration de deux méthodes d'initialisation des query points utilisés par CoTracker 3 pour le suivi visuel du rein. (en haut) Disposition selon une grille régulière. (en bas) Projection 3D-2D des sommets du maillage du jumeau numérique recalé. Source : auteur et A. Samadi.

5.11 Expériences

L'évaluation des méthodes proposées a été réalisée selon la chronologie de leur développement, du suivi dynamique jusqu'à l'application aux données pré-cliniques et cliniques dans le cadre de la NPRA. Parallèlement, le module de segmentation automatique a fait l'objet d'une analyse spécifique en tant que composant clé du suivi visuel.

5.11.1 Segmentation automatique du parenchyme rénal

Le module de segmentation automatique, décrit en section 5.4, est issu d'une série d'optimisations portant simultanément sur l'architecture du réseau de neurones, le choix des hyperparamètres et la constitution de l'ensemble de données d'entraînement.

Dans un premier temps, deux évaluations comparatives au gold standard, défini par un chirurgien urologue expert, ont été menées respectivement en 2022 et 2023. Celles-ci avaient pour objectif principal d'analyser la variabilité de la labellisation du parenchyme rénal et de répondre à la question suivante : la tâche de labellisation doit-elle impérativement être confiée à des professionnels de santé ? L'étude a mobilisé cinq profils distincts d'annotateurs : deux chirurgiens urologues juniors, deux médecins non spécialistes, deux techniciens/ingénieurs impliqués dans le projet, ainsi que deux novices sans formation médical, en plus du chirurgien expert. La réflexion a également pris en compte la complexité visuelle des images et leur contexte opératoire, afin de favoriser un dialogue approfondi entre les équipes cliniques et techniques. Ces échanges ont conduit à la formalisation de règles d'annotation consolidées, telles que présentées en section 5.4.3.

La seconde phase d'évaluation a élargi ce dispositif, à la fois en nombre d'images et en nombre d'annotateurs impliqués. L'ensemble de données, constitué de 454 images extraites manuellement de 29 NPRA, est illustré en figure 5.24. Les images correspondent aux phases chirurgicales où la RA est cliniquement pertinente pour l'assistance au geste (cf. section 2.2). Les données ont été classées en trois niveaux de complexité visuelle croissante, présentées en figure 5.24. Cet indicateur, bien que subjectif, permet d'anticiper la difficulté de labellisation et s'appuie sur le type d'élément en occultation avec la surface rénale. Le premier niveau regroupe 97 images dites simples (21,4 %), où le parenchyme rénal est aisément identifiable. Le second, correspond à une complexité intermédiaire et inclut les cas où la graisse recouvre la surface rénale (78 images, 17,2 %), la présence de sang ou de caillots (66 images, 14,5 %) ou d'instruments, compresses, clips en avant de la zone d'intérêt (98 images, 21,6 %). Le troisième niveau rassemble 115 images dites complexes (25,3 %), caractérisées par la combinaison d'au moins deux facteurs perturbateurs.

Un contrôle de diversité a été effectué sur l'ensemble de données au moyen de la métrique SSIM, appliquée en comparaison croisée. La similarité moyenne de 35,9 %, confirme la représentativité et limite le risque de redondance.

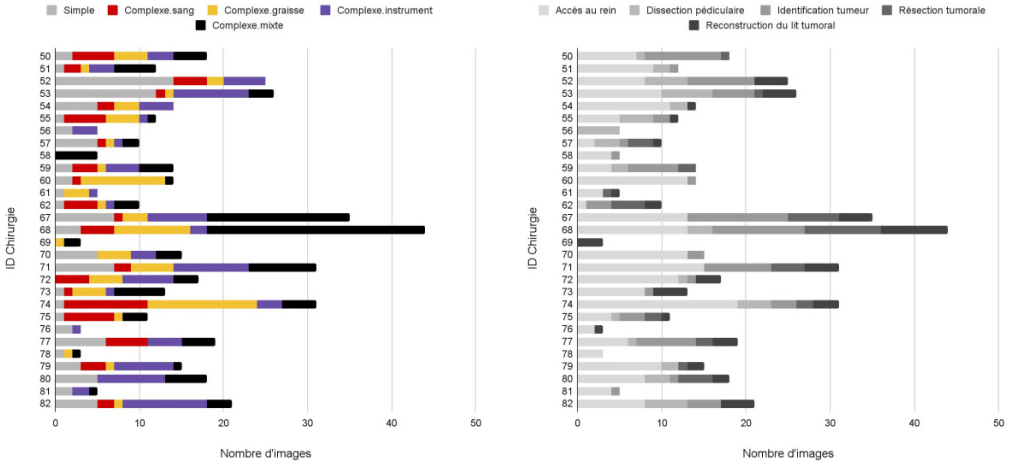


FIGURE 5.24 – Distributions des images composant l'ensemble de données pour l'évaluation comparative inter-annotateurs de la labellisation du parenchyme rénal : à gauche par complexité visuelle, à droite par phases chirurgicales. Source : auteur.

Cinq profils d'annotateurs ont ainsi été comparés au gold standard : le chirurgien urologue expert. Tous ont bénéficié d'une formation initiale standardisée portant sur les outils de segmentation et sur les règles de labellisation spécifiques aux images peropératoires de chirurgie rénale. Cette procédure reprend celle décrite pour la production de données annotées en section 5.4.3.

Les performances de chaque annotateur par rapport au chirurgien urologue expert ont été quantifiées à l'aide de trois indicateurs : spécificité, sensibilité et indice de Dice. Les résultats, rapportés dans le tableau 5.3, montrent une spécificité moyenne comprise entre 85 % et 97 %, une sensibilité entre 82 % et 93 % et un indice de Dice entre 81 % et 88 %. La durée moyenne de labellisation par image varie fortement selon les profils, de 31 secondes à plus de 4 minutes, confirmant que l'expertise thématique contribue à réduire ce temps.

Les analyses ont également été stratifiées selon la complexité visuelle et les phases chirurgicales, comme illustré en figures 5.25 et 5.26. Comme attendu, les performances diminuent pour les images complexes, en particulier en présence de sang ou d'une combinaison de plusieurs facteurs perturbateurs. La difficulté s'accroît également avec l'avancement de l'intervention, rendant l'interprétation des contours tumoraux sur une image isolée plus délicate. Malgré ces écarts, les performances des annotateurs non cliniciens restent comparables à celles des médecins. Les règles de labellisation ont donc été affinées afin de réduire l'écart avec le chirurgien expert, et de nouvelles classes ont été introduites afin de lever certaines ambiguïtés, comme exposé en section 5.4.3.

	Durée de labellisation par image (s)	Spécificité (%)	Sensibilité (%)	Indice de Dice (%)
Chirurgien junior 1	78 ± 58	96,5 ± 7,0	93,3 ± 21,1	88,1 ± 24,2
Chirurgien junior 2	53 ± 56	84,7 ± 7,1	91,5 ± 19,8	83,6 ± 23,3
Médecin 1	264 ± 201	93,0 ± 8,3	82,3 ± 20,0	81,6 ± 23,7
Médecin 2	142 ± 89	93,5 ± 8,0	90,6 ± 21,9	81,3 ± 24,9
Equipe technique 1	31 ± 17	96,5 ± 5,8	87,3 ± 22,2	83,8 ± 23,3
Equipe technique 2	50 ± 30	96,3 ± 6,1	86,4 ± 23,3	83,3 ± 23,7
Novice 1	236 ± 127	94,9 ± 7,2	86,6 ± 21,9	81,2 ± 23,9
Novice 2	277 ± 141	92,6 ± 8,7	92,3 ± 22,3	80,8 ± 25,0

TABLE 5.3 – Résultats de l'étude comparative inter-annotateur, moyennes et écart-types pour 454 images. Chaque annotateur est comparé à un chirurgien expert (gold standard) à l'aide des indicateurs de spécificité, sensibilité et indice de Dice. Les temps de labellisation par image sont également donnés. Source : auteur.

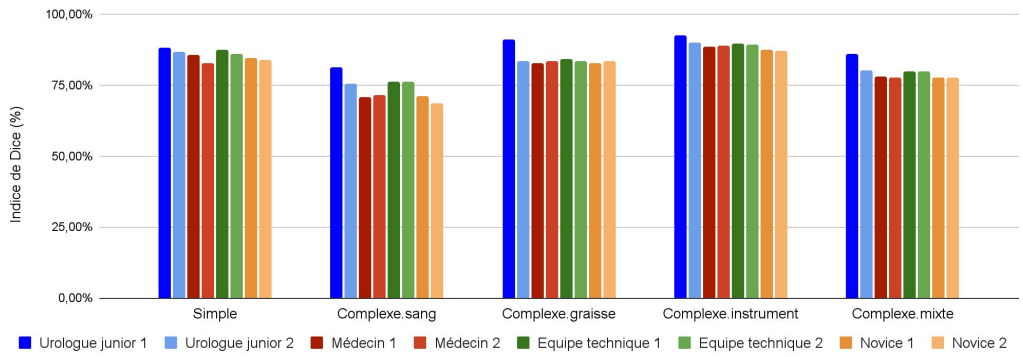


FIGURE 5.25 – Indice de Dice moyen évalué entre le chirurgien expert et les différents annotateurs en fonction de la complexité visuelle des images labellisées. Source : auteur.

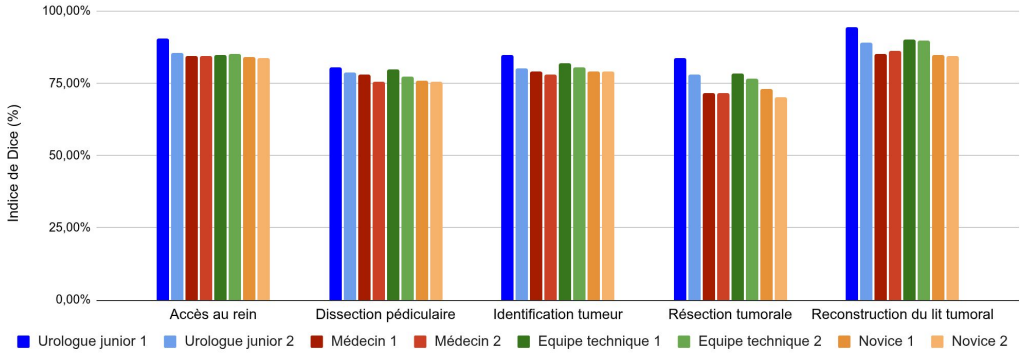


FIGURE 5.26 – Indice de Dice moyen évalué entre le chirurgien expert et les différents annotateurs en fonction de la phase chirurgicale associée aux images labellisées. Source : auteur.

L'analyse des accords et désaccords inter-annotateurs a été approfondie grâce aux cartes de chaleur présentées en figure 5.27, qui ont conforté les ajustements méthodologiques.

Ces résultats suggèrent qu'une labellisation qualitative peut être réalisée par des annotateurs non médecins, sous réserve d'un protocole de formation et de contrôle qualité. Ce constat a orienté la stratégie de constitution de l'ensemble de données vers une labellisation à grande échelle confiée majoritairement à des annotateurs non cliniciens, associés à des processus de vérification systématique (cf. section 5.4.3). Cette stratégie permet d'accroître significativement le volume de données disponibles pour l'apprentissage tout en déchargeant les cliniciens, dont le temps est prioritairement consacré aux soins.

Une première recherche d'architecture et d'hyperparamètres est réalisée sur l'ensemble de données DSPT (6 980 images), afin de permettre des itérations rapides. Plusieurs encodeurs d'image pré-entraînés sur ImageNet [Deng 2009] ont été comparés dans un cadre rigoureux de reproductibilité (graines fixées), incluant ResNet-34 [He 2016], MobileNet-v2 [Sandler 2018] et EfficientNet-B3 [Tan 2019b], combinés à des architectures UNet [Ronneberger 2015] et UNet++ [Zhou 2019b], et entraînés avec les optimiseurs SGD [Sutskever 2013], Adam [Adam 2014] et AdamW [Loshchilov 2017]. L'analyse des pertes de validation et des métriques sur les cinq classes prédites a conduit à retenir l'encodeur EfficientNet-B3 intégré dans une architecture UNet et associé à l'optimiseur AdamW. Ce réseau de neurones multiclasse, présenté en section 5.4.4, a ensuite été entraîné.

Dans une seconde étape, l'entraînement est élargi à l'ensemble de données DSP, composé de 41 160 images labellisées (32 928 pour l'entraînement et 8 232 pour la validation). L'impact de la taille de l'ensemble d'entraînement sur les performances de la segmentation est évalué à partir de sous-échantillons croissants, comme illustré en figure 5.28. Les performances maximales sont obtenues avec DSP au complet.

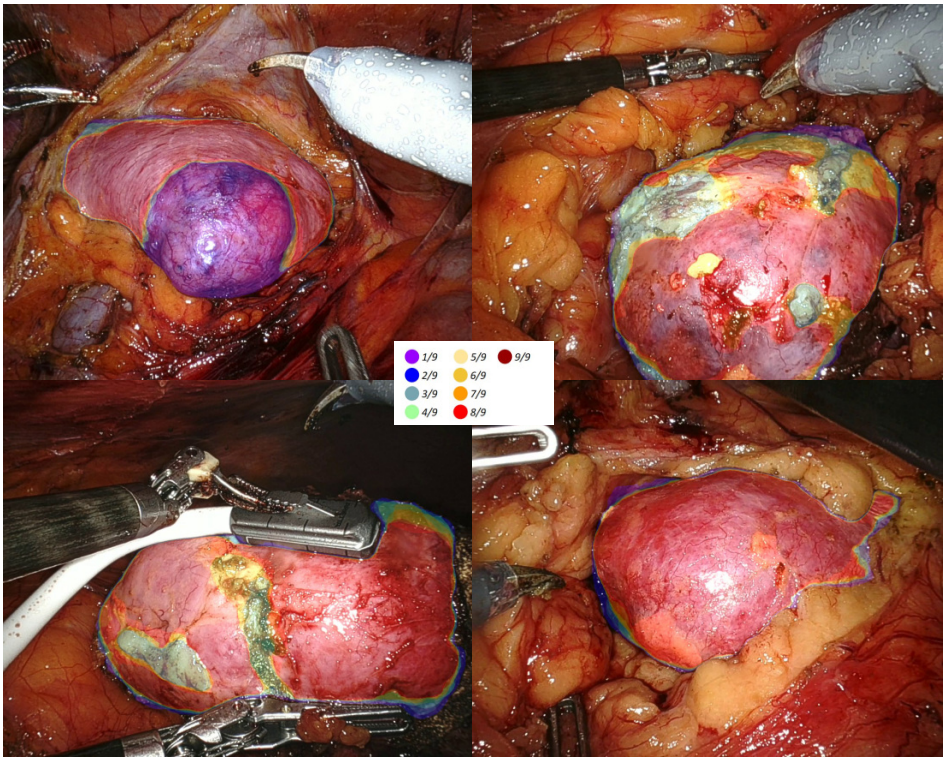


FIGURE 5.27 – Exemples de cartes de chaleur indiquant la concordance des segmentations inter-annotateurs pour la surface du parenchyme rénal. Violet, zone labellisée par un seul annotateur. Rouge foncé, zone labellisée par tous les annotateurs. Source : auteur.

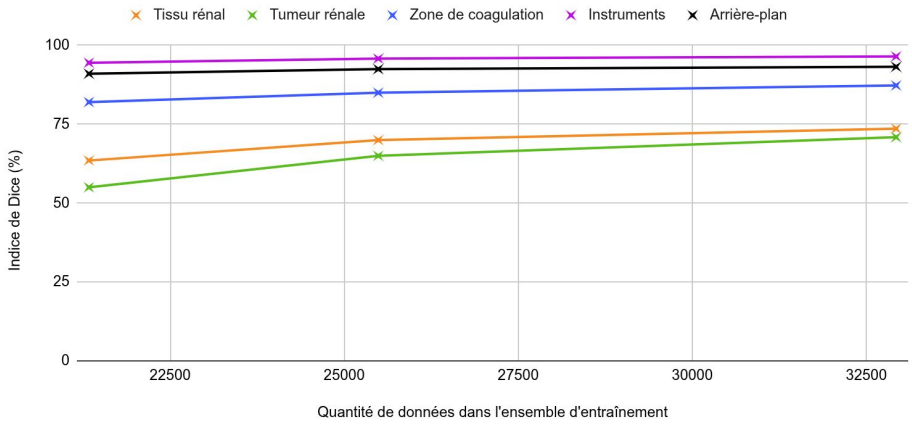


FIGURE 5.28 – Évolution de la performance (indice de Dice évalué sur l'ensemble de données DSPT) en fonction de la quantité de données utilisée pour l'entraînement du réseau de neurones. Source : auteur.

L'évaluation initiale est réalisée sur l'ensemble de données DSPT. Il comprend des images issues de phases chirurgicales où la RA est cliniquement pertinente. Les résultats, reportés dans le tableau 5.4, mettent en évidence une spécificité élevée pour toutes les classes (97,4 % en moyenne), ce qui indique une capacité robuste du réseau de neurones à éviter les faux positifs. En revanche, la sensibilité est plus variable, ce qui traduit la difficulté à segmenter correctement le tissu rénal et de la tumeur, dont les contours sont incertains et l'apparence est très variable entre les patients. Les masques obtenus par segmentation automatique sont illustrés en figure 5.29. Les indices de Dice suivent la même tendance. Ces observations corroborent la variabilité inter- et intra-patient, notamment liée à la proximité visuelle entre tissu sain, tissu tumoral et graisse périrénale.

	Tissu rénal	Tumeur	Zone de coag.	Instruments	Arrière-plan
Spécificité (%)	97,8 ± 3,9	99,1 ± 2,8	98,6 ± 2,9	99,2 ± 1,3	92,8 ± 10,6
Sensibilité (%)	74,7 ± 34,0	71,8 ± 41,8	89,7 ± 24,4	96,5 ± 7,8	93,9 ± 9,1
Indice de Dice (%)	73,6 ± 33,8	70,9 ± 41,7	87,3 ± 24,4	96,5 ± 7,5	93,2 ± 9,0

TABLE 5.4 – Évaluation du réseau de neurones proposé sur l'ensemble de données DSPT selon les cinq classes prédites : Tissu rénal (parenchyme + graisse), Tumeur rénale, Zone de coagulation, Instruments et Arrière-plan. Source : auteur.

Une évaluation complémentaire dans le contexte applicatif final de la RA en NPRA est réalisée. Pour rappel, la segmentation automatique est utilisée principalement sur les IC afin de réaliser un filtrage spatial des points clés détectés (cf. section 5.4.1). Dans l'ensemble DSPT, 58 images ayant servi d'IC sont isolées. Les résultats, présentés dans le tableau 5.5, montrent une amélioration notable de la sensibilité (+9,4 points) et de l'indice de Dice (+9,6 points) par rapport à l'évaluation générale. Cette évolution positive peut s'expliquer par des conditions de visibilité favorables, nécessaires au recalage 3D-2D de l'IC.

	Tissu rénal	Arrière-plan
Spécificité (%)	96,8 ± 3,2	87,1 ± 0,15
Sensibilité (%)	84,2 ± 18,0	98,3 ± 1,7
Indice de Dice (%)	83,4 ± 0,15	97,3 ± 2,0

TABLE 5.5 – Évaluation du réseau de neurones proposé sur des Images Clés (IC) selon les classes principales prédites : Tissu rénal (parenchyme + graisse) et Arrière-plan. Source : auteur.

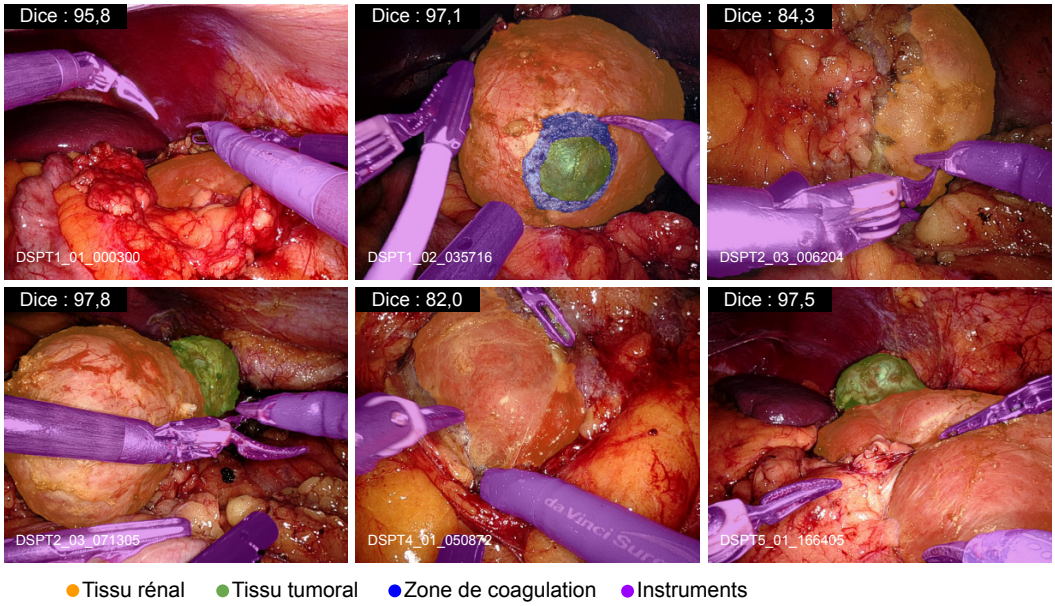


FIGURE 5.29 – Exemples de segmentations automatiques multiclasse en Néphrectomie Partielle Robot Assistée (NPRA). L'indice de Dice pour le tissu rénal est donné pour chaque image segmentée. Source : auteur.

Enfin, des optimisations sont appliquées aux poids des neurones par quantisation, en vue d'une utilisation en production dans un système de RA pour la NPRA (cf. section 5.4.5). Les performances selon différents moteurs d'inférence et représentations des poids sont présentées dans les tableaux 5.6 et 5.7 pour un ordinateur Ubuntu 24.04 équipé d'un processeur Intel Core™ i7-13620H et d'un GPU Nvidia RTX 4070M. Les résultats montrent qu'avec TensorRT en FP16, le temps d'inférence moyen est réduit à 2,70 millisecondes par image, soit un gain de 79 % par rapport à Pytorch 2.7.1, sans perte significative de performance.

	Pytorch 2.7.1	**Onnxruntime GPU**	**TensorRT FP32**
Temps d'inférence (ms)	12,99 ± 5,09	9,04 ± 0,12	5,44 ± 0,04
Poids du modèle (MB)	53,2	50,3	67,4

TABLE 5.6 – Évaluation de la performance (temps d'inférence évalué sur l'ensemble de données DSPT) du réseau de segmentation pour trois moteurs d'inférence différents, pour un ordinateur Ubuntu 24.04 équipé d'un processeur Intel Core™ i7-13620H et d'un GPU Nvidia RTX 4070M. Source : auteur.

	FP32	FP16	UINT8
Temps d'inférence (ms)	5,44 ± 0,04	2,70 ± 0,03	1,97 ± 0,02
Indice de Dice sur DSPT (%)	73,6 ± 33,8	73,6 ± 33,8	71,1 ± 41,2
Poids du modèle (MB)	67,4	32,5	14,1

TABLE 5.7 – Évaluation de la performance (temps d'inférence et indice de Dice évalués sur l'ensemble de données DSPT) du réseau de segmentation pour trois représentations différentes des poids des neurones, pour un ordinateur Ubuntu 24.04 équipé d'un processeur Intel Core™ i7-13620H et d'un GPU Nvidia RTX 4070M. Source : auteur.

5.11.2 Apport de la méthode dynamique

La méthode de suivi visuelle dynamique (cf. section 5.6) composée d'une BIC et d'un sélecteur de PIC est comparée à la méthode statique (cf. section 5.5) sur sept vidéos endoscopiques de myomectomie. Ces séquences haute définition, d'une résolution de 1920×1080 pixels et cadencées entre 25 et 60 Hz, couvrent une diversité de situations peropératoires incluant des mouvements rigides et non rigides, ainsi que des perturbations visuelles liées aux instruments, à la fumée et aux saignements. Le critère principal d'évaluation est le pourcentage d'images pour lesquelles l'organe cible, ici l'utérus, est correctement suivi. Les expériences ont été réalisées sur un ordinateur équipé d'Ubuntu 18.04, d'un processeur Intel Core i9-10900K et d'un GPU Nvidia RTX 2080Ti.

Sur le plan technique, cette première version du système s'appuie sur trois CS. Le CS1 qui évalue le résidu quadratique moyen de l'erreur de reprojection, le CS2 qui caractérise l'accélération de la caméra et le CS3 qui estime la netteté de l'image à partir de la variance du Laplacien. La performance du suivi visuel est monitorée par deux CD : le CD1, défini par le nombre de correspondances de points clés entre l'image courante et l'IC sélectionnée et le CD2, fondé sur le nombre d'inliers obtenus lors de l'estimation de pose par PnP RANSAC. Afin de limiter la charge computationnelle et de prévenir les ralentissements, la taille de la BIC est limitée à un maximum de 50 IC. Les paramètres et les seuils sont précisés en section 5.6.

Les résultats expérimentaux sont synthétisés dans le tableau 5.8. Ils mettent en évidence une amélioration globale du nombre d'images suivies grâce à la méthode dynamique, avec un gain moyen de 11,19 % par rapport à la méthode statique. L'ampleur de ce gain dépend toutefois de la nature des séquences chirurgicales et des mouvements de l'organe. L'amélioration est plus marquée en présence de déformations ou lorsque de larges zones intra-abdominales sont explorées.

Vidéo	Nb. Images	Suivi statique (%)	Suivi dynamique (%)	Différence (%)
1	1892	55,57	75,15	+19,58
2	1554	95,75	95,91	+0,19
3	2207	32,02	73,17	+41,15
4	1475	80,24	80,30	+0,06
5	5484	73,09	77,94	+4,84
6	514	85,78	96,09	+10,31
7	3250	96,12	98,34	+2,22

TABLE 5.8 – Comparaison quantitative des méthodes de suivi visuel statique et dynamique, en pourcentage d'images correctement suivies. Source : auteur.

A titre illustratif, les résultats détaillés pour la vidéo n°3 sont donnés en figure 5.30. Dans cette séquence, la caméra endoscopique et l'utérus subissent des déplacements extrêmes, à la fois verticaux et horizontaux, l'utérus est même observé de manière oblique en raison de sa manipulation. La méthode statique a échoué après 45 secondes, tandis que la méthode dynamique a maintenu le suivi pour toutes les images où l'organe est suffisamment rigide par rapport à son modèle peropératoire, hypothèse centrale de l'approche. En termes de taux de rafraîchissement, les deux méthodes sont équivalentes, avec des fréquences moyennes de 47 Hz et 46 Hz respectivement, largement supérieures au flux vidéo monoculaire à 25 Hz, sans différence perceptible sur le plan de la fluidité visuelle.

Ces résultats montrent que la mise à jour de la BIC améliore significativement la robustesse du suivi tout en gérant efficacement de nouveaux points de vue et des changements d'apparence au cours de l'intervention. Par ailleurs, il est important de noter que ces expériences, menées rétrospectivement sur des vidéos, ne tiennent pas compte de la réaction de l'utilisateur en cas d'échec du suivi visuel. En pratique, un chirurgien ayant besoin du guidage par RA adapterait sa manipulation pour revenir vers une configuration spatiale antérieurement augmentée, ce qui pourrait encore renforcer l'efficacité de la méthode dynamique. Cette hypothèse a été confirmée lors d'essais en conditions réelles au bloc opératoire (cf. section 2.2.3), où il est possible d'observer une adaptation réciproque du système et de l'utilisateur.

Enfin, la marge d'amélioration reste dépendante de la composante statique de BIC, constituée de données issues du recalage 3D-2D initial utilisant le SfM. Ces IC sont des points d'ancrage forts mais immuables, qui influencent directement l'initialisation du suivi et la qualité des résultats.

Le système de suivi dynamique est entièrement basé sur des composants logiciels. Il peut ainsi être intégré facilement à tous les dispositifs de chirurgie assistée par ordinateur exploitant un flux vidéo et être adapté en fonction des besoins grâce à l'ajout de critères supplémentaires.

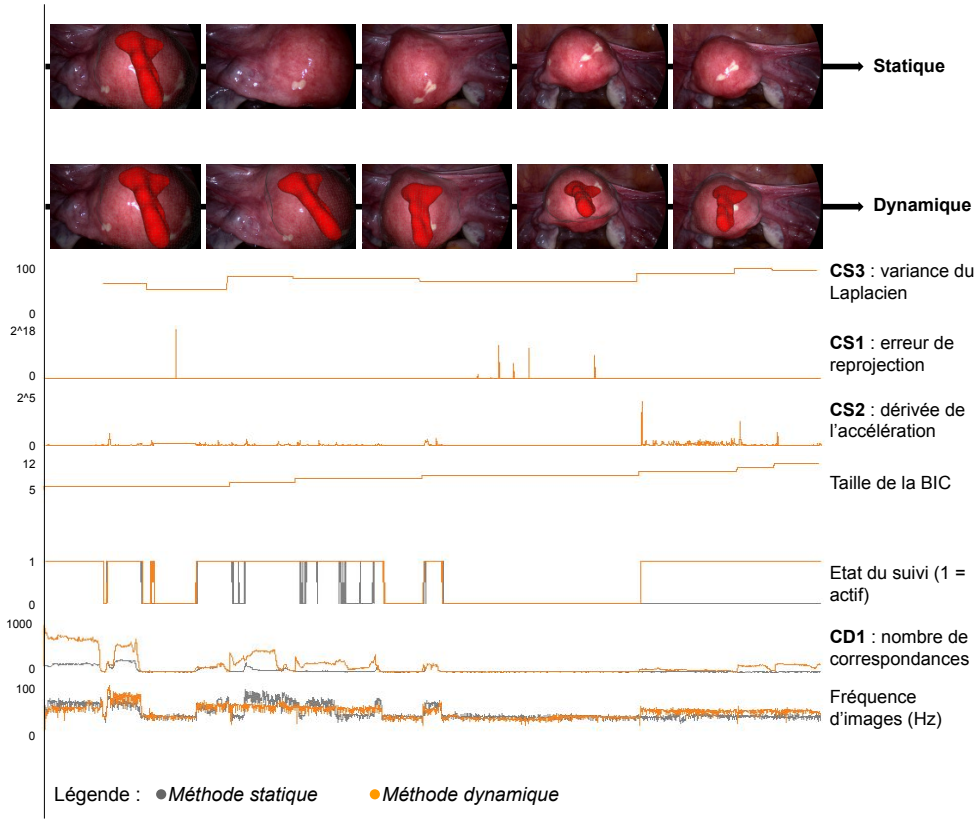


FIGURE 5.30 – Comparaison détaillée entre méthode statique et dynamique pour la séquence vidéo n°3. Les trois Critères de Sélection (CS) sont tracés en fonction du temps ainsi que l'évolution de la taille de la Base d'Images Clés (BIC), de l'état binaire du suivi, du Critère de Déclenchement (CD1) et de la performance en termes de fréquence d'image. Source : auteur.

5.11.3 Apport de la stéréovision

Nous avons évalué la méthode de suivi visuel dynamique stéréoscopique décrite en section 5.7 au travers d'une étude d'ablation portant sur les deux étapes constitutives en vue d'une application clinique en NPRA : (1) la segmentation sémantique du parenchyme rénale et (2) l'appariement contraint par la géométrie épipolaire gauche-droite. Parallèlement, une comparaison directe entre le suivi dynamique stéréoscopique et son pendant monoculaire est également réalisée.

L'évaluation s'appuie sur trois cas cliniques de NPRA provenant du système de télémanipulation Da Vinci Xi (Intuitive Surgical) équipé d'un endoscope stéréoscopique 30°. Un étalonnage géométrique stéréoscopique a été effectué sur la base d'un modèle pinhole, puis les paires d'images stéréoscopiques ont été rectifiées selon la procédure standard (cf. section 2.2.6). Ces expériences ont été réalisées sur un ordinateur Ubuntu avec un processeur Intel Core™ i7-10870H et un GPU Nvidia RTX 2060M. Les résultats présentés dans le tableau 5.9 correspondent aux moyennes calculées sur les trois cas cliniques.

Les analyses montrent que l'utilisation de la stéréovision augmente le pourcentage d'images suivies de 62,94 points de pourcentage par rapport au suivi monoculaire, soit une hausse relative de 209,8 %. L'intégration de la segmentation du parenchyme rénal améliore la robustesse, réduisant l'écart type du nombre de correspondances disponibles pour le suivi par un facteur d'environ sept. En l'absence de segmentation, la détection des points clés s'effectue sur l'intégralité de l'image ; si cela accroît le nombre brut de points disponibles après filtrage, cela ne signifie pas nécessairement une meilleure qualité de suivi, puisque ces points sont souvent associés à des structures hors surface rénale (cavité abdominale et instruments chirurgicaux) et ne contribuent pas systématiquement à une augmentation du ratio images suivies / images totales.

L'appariement gauche/droit via l'utilisation de la contrainte épipolaire s'avère efficace pour accroître le nombre d'images suivies, effet d'autant plus marqué lorsque la segmentation du rein n'est pas utilisée. Enfin, bien que tous les modes testés présentent des temps d'exécution compatibles avec une utilisation clinique en temps réel, la segmentation rénale demeure le composant le plus coûteux en calcul. Selon le moteur d'inférence retenu (cf. section 5.11.1), le temps nécessaire à la prédiction pour une image est compris entre 3 et 33 millisecondes, ce qui reste compatible avec une application temps réel.

Méthode	Seg. du rein	Appariement Gauche/Droite	Nb. correspondances	Images suivies (%)	Fréquence (Hz)
Stéréo	Non	Non	91 ± 101	62,16	84,61
Stéréo	Non	Oui	47 ± 26	92,94	86,29
Stéréo	Oui	Non	16 ± 13	59,35	27,35
Stéréo	Oui	Oui	27 ± 15	66,57	28,16
Mono	Non	Non	45 ± 51	30,00	150,62
Mono	Oui	Non	9 ± 7	18,65	36,10

TABLE 5.9 – Étude d'ablation des composants principaux de la méthode de suivi dynamique stéréoscopique et report des valeurs de performance pour un ordinateur Ubuntu avec un processeur Intel Core™ i7-10870H et un GPU Nvidia RTX 2060M. Source : auteur.

5.11.4 Apport de la correction de la perspective

Nous avons évalué la méthode de CP en détectant et en décrivant des points clés SIFT [Lowe 2004] dans les textures issues de l'aplatissement du maillage tridimensionnel peropératoire (cf. section 5.8) et en les comparant aux points clés extraits des images originales après correction de la distorsion. Les appariements sont calculés entre une IC unique, définie en position neutre, et plusieurs images acquises lors du mouvement de la caméra endoscopique. Afin d'assurer une comparaison équitable, seuls les points clés localisés à la surface du parenchyme rénal par la segmentation sémantique automatique (cf. section 5.4) sont pris en compte. Deux expériences sont menées, l'une sur des images semi-synthétiques et l'autre sur des images réelles de NPRA.

Les images semi-synthétiques sont générées à partir d'un modèle tridimensionnel reconstruit par SfM avec le logiciel open source Meshroom (Alice VISion) [Griwodz 2021], à partir d'une séquence réelle de NPRA, comme illustré dans la figure 5.31. Les déplacements de la caméra endoscopique stéréoscopique sont simulés sous le logiciel open source Blender, pour créer différents angles de vue. Ce cadre expérimental permet l'utilisation de données de vérité terrain, incluant les paramètres intrinsèques et extrinsèques de la caméra stéréoscopique, les poses de l'endoscope, les cartes de profondeur et les masques de segmentation sémantique. Une extension spécifique développée sous Blender a été utilisée pour produire ces données.

Les résultats obtenus sur les images semi-synthétiques sont présentés en figure 5.32. À mesure que l'angle relatif entre la caméra et l'IC neutre augmente, la distorsion de perspective se renforce et la qualité des appariements se dégrade nettement. L'introduction de la CP améliore le ratio et le nombre de correspondances correctes pour des angles supérieurs à 20°. En revanche, pour des angles plus

faibles, les images originales offrent de meilleures performances, car les traitements correctifs appliqués à l'image altèrent légèrement le signal et perturbent les descripteurs SIFT, bien que leur invariance reste suffisante pour de faibles changements de perspective. Au-delà de 20°, la CP devient toutefois déterminante.

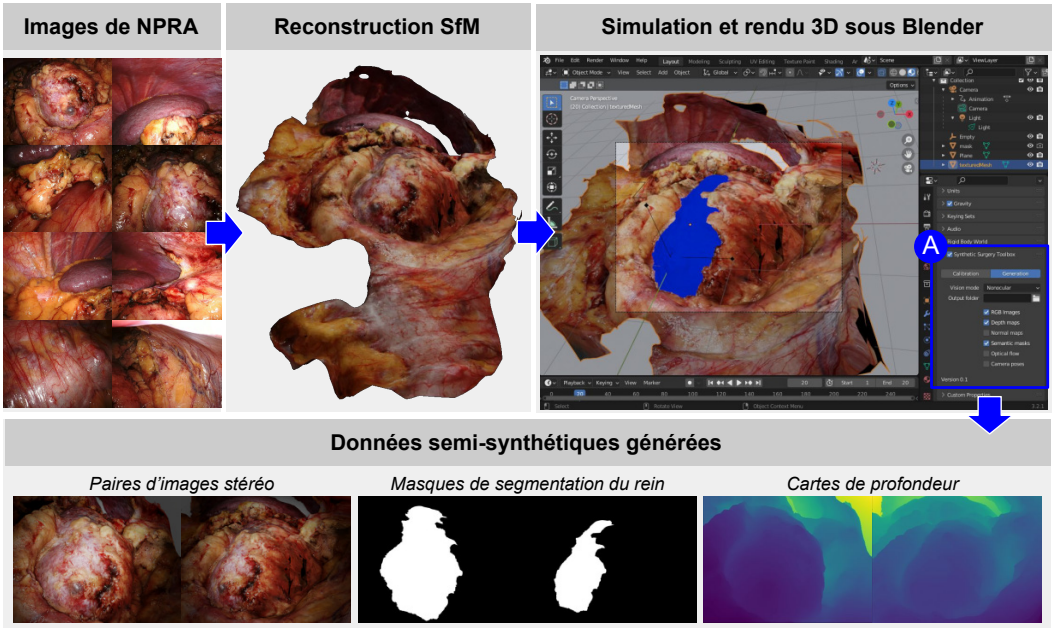


FIGURE 5.31 – Proposition de génération de données semi-synthétiques à partir d'images réelles de Néphrectomie Partielle Robot Assistée (NPRA) et d'une reconstruction tridimensionnelle via Structure-from-Motion (SfM). (A) Extension Blender spécialement conçue pour les images stéréoscopiques et la génération de données semi-synthétiques en Chirurgie Mini-Invasive (CMI). Source : auteur.

Parallèlement, une évaluation de l'approche de CP avec la méthode de détection et description basée sur l'apprentissage profond SuperGlue [Sarlin 2020] est réalisée. Les performances obtenues sont comparables à celles de SIFT [Lowe 2004], mais le coût computationnel est en moyenne six fois plus élevé, avec un taux de rafraîchissement inférieur à 15 Hz, ce qui rend cette approche incompatible avec une exécution en temps réel.

Ainsi, l'utilisation de la CP est essentielle pour les appariements complexes impliquant des angles de vue supérieurs à 20°. Dans le cas de faibles angles, l'impact de la CP reste marginal puisque le nombre de points clés disponibles est déjà suffisant au suivi. Comme illustré en figure 5.32, l'extraction et la description des caractéristiques dans le domaine aplati réduisent la distance entre descripteurs correspondants, renforçant l'invariance aux déformations de perspective. En moyenne, la CP accroît donc la robustesse et la discriminabilité des descripteurs classiques comme SIFT.

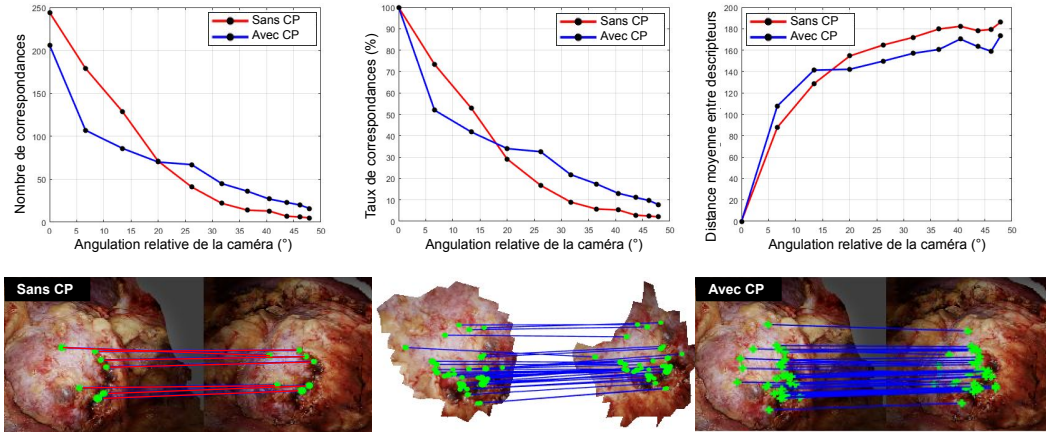


FIGURE 5.32 – (première ligne) Évaluation de la méthode de Correction de la Perspective (CP) sur des images semi-synthétiques, en termes de nombre de correspondances correctes, de ratio de correspondances correctes et de distance moyenne du descripteur SIFT par rapport à l'angle relatif de la caméra. (seconde ligne) Exemple avec un angle de caméra relatif de $43,7°$: 7 correspondances correctes (sur 10) obtenues sans CP et 31 correspondances correctes (sur 31) avec CP. Source : auteur.

Les mêmes expériences ont été réalisées sur des images réelles de NPRA, dont les résultats sont visibles en figure 5.33. Pour des angles de caméra importants entre l'IC et l'image courante, la CP multiplie par plus de trois le nombre de correspondances correctes. Dans l'ensemble des expériences, semi-synthétiques comme réelles, nous observons que la CP diminue légèrement les performances d'appariement dans les cas avec une faible angulation relative, mais améliore considérablement les performances d'appariement lorsque la distorsion de la perspective prédomine.

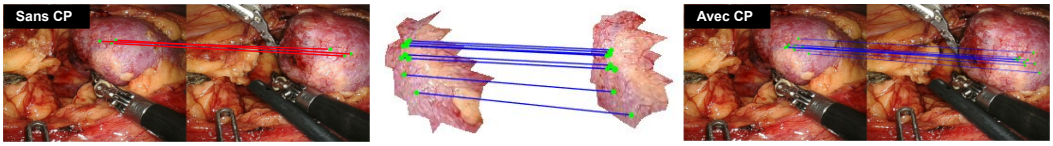


FIGURE 5.33 – Évaluation de la méthode de Correction de la Perspective (CP) sur des images réelles de Néphrectomie Partielle Robot Assistée (NPRA). 3 correspondances correctes (sur 4) sans CP et 10 correspondances correctes (sur 10) avec CP. Source : auteur.

5.11.5 Évaluation en conditions peropératoires

La méthode de suivi dynamique a été utilisée de manière rétrospective en per-opératoire, c'est-à-dire sans guidage direct du geste mais avec génération en temps réel des images augmentées, sur plus de vingt-cinq cas cliniques de NPRA dans le cadre des essais de faisabilité de RA (cf. section 2.2.3). Des exemples d'images augmentées générées au bloc opératoire sont données en figure 5.34.

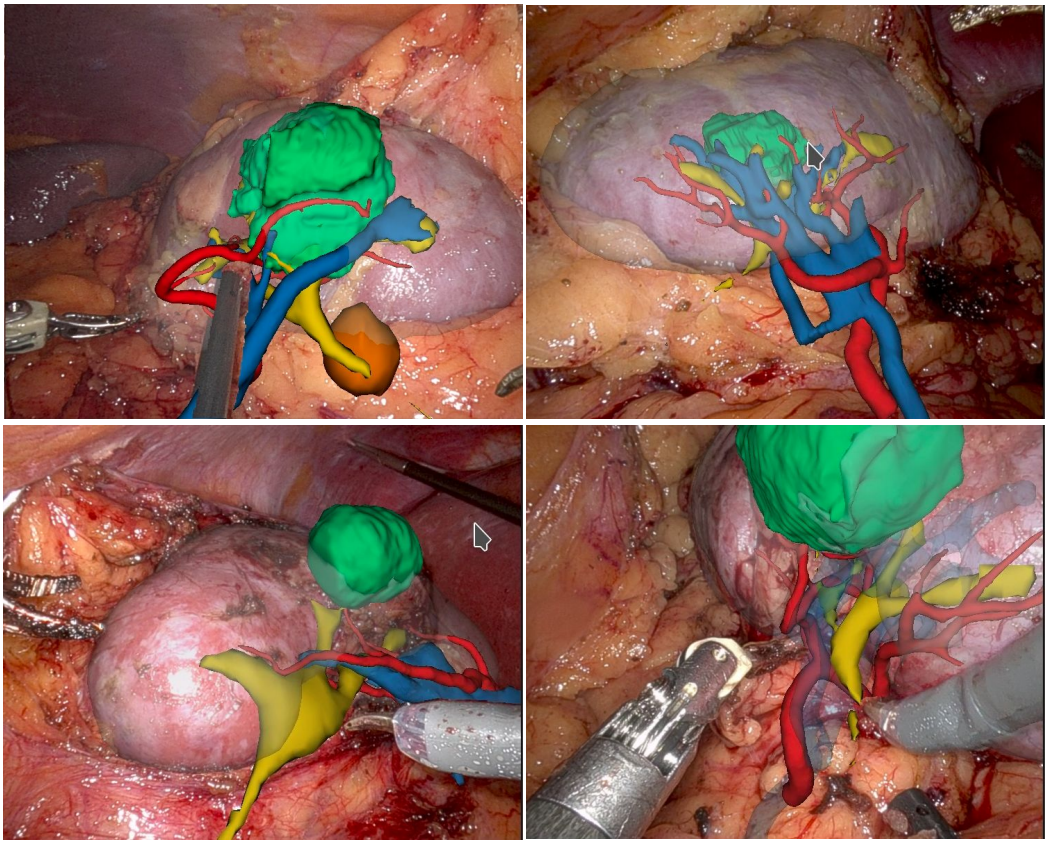


FIGURE 5.34 – Images augmentées en Néphrectomie Partielle Robot Assistée (NPRA) guidée par Réalité Augmentée (RA). Source : auteur.

Trois types d'évaluations ont été conduites. Tout d'abord, une évaluation in situ au bloc opératoire a permis de recueillir les retours des chirurgiens sur la qualité du recalage 3D-2D du jumeau numérique sur la vue peropératoire, ainsi que sur leur charge mentale via les échelles NASA-TLX [Hart 1988] et Surg-TLX [Wilson 2011], au cours des sessions d'essais de faisabilité. Bien que le système de RA ne soit pas encore entièrement automatique, son utilisation devient de plus en plus aisée, au point que le chirurgien principal peut désormais l'employer de manière autonome.

Des indicateurs temporels ont été collectés sur les deux étapes préalables indispensables à la RA, à savoir l'étalonnage géométrique de l'endoscope et le recalage 3D-2D manuel. Actuellement, ces procédures nécessitent respectivement moins d'une minute et entre trois et cinq minutes, ce qui demeure compatible avec les contraintes organisationnelles du bloc opératoire au regard du bénéfice clinique attendu de la RA.

Ensuite, une comparaison quantitative a été conduite entre les trajectoires tridimensionnelles issues du suivi visuel dynamique et une semi-vérité terrain labellisée de manière semi-automatique (cf. section 4.6.2). Sept vidéos peropératoires haute résolution (1280 × 1024 pixels, 60 Hz) de NPRA réalisées à l'aide du système de télémanipulation chirurgicale Da Vinci Xi (Intuitive Surgical) sont exploitées, soit 299 659 images recalées. Le jumeau numérique ainsi que les données d'étalonnage correspondantes sont disponibles. Les données ont été labellisées de manière rétrospective par une équipe de chirurgiens urologues, comme détaillé en sections 4.6.2. Les trajectoires tridimensionnelles obtenues sont comparées en termes d'erreurs de pose en rotation et en translation. Les résultats sont donnés pour les six degrés de liberté dans le tableau 5.10 et un exemple visuel est présenté en figure 5.35. Lors du suivi visuel dynamique, les erreurs moyennes en rotation et en translation sont respectivement inférieures à un degré et à cinq millimètres.

	Rx (°)	Ry (°)	Rz (°)	Tx (mm)	Ty (mm)	Tz (mm)
Moyenne	0,51	0,74	0,42	3,26	4,93	3,36
Ecart-type	0,49	0,61	0,43	3,31	4,17	3,61
Minimum	0,00	0,00	0,00	0,00	0,00	0,00
Maximum	3,15	2,89	1,92	14,74	14,99	14,79

TABLE 5.10 – Résultats expérimentaux issus de la comparaison des poses de semi-vérité terrain avec celles du suivi visuel dynamique pour sept cas de Néphrectomie Partielle Robot Assistée (NPRA). Source : auteur.

Enfin, une évaluation quantitative a porté sur l'alignement des contours tumoraux à la surface du parenchyme rénal. Ce test vise à objectiver la pertinence du suivi visuel dynamique en conditions réelles et représentatives de la NPRA. Nous avons ciblé spécifiquement l'erreur de recalage de la tumeur, paramètre critique pour la fiabilité du guidage peropératoire et la précision du geste chirurgical. Quatre cas de tumeurs exophytiques (ID 1, 3, 5 et 6) sont sélectionnés. La comparaison est alors directe entre le contour projeté de la tumeur (issu du modèle augmenté) et le contour réel de la tumeur observé dans l'image peropératoire. Quarante mesures, dix par cas, révèlent une erreur moyenne d'alignement du contour de la tumeur de $5,6 \pm 2,3$ millimètres, ce qui est comparable aux erreurs d'alignement manuel observées dans des conditions non contraintes, généralement comprises entre trois et

neuf millimètres (cf. section 4.7.2) et inférieure au seuil clinique de dix millimètres (cf. section 4.2).

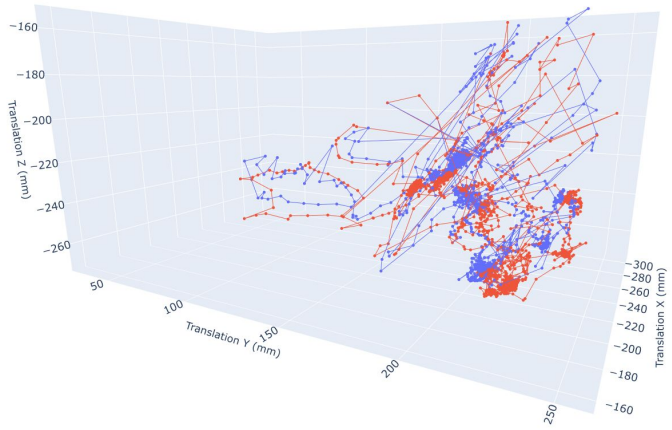


FIGURE 5.35 – Exemple de trajectoires tridimensionnelles de la caméra endoscopique issue d'une semi-vérité terrain (en rouge) et du suivi visuel dynamique (en bleu) lors d'une Néphrectomie Partielle Robot Assistée (NPRA). Source : auteur.

En complément, une évaluation pré-clinique est en cours sur phantoms, modèles de reins physiques à taille réelle obtenus par fabrication additive. Les données ont été collectées dans des conditions très proches des conditions réelles via l'utilisation d'un endoscope stéréoscopique 30° Da Vinci Xi (Intuitive Surgical). Les séquences expérimentales simulent des trajectoires de caméra proches de celles rencontrées en peropératoire. Les scénarios d'essai visent à mesurer l'erreur de recalage en fonction de l'amplitude des déformations de l'organe cible, générées virtuellement sur le maillage numérique.

L'expérience utilisateur est également abordée sous l'angle de la perception visuelle, facteur déterminant pour l'adoption et la performance de la RA (cf. section 3.4.3). Afin de renforcer la perception de profondeur et d'assurer une visualisation précise de la pointe des instruments, ceux-ci sont restitués systématiquement au premier plan, comme illustré en figure 5.36. Sur le plan technique, un réseau de neurones de segmentation sémantique (cf. section 5.4) produit un masque des instruments pour chaque image du flux vidéo. Ce masque est chargé dans la mémoire tampon du stencil buffer OpenGL afin d'exclure les zones segmentées du rendu graphique. En d'autre termes, le rendu du maillage du jumeau numérique est découpé en fonction de la segmentation.

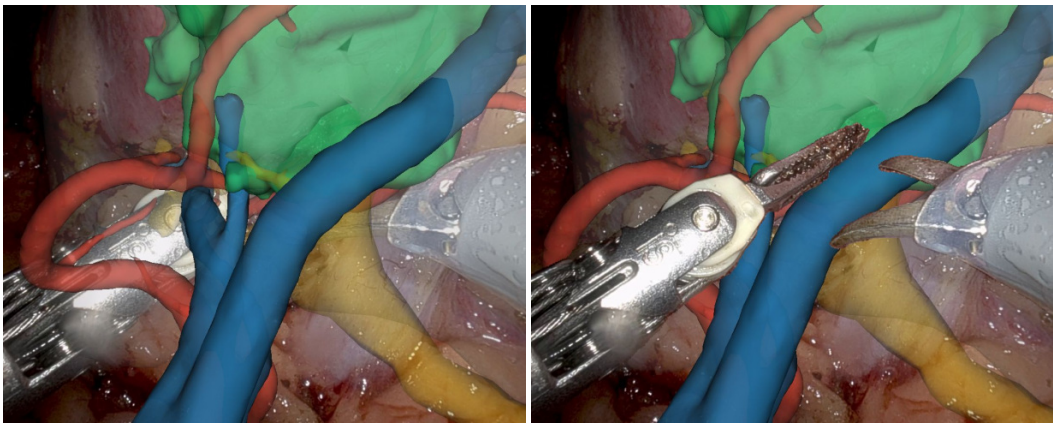


FIGURE 5.36 – Deux modes d'affichage de la Réalité Augmentée (RA) en Néphrectomie Partielle Robot Assistée (NPRA) : à gauche sans aucun post-traitement ; à droite avec post-traitement de l'instrumentation chirurgicale restituée au premier plan. Source : auteur.

Sur les aspects de perception, une autre évaluation est en cours pour examiner l'influence du mode de visualisation (monoculaire 2D ou binoculaire 3D) sur la performance de recalage initial. Un groupe de chirurgiens devra réaliser des recalages sur des séquences vidéo préenregistrées dans la console du Da Vinci Xi (Intuitive Surgical), avec affichage en RA dans le TilePro (Intuitive Surgical). Les mesures porteront sur la précision, la répétabilité, la reproductibilité et le temps nécessaire à la réalisation du recalage.

5.12 Conclusion

Nous avons proposé et évalué une méthode de suivi visuel d'organe en temps réel, qui utilise une base d'images clés dynamique, spécifiquement conçue pour la chirurgie augmentée. L'approche intègre un gestionnaire de base d'images qui assure la cohérence temporelle et spatiale du suivi, ainsi qu'un sélecteur automatique d'images clés qui permet d'adapter la base au cours de l'intervention. L'extension à la stéréovision, via l'utilisation de la géométrie épipolaire et l'introduction d'une correction de la perspective représentent des solutions techniques intéressantes afin d'améliorer la robustesse du système face aux variations de point de vue.

Les expérimentations confirment la faisabilité et l'intérêt clinique de l'approche. L'évaluation inter-annotateurs de la segmentation du parenchyme rénal, menée sur 454 images issues de 29 NPRA, a montré des indices de Dice compris entre 81 % et 88 %, avec des temps de labellisation variant de 31 secondes à plus de 4 minutes selon les profils. Le réseau de segmentation multiclasse, entraîné sur 41 160 images, atteint une spécificité moyenne de 97,4 %. Sur les images clés, la sensibilité et l'indice de Dice sont respectivement de 84,2 % et 83,4 % pour le tissu rénal. Après quantisation,

le temps d'inférence moyen descend à 2,70 millisecondes par image avec TensorRT FP16, soit un gain de 79 %.

La méthode dynamique de suivi visuel améliore en moyenne de 11,19 % le nombre d'images suivies par rapport à la méthode statique, avec un gain maximal de +41,15 % selon les séquences. L'usage de la stéréovision accroît ce taux de 62,94 points, ce qui correspond à une augmentation relative de 209,8 %. La correction de la perspective apporte un bénéfice net pour des angles supérieurs à 20°, multipliant par plus de trois le nombre de correspondances correctes sur des données réelles.

En conditions peropératoires, la méthode a été appliquée rétrospectivement sur plus de 25 cas cliniques de NPRA. Sur sept vidéos haute résolution (299 659 images recalées), l'erreur moyenne de pose par rapport à une semi-vérité terrain est inférieure à 1° en rotation et 5 millimètres en translation. L'alignement des contours tumoraux exophytiques présente une erreur moyenne de $5,6 \pm 2,3$ millimètres, inférieure au seuil clinique de 10 millimètres. Les temps opératoires consacrés à l'étalonnage géométrique et au recalage 3D-2D manuel restent courts, respectivement inférieurs à une minute et compris entre trois et cinq minutes, ce qui valide l'intégrabilité du système au flux opératoire.

Ces résultats ouvrent de nombreuses perspectives de développement méthodologique et applicatif. De nouveaux CS, comme une couverture spatiale explicite des IC, pourraient renforcer la robustesse et la pertinence de la méthode dynamique, en améliorant la représentativité et la diversité des vues de référence. Dans cette optique, l'identification des PIC pourrait être confiée à un réseau de neurones dédié, sans nécessiter une description explicite des CS. Ce réseau serait conçu pour estimer, dans sa forme la plus simple, un état binaire par classification ou, dans une version plus avancée utilisant la régression, un score de pertinence d'ajout de la PIC à la BIC.

La prise en compte des déformations non rigides, particulièrement complexe en conditions temps réel, pourrait quant à elle s'appuyer sur des approches d'apprentissage profond, notamment à travers des réseaux de régression capables d'estimer directement les champs de déformation, ou des architectures telles que CoTracker qui permettent de prédire les correspondances locales dans l'espace bidimensionnel de l'image ou tridimensionnel de la scène chirurgicale. L'intégration de telles méthodes ouvrirait la voie à une adaptation plus fine du jumeau numérique augmenté au comportement tissulaire réel, tout en maintenant une compatibilité avec les contraintes temps réel indispensables à la RA peropératoire.

Les approches d'apprentissage continu, de transfert de domaine et d'active learning représentent également une piste particulièrement prometteuse. Elles permettraient au système de s'adapter progressivement à des conditions variées, à partir de quantités de données limitées ou issues de simulations photoréalistes. De plus, l'ajustement dynamique des RNA en fonction des différentes phases chirurgicales constitue un levier d'amélioration notable pour réduire la variabilité inter-procédure. Ces stratégies contribueraient ainsi à une adaptation plus fine aux variations anatomiques, instrumentales ou contextuelles.

Par ailleurs, le suivi visuel d'organes et d'instruments ouvre un champ de recherche plus vaste vers des solutions de suivi multi-objet. Celles-ci visent à gérer simultanément plusieurs cibles, qu'elles soient indépendantes ou liées par des contraintes cinématiques. L'objectif serait de concevoir un système de suivi capable d'assurer la continuité spatiale et temporelle d'un ensemble d'entités interagissant au sein de la scène chirurgicale. Cette évolution nécessitera l'élaboration de modèles d'interaction et de contraintes géométriques ou mécaniques reliant les trajectoires des différents objets, par exemple entre un organe et ses structures vasculaires qui le perfusent. Ces modèles permettront d'intégrer une cohérence physique et anatomique dans le processus de suivi.

Enfin, l'ensemble de ces améliorations, couplé aux évaluations précliniques actuellement en cours, jouera un rôle déterminant dans la validation et l'intégration future de la méthode en conditions in vivo, ouvrant la voie à un suivi visuel chirurgical adaptatif et contextuellement cohérent.

Chapitre 6

Conclusion et travaux futurs

Ce manuscrit a exposé l'ensemble de nos contributions relatives à la chaîne de traitement amont ainsi qu'au suivi visuel monoculaire et stéréoscopique, en conditions synchrones et asynchrones, au sein d'un environnement chirurgical dynamique. Ces travaux s'inscrivent dans la continuité des développements d'outils de CAO et de RA dédiés à la CMI, avec pour application clinique principale la NPRA. L'objectif général poursuivi a consisté à fiabiliser et automatiser les étapes essentielles du pipeline de RA chirurgicale, en levant trois verrous scientifiques : la synchronisation stéréoscopique, le recalage 3D–2D automatique et le suivi visuel dynamique d'organe en temps réel. La dimension industrielle, fortement présente dans le cadre de cette thèse CIFRE, a orienté la démarche scientifique vers une finalité produit et utilisateur, à moyen terme. Les développements décrits au fil des chapitres traduisent ainsi une trajectoire scientifique cohérente, depuis la maîtrise des flux d'acquisition jusqu'à la stabilité du rendu augmenté, tout en ouvrant des perspectives de recherche à la fois méthodologiques, techniques et cliniques.

Sommaire

6.1 Conclusions 281
6.1.1 Objectifs et trajectoire des travaux 281
6.1.2 Synchronisation de flux vidéo stéréoscopiques en chirurgie endoscopique 281
6.1.3 Estimation automatique d'une pose globale par apprentissage profond 282
6.1.4 Suivi visuel d'organe en temps réel par points clés 283
6.2 Travaux futurs 283
6.2.1 Perspectives sur la synchronisation de flux stéréoscopiques . . 283
6.2.2 Perspectives sur le recalage 3D-2D automatique 284
6.2.3 Perspectives sur le suivi visuel d'organe 285

6.1 Conclusions

6.1.1 Objectifs et trajectoire des travaux

L'ambition initiale de ces travaux était de contribuer à la définition et à la mise en œuvre d'un pipeline complet de RA pour la CMI des organes raisonnablement rigides, en particulier le rein dans le contexte de la NPRA. Ce pipeline devait assurer la cohérence temporelle, spatiale et dynamique entre les données anatomiques préopératoires (jumeau numérique du patient) et la scène endoscopique peropératoire. Sa conception reposait sur la maîtrise successive de trois étapes interdépendantes : l'acquisition et la synchronisation des flux vidéo stéréoscopiques, le recalage initial 3D-2D et le suivi visuel d'organe.

La stratégie méthodologique adoptée s'est appuyée sur une progression incrémentale, ou chaque contribution a été conçue comme une composante complémentaire visant la consolidation du pipeline global. Les premiers travaux décrits ont porté sur la fiabilisation des données d'entrée, notamment par la synchronisation temporelle des flux stéréoscopiques. Le second axe a ciblé la mise en correspondance spatiale entre le jumeau numérique préopératoire et la scène chirurgicale réelle, via un recalage 3D-2D automatique initial. Enfin, le troisième axe a assuré le maintien de l'alignement au cours du temps grâce à un suivi visuel robuste et en temps réel, garantissant la stabilité et la continuité de la superposition pendant l'intervention.

Ces travaux s'inscrivent dans une démarche expérimentale pragmatique, visant non seulement la validation algorithmique mais aussi la faisabilité clinique. Chaque composant développé a été évalué sur des données réelles, attestant de la pertinence des choix méthodologiques. Ce cheminement global a permis de structurer une chaîne complète de traitement d'images peropératoires, intégrable dans un prototype fonctionnel de RA chirurgicale. L'objectif n'était pas d'explorer toutes les dimensions possibles du pipeline global, mais de tracer une trajectoire expérimentale claire vers un système de RA clinique adapté au contexte applicatif de la NPRA.

6.1.2 Synchronisation de flux vidéo stéréoscopiques en chirurgie endoscopique

Le premier axe, présenté au chapitre 3, portait sur la synchronisation des flux gauche et droit d'une caméra endoscopique stéréoscopique. Ce problème, souvent sous-estimé, influe directement sur la qualité de la perception tridimensionnelle et sur la fiabilité des traitements de stéréovision. Une désynchronisation, même de quelques millisecondes, peut induire des erreurs significatives d'estimation des paramètres extrinsèques d'un système stéréoscopique ou de calcul de la disparité, et compromettre les performances du recalage ou du suivi visuel.

Nous avons proposé un mécanisme de synchronisation automatique exploitant la composante spéculaire des images stéréoscopiques. Cette approche repose sur la détection et la description des taches spéculaires visibles dans les flux gauche et droit, et permet la définition de SCGD robustes aux variations de texture et d'illumination. Un ensemble de données dédié, acquis en conditions réelles de NPRA et de RHRA,

a permis d'évaluer les performances du système dans un contexte peropératoire réaliste. Les résultats ont mis en évidence la capacité du SCGD spéculaire à corriger efficacement les désalignements temporels avec une marge d'erreur inférieure à la résolution temporelle des caméras chirurgicales actuelles, assurant ainsi l'utilisation de flux synchrones dans les algorithmes de vision tridimensionnelle avals.

À notre connaissance, il s'agit de la première méthode de synchronisation entièrement automatique fondée sur l'exploitation de caractéristiques photométriques spéculaires intrinsèques à la scène endoscopique. Sa robustesse et son indépendance vis-à-vis du matériel d'acquisition en font une solution générique et industrialisable.

6.1.3 Estimation automatique d'une pose globale par apprentissage profond

Le second axe, détaillé au chapitre 4, a concerné le recalage 3D-2D automatique. Cette étape du pipeline de RA chirurgicale assure la mise en correspondance initiale entre le jumeau numérique préopératoire et une image peropératoire. Les approches traditionnelles de recalage 3D-2D peropératoires sans marqueur, reposent généralement sur des primitives anatomiques explicites ou sur des interactions manuelles, impliquant soit une détection précise de structures anatomiques souvent peu discriminantes en NP, soit une dépendance à une intervention humaine chronophage, peu compatible avec la réalité clinique.

Pour surmonter ces limitations, nous avons conçu une approche sans marqueur basée sur l'apprentissage profond bout en bout, qui permet l'estimation directe de la pose tridimensionnelle de l'organe cible à partir d'images monoculaires. L'architecture proposée combine un encodeur convolutionnel à un module de régression de pose global, entraîné sur des données annotées issues d'interventions réelles de NPRA. Ce réseau de neurones apprend implicitement la géométrie et l'apparence de l'organe, sans recourir à des primitives explicites, et produit une estimation plausible même dans des conditions d'éclairage complexes ou en présence d'outils chirurgicaux.

Les évaluations menées sur un ensemble de données multi-patients restreint (7 chirurgies) ont montré que la précision atteinte permet une superposition en RA conforme aux exigences de la CGI. Ces résultats préliminaires démontrent la faisabilité d'un recalage entièrement automatique, bien qu'un renforcement de la robustesse et une validation clinique élargie demeurent nécessaires.

6.1.4 Suivi visuel d'organe en temps réel par points clés

Le troisième axe, exposé au chapitre 5, a porté sur le suivi visuel d'organe, maillon final du pipeline de RA, qui garantit la cohérence spatiale de l'augmentation au cours du geste chirurgical. L'approche proposée repose sur une BIC dynamique, qui permet d'assurer un suivi stable et adaptatif dans le temps. Cette méthode utilise le suivi par appariement de points clés entre l'image courante et une IC de la BIC. La BIC est mise à jour de manière incrémentale via l'ajout automatique d'IC, afin de gérer efficacement les variations de texture, de perspective et les occultations partielles en contexte peropératoire.

L'intégration d'une extension stéréoscopique et d'une correction automatique de la perspective a permis d'améliorer la fidélité de l'estimation des mouvements relatifs entre la caméra et l'organe. En complément, une segmentation automatique du parenchyme rénal, obtenue par apprentissage profond, a permis d'accroître la spécificité du suivi et de réduire la sensibilité aux éléments de l'environnement opératoire. Les validations sur données peropératoires ont confirmé la stabilité du suivi sur des séquences prolongées et la cohérence de la superposition du jumeau numérique sur l'image endoscopique dans des situations dynamiques complexes.

Ces résultats confirment la possibilité d'un suivi entièrement automatique et en temps réel, sans recalage intermédiaire, et constitue donc une étape supplémentaire vers une RA chirurgicale cliniquement exploitable, fluide et avec une dérive temporelle maîtrisée.

6.2 Travaux futurs

Les perspectives ouvertes par ces travaux s'inscrivent dans une continuité naturelle visant à enrichir, généraliser et intégrer les modules développés dans un système peropératoire complet de CGI. Les directions principales portent sur l'optimisation et la généralisation de la synchronisation aux sources multimodales, l'enrichissement du recalage par la prise en compte des déformations locales et la gestion des cas sans visibilité directe de l'organe cible, ainsi que l'évolution du suivi vers une implémentation adaptée aux cibles multiples, intégrant des faisceaux d'informations multi-échelle.

6.2.1 Perspectives sur la synchronisation de flux stéréoscopiques

Concernant la synchronisation, les travaux futurs pourront viser à élargir le cadre initial limité aux flux stéréoscopiques vers une approche véritablement multimodale et interopérable, apte à gérer de manière cohérente les flux issus de différentes modalités peropératoires comme la fluorescence ou l'échographie.

Une première orientation pourra consister à développer un réseau de neurones capable d'estimer directement le décalage temporel en temps réel entre canaux stéréoscopiques, sans recourir à une détection préalable. Une extension ultérieure vers la synchronisation de flux hétérogènes pourrait s'appuyer sur les corrélations spatio-

temporelles entre structures anatomiques communes ou éléments instrumentaux. L'intégration de ce module de synchronisation au sein des plateformes CAO représente par ailleurs une étape décisive vers sa validation et son adoption.

Face à la multiplication des systèmes de CGI et à la croissance du nombre de flux vidéo chirurgicaux, l'interopérabilité des solutions devient urgente. La standardisation des protocoles de partage vidéo pourrait offrir une voie vers un système global en boucle fermée, minimisant par conception les désalignements induits par le matériel ou le réseau, et garantissant ainsi une cohérence spatio-temporelle continue. Une autre approche consisterait à centraliser les applications de CGI sur une plateforme unifiée, simplifiant l'intégration et la maintenance des composants.

6.2.2 Perspectives sur le recalage 3D-2D automatique

Les perspectives liées au recalage concernent la transition d'un modèle rigide vers un recalage pleinement adaptatif, capable d'intégrer les déformations locales et physiologiques observées au cours de l'intervention. La prise en compte explicite de ces déformations, qu'elles résultent de la respiration, de la manipulation instrumentale ou des variations de pression intra-abdominale, demeure un défi scientifique et technologique majeur.

Une première orientation de recherche consisterait à coupler les méthodes d'apprentissage profond existantes à des modèles biomécaniques simplifiés ou des réseaux de neurones pour la simulation de déformations basées sur la physique. Ce couplage permettrait de mieux contraindre la pose estimée par des régularisations anatomiques ou tissulaires issues des modèles préopératoires, tout en préservant la vitesse d'inférence nécessaire au temps réel.

Une seconde voie viserait à exploiter davantage la topologie tridimensionnelle des structures anatomiques plutôt que de se limiter à une analyse essentiellement planaire des textures.

Une troisième orientation porterait sur le recalage sans visibilité directe de l'organe cible, reposant sur l'exploitation d'indices contextuels visibles tels que les instruments, organes adjacents ou zones anatomiques de référence. Ce scénario se rencontre typiquement en NP avant le dégraissage, où un guidage par RA pourrait faciliter l'approche de la tumeur.

Enfin, le couplage d'une modalité complémentaire comme l'échographie peropératoire aux informations visuelles issues de l'endoscope pourrait contribuer à mieux contraindre le recalage 3D–2D. Une telle combinaison multimodale permettrait d'enrichir la compréhension spatiale et d'améliorer la robustesse du processus, notamment en présence de zones peu texturées ou temporairement occultées. Néanmoins, en NP, l'usage de cette modalité d'imagerie demeure encore limité, en raison du surcoût financier associé et de l'expertise d'interprétation requise.

Pour aller plus loin, les difficultés inhérentes à l'évaluation sur données réelles devront également être surmontées afin de valider expérimentalement et cliniquement les méthodes développées, condition indispensable à leur transfert vers une application peropératoire fiable et automatisée.

6.2.3 Perspectives sur le suivi visuel d'organe

Le suivi visuel d'organe ouvre un vaste champ de recherche vers des solutions de suivi capables de gérer simultanément plusieurs cibles anatomiques ou instrumentales, qu'elles soient indépendantes ou reliées par des contraintes cinématiques. L'objectif sera de concevoir un système de suivi multi-objet, coopératif et contextuel, capable d'assurer la continuité spatiale et temporelle d'un ensemble d'entités évoluant dans la scène chirurgicale.

Cette évolution nécessitera l'élaboration de modèles d'interaction et de contraintes cinématiques permettant de relier les trajectoires des différents objets, qu'il s'agisse par exemple des relations entre un rein et ses structures vasculaires ou entre un instrument et la zone opérée.

Par ailleurs, l'intégration de faisceaux d'informations multi-échelle, à la fois contextuelle (vue globale de la scène) et locale (texture et reliefs spécifiques), représente une perspective pour améliorer la robustesse et la résilience du suivi dans des environnements complexes et dynamiques.

Enfin, le couplage entre les méthodes classiques de vision par ordinateur et les méthodes d'apprentissage profond semble prometteur. Dans les approches actuelles, ce couplage est le plus souvent réalisé sous forme de chaînage, par exemple en utilisant la sortie d'un réseau de neurones comme entrée à une étape ultérieure de détection ou d'appariement de points clés. Une intégration plus intrinsèque permettrait toutefois d'exploiter pleinement leurs complémentarités, en conjuguant l'explicabilité et la stabilité métrologique des méthodes traditionnelles avec la rapidité de calcul et la puissance de représentation de l'apprentissage profond.

À l'instar du recalage, l'établissement d'une méthodologie rigoureuse d'évaluation sur données réelles demeure un axe essentiel de progression. Cette étape est indispensable pour valider la robustesse, la précision et la transférabilité clinique des approches de suivi, et ainsi rapprocher davantage la RA chirurgicale d'une intégration complète dans la pratique chirurgicale quotidienne.

En conclusion, les travaux présentés dans cette thèse posent les bases méthodologiques et expérimentales d'un pipeline complet de RA pour la CMI abdomino-pelvienne, applicable aux organes raisonnablement rigides tel que le rein. Les perspectives identifiées ouvrent la voie à des développements intégrant plus étroitement l'intelligence artificielle, la modélisation biomécanique et la vision par ordinateur, vers un futur où la RA chirurgicale deviendra un outil de guidage peropératoire standardisé, fiable et profondément intégré à la pratique chirurgicale.

Annexes

Annexe A

Liste des Diplômes d'Études Spécialisées (DES) définis par l'arrêté du 21 avril 2017

Diplômes d'études spécialisées de la discipline médicale :

- Anatomie et cytologie pathologiques ;
- Anesthésie-réanimation ;
- Dermatologie et vénéréologie ;
- Endocrinologie-diabétologie-nutrition ;
- Génétique médicale ;
- Gériatrie ;
- Gynécologie médicale ;
- Hématologie ;
- Hépato-gastro-entérologie ;
- Maladies infectieuses et tropicales ;
- Médecine cardiovasculaire ;
- Médecine d'urgence ;
- Médecine et Santé au travail ;
- Médecine générale ;
- Médecine intensive-réanimation ;
- Médecine interne et immunologie clinique ;
- Médecine légale et expertises médicales ;
- Médecine nucléaire ;
- Médecine physique et de réadaptation ;
- Médecine vasculaire ;
- Néphrologie ;
- Neurologie ;
- Oncologie ;
- Pédiatrie ;
- Pneumologie ;
- Psychiatrie ;
- Radiologie et imagerie médicale ;
- Rhumatologie ;
- Santé publique.

Diplômes d'études spécialisées de la discipline chirurgicale :

- Chirurgie maxillo-faciale ;
- Chirurgie orale ;
- Chirurgie orthopédique et traumatologique ;
- Chirurgie pédiatrique ;
- Chirurgie plastique, reconstructrice et esthétique ;
- Chirurgie thoracique et cardiovasculaire ;
- Chirurgie vasculaire ;
- Chirurgie viscérale et digestive ;
- Gynécologie obstétrique ;
- Neurochirurgie ;
- Ophtalmologie ;
- Oto-rhino-laryngologie - chirurgie cervico-faciale ;
- Urologie.

Annexe B

Classification d'Altemeier

Classe d'Altemeier	Type d'intervention
Classe I : chirurgie propre	Incisions primitivement fermées non drainées, non traumatiques, sans inflammation ni faille dans la technique d'asepsie, en l'absence d'ouverture de l'oropharynx, du tube digestif, de l'appareil génito-urinaire ou des voies respiratoires (viscères creux).
Classe II : chirurgie propre-contaminée	● Ouverture de l'appareil génito-urinaire en l'absence d'uroculture positive. ● Ouverture des voies respiratoires, du tube digestif dans de bonnes conditions et sans contamination anormale. ● Ouverture de l'oropharynx ou des voies biliaires en l'absence de bile infectée. ● Ruptures minimes d'asepsie et drainages mécaniques.
Classe III : chirurgie contaminée	● Plaies traumatiques récentes (datant de moins de 4 heures). ● Ouverture du tractus biliaire ou génito-urinaire en présence de bile ou d'urines infectées. ● Contaminations importantes par le contenu du tube digestif. ● Ruptures majeures d'asepsie. ● Interventions en présence d'inflammation aiguë sans pus.
Classe IV : chirurgie sale	● Plaies traumatiques souillées ou traitées de manière retardée (datant de plus de 4 heures). ● Présence de tissus dévitalisés, d'inflammation bactérienne avec pus, de contamination fécale ou de corps étranger. ● Viscères perforés.

Annexe C

Liste des solutions de chirurgie robotique sur le marché et en cours de développement en 2024

Liste des abréviations utilisées :

- Microchir. : Microchirurgie
- Chir. Pedia. : Chirurgie Pédiatrique
- Chir. Plast. : Chirurgie Plastique
- Ophta. : Ophtalmologie
- Cardio. : Cardiologie
- Neuro chir. : Neurochirurgie
- Chir. Rachis : Chirurgie du Rachis
- Ortho. : Orthopédie
- Chir. Thora. : Chirurgie Thoracique
- Gynéco. : Gynécologie
- AV : A venir
- R : Réalisable
- TC : Télé-chirurgie
- CM : Co-manipulateur
- AG : Assistant de guidage

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Acusurgical SAS	Luca	France				AV										TC	Non
Agilis Robotics	Agilis Robotics System	Chine										AV	AV	AV			Oui
AiM Medical Robotics		Etats-Unis						R								AG	Non
Aktormed GMBH	SOLO ASSIST II	Allemagne					R				R		R	R	R	CM	Oui
Alphatec	REMI	Etats-Unis							R							AG	Oui
Amplitude	Andy	France								R						CM	Non
AOT	CARLO	Suisse								R						AG	Non
Artedrone		France					AV										Non
Asahi Intecc Co., Ltd	ANSUR	Japon													R	CM	Oui
Asensus Surgical (TransEnterix) (Karl Storz)	Senhance	Etats-Unis		R							R		R	R	R	TC	Oui
Asensus Surgical (Karl Storz)	LUNA	Etats-Unis											AV	AV	AV	TC	Non
Avateramedical GMBH	AVATERA	Allemagne		AV									AV	R	R	TC	Oui
BJBraun	Aesculap Aeos	Allemagne	R					R	R			R				AG	Oui
Baihui Weikang Technology	REMEBOT Dental	Chine										R				AG	Non
Baihui Weikang Technology	REMEBOT	Chine						R								AG	Non
Beijing Hurwr Medical	HURWA Surgical Robot	Chine								R						AG	Oui
Beijing Rossum Robot Technology Co	ROSSUM	Chine								AV						AG	Non

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	VIscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Beijing Surgerii Technology C	SURGERII SYSTEM	Chine									AV		AV	AV	AV	TC	Non
Beijing WeMed	ETCath	Chine					R									AG	Non
Beyeonics Surgical	Maverick	Israël						R	R							AG	Non
Beyeonics Vision	Beyeonics One	Israël				R										AG	Non
BHS Technologies	Robotic Scope	Autriche						R	R			R				AG	Oui
Biobot Surgical Pte Ltd	iSR'obot Monalisa	Singapour												R		AG	Oui
Bionaut Labs		Etats-Unis						AV								TC	Non
Borns Medical Robotics	Simphoni	Chine														TC	Non
Brain Navi Biotechnology	NaoTrac	Taïwan						R								AG	Oui
Brainlab	CIRQ	Allemagne							R							AG	Oui
Caranx Medical	TAVIPILOT 1	France					AV									AG	Non
Caranx Medical	OBELIA	France											AV				Non
Cardiawave		France															Non
Carvature Medical		Israël							AV							TC	Non
CAScination AG	HEARO	Suisse										R				AG	Oui
Chunli	Yangze River	Chine								R						CM	Non
CMR Surgical	VERSIUS	Royaume-Uni		AV							R		R	R	R	TC	Oui
Collin Medical	RobOtol	France										R				AG	Oui
Columbris MX	Columbris ELS	Chine										AV	AV			TC	Non

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Columbris MX	Columbris SP	Chine											AV	AV	AV	TC	Non
Convergence Medical Robotics	V01	Etats-Unis								AV						TC	Non
CORIN Group	ApolloKnee	Royaume-Uni								R						AG	Oui
CORIN Group	OMNIBotics	Royaume-Uni								R						AG	Oui
Cornerstone robotics	SENTIRE	Chine											AV	AV	AV	TC	Non
Curexo	CUVIS-Joint	Corée du Sud								R						AG	Oui
Curexo	CUVIS-Spine	Corée du Sud							R							AG	Oui
Cyber Surgery	Alaya	Espagne							R							AG	Non
Dentium		Chine										AV				AG	Non
Dentroid		Australie										AV					Non
DEX Surgical	DEX	France									R		R	R	R	CM	Oui
DigiSurge		France					AV										Non
Distalmotion	DEXTER	Suisse											R	R	R	TC	Oui
DLR	MiroSurge	Allemagne											AV			TC	Non
eCential Robotics	CoBot	France							R	R						AG	Oui
EDAP TMS SA	Focal One	France												R		AG	Oui
Edge Medical	MP1000	Chine									R		R	R	R	TC	Oui
Edge Medical	SP1000	Chine													AV	TC	Non
Elmed Medical	AVICENNA ROBOFLEX	Turquie												R		TC	Oui
EndoControl	JAIMY Advance	France		R									R	R	R	CM	Oui

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	VIscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
EndoControl	VIKY	France		R							R		R	R	R	CM	Oui
EndoQuest Robotics (Columbris MX)	Columbris ELS	Etats-Unis										AV	AV			TC	Non
EndoQuest Robotics (Columbris MX)	Columbris SP	Etats-Unis											AV	AV	AV	TC	Non
Extremis Robotics		France								AV							Non
FlexDex, Inc	FlexDex System	Royaume-Uni		R									R	R	R	CM	Oui
Flux Robotics	Flux One	Pays-Bas					AV									AG	Non
ForSight Robotics	ORYOM	Israël				AV											Non
Fortimedix Surgical B.V.	SymphonX Surgical Platform	Pays-Bas											R	R	R	CM	Oui
FreeHand LDT	FREEHAND V1.2	Royaume-Uni					R		R		R		R	R	R	CM	Oui
Ganymed Robotics		France								AV							Non
GBI	AUTOLAP	Etats-Unis											R	R	R	CM	Oui
Globus Medical Inc.	EXCELSIUS GPS	Etats-Unis						R	R							AG	Oui
Hanglok-Tech	SGNIO	Chine									AV					AG	Non
Hangzhou Jianjia Medical Technology Co., Ltd	Arthrobot	Chine						R								AG	Oui
Hangzhou Lancet Robotics	RobPath	Chine								R						AG	Non

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Harbin Sagebot Intelligent Medical Equipment Co., Ltd/ Kangduo	Kangduo	Chine									R		R	R	R	TC	Oui
HeartLander Surgical		Etats-Unis					AV										Non
HIWIN Healthcare	MTG-H100	Taïwan									R		R	R	R	CM	Oui
Horizon Surgical Systems	Polaris	Etats-Unis				AV										AG	Non
Human Xtensions	HandX	Israël									R		R	R	R	CM	Oui
Identabotics		Egypte										AV				AG	Non
INEN Robotics		France											AV				Non
Inetec	NERO	Suisse						AV								TC	Non
Interventional Systems	MICROMATE	Autriche						R	R	R	R		R	R		AG	Oui
Intuitive Surgical	DA VINCI SP	Etats-Unis									AV	R	R	R	R	TC	Oui
Intuitive Surgical	DA VINCI X	Etats-Unis											R	R	R	TC	Oui
Intuitive Surgical	DA VINCI Xi	Etats-Unis		R			R				R	R	R	R	R	TC	Oui
Intuitive Surgical	ION	Etats-Unis									R					TC	Oui
iotaMotion	IotaSOFT	Etats-Unis										R				TC	Non
J&J_Auris Health	MONARCH	Etats-Unis									R			R		TC	Oui
J&J_DePuy Synthes	VELYS	Etats-Unis								R						AG	Oui
J&J_Tinavi	TiRobot	Chine						R	R	R						AG	Oui

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	VIscérale	Urologie	Gynéco.	Type de contrôle	Sur le marché
J&J_Verb Surgical	Ottava	Etats-Unis											AV	AV	AV	TC	Non
Jianjia	ARTHROBOT	Chine								R						CM	Non
Jianjia	THETA	Chine										R				AG	Non
Karl STORZ	ARTip CRUISE & VITOM 2D/3D	Allemagne	R	R	R		R	R	R		R	R		R	R	AG	Oui
Keranova	FemtoMatrix	France				AV										AG	Non
Koh Young	Kymero	Corée du Sud						R								AG	Oui
Ku Leuven	MYNUTIA	Belgique				AV										AG	Non
LARC Robotics		Etats-Unis												AV		TC	Non
LEM Surgical	SAPIEN system	Suisse							AV				AV	AV	AV	TC	Non
Levita Magnetics	MARS	Chili											R	R		CM	Oui
Levita Magnetics	Levita Magnetic Surgical System	Chili											R	R		CM	Oui
LungHealth MedTech	Unicorn	Chine									AV					TC	Non
Lupin Dental		France										AV					Non
Maestro Surgical		Chine							AV								Non
Masmec Biomed	Pyxis	Italie							AV							AG	Non
Medical Devices Corner		Etats-Unis														AG	Non
Medicaroid	HINOTORI	Japon				AV					AV		R	R	R	TC	Oui
Medrobotics	FLEX	Etats-Unis										R	R	R	R	TC	Oui

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Medtronic	Hugo RAS	Etats-Unis		AV			AV				AV		AV	R	R	TC	Oui
Medtronic	MAZOR X Stealth Station	Etats-Unis							R							AG	Oui
Medtronic	Stealth Autoguide	Etats-Unis						R								AG	Oui
Meere Company	REVO-I	Corée du Sud										R	R	R	R	TC	Oui
Meio Medical	Titian	Chine					R										Non
Memic	Hominis	Chine													R	TC	Oui
Mendaera		Etats-Unis															Non
Microbot Medical	Liberty	Etats-Unis					R									AG	Non
MicroPort	HONGHU	Chine								R						AG	Oui
MicroPort	TOUMAI	Chine									R		R	R	R	TC	Oui
MicroPort	SkyWalker	Chine								AV						CM	Non
MicroPort	Spine surgical robot	Chine							AV								Oui
MicroPort	Trans-bronchial surgical robot	Chine									AV					TC	Oui
MicroPort	TAVR Surgical robot	Chine					AV										Oui
Microsure	MUSA	Pays-Bas	R	AV	R											TC	Oui
MMI (Medical Microinstruments)	SYMANI	Italie	R	R	R	AV		AV		AV		R				TC	Oui
Momentis Surgical (Memic)	Anovo (Hominis)	Israël													R	TC	Oui

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	VIscérale	Urologie	Gynéco.	Type de contrôle	Sur le marché
Monogram robot	Monogram Orthopedics	Etats-Unis								AV						AG	Non
Moon Surgical	Maestro	France									R	R	R	R	R	CM	Oui
Moray Medical	Coral	Etats-Unis					AV									AG	Non
Nanoflex Robotics		Suisse					AV										Non
NDR Medical	ANT-C	Singapour									AV					AG	Non
Neocis	Yomi Dental Robot	Etats-Unis										R				AG	Oui
Neuro42		Etats-Unis						AV									Non
Noah Medica	Galaxy System	Etats-Unis									R					TC	Oui
NuVasive	PULSE	Etats-Unis							R							AG	Oui
Ocutrx	OR-Bot	Etats-Unis				R										AG	Non
Olympus	ORBEYE	Japon	R	R	R			R	R		R	R		R		AG	Oui
OphtoRobotics	OR-System	Suisse				AV											Non
OrbSurgical	NeuroArm	Canada						AV								TC	Non
Panda Surgical		Royaume-Uni						AV									Non
Perceptive		Etats-Unis										AV				AG	Non
Perfint Healthcare Pvt Ltd	Maxio III	Inde									R		R	R	R	AG	Oui
Perfint Healthcare Pvt Ltd	Robio EX	Inde									AV		AV	AV	AV	AG	Oui
Perlove Medical	PL300B	Chine							R							AG	Oui
Point Robotics	The Kinguide system	Taïwan							R	AV						AG	Oui

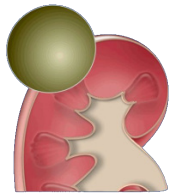
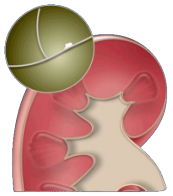
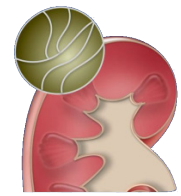
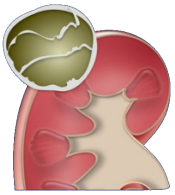
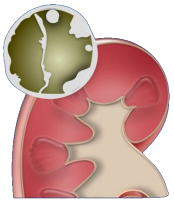
Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Preceyes B.V (ZEISS)	PRECEYES	Pays-Bas				R										TC	Oui
Precision Robotics		Royaume-Uni															Non
Procept BioRobotics	AQUABEAM	Etats-Unis												R		TC	Oui
Procept BioRobotics	Hydros	Etats-Unis												R		TC	Non
Quantum Surgical	EPIONE	France									R		R	R		AG	Oui
Renishaw	Neuromate	Royaume-Uni						R								AG	Oui
Restoration Robotics	ARTAS iX	France			R											AG	Oui
Riverfield	Saroa	Japon														TC	Oui
Rob Surgical	Bitrack	Espagne									AV		AV	AV	AV	TC	Non
Robeauté		France						AV								TC	Non
Robocath	R-ONE	France					R									AG	Oui
Roen Surgical	Zamenix R	Corée du Sud												R		TC	Non
Ronna Medical	RONNA G6	Croatie					R										Non
Ronovo Surgical	CARINA	Chine												AV	AV	TC	Non
Sagebot		Chine							AV			AV		AV			Non
Shenzhen Aibo Medical		Chine					R									AG	Non
Shenzhen Xinjunte	ORTHBOT	Chine							R							AG	Oui
Siemens	Corpath GRX	Allemagne					R	R								AG	Oui
Sinamed	SINAFLEX	Iran		AV			AV			AV	AV	AV	AV	AV	AV	TC	Non
Sinovation Medical	Sinobot X1	Chine						R								AG	Oui

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	VIscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
Smith&Nephew	CORI	Royaume-Uni								R						AG	Non
Smith&Nephew	NAVIO FPS	Royaume-Uni								R						AG	Oui
Soteria Medical	SOTERIA	Pays-Bas												R		TC	Oui
SpineGuard		Etats-Unis							AV							AG	Non
SSInnovations / Surgical Robotic Solutions	SSI Mantra	Inde					R				R	R	R	R	R	TC	Oui
Sterlab	ILY	France												R		TC	Non
Sterotaxis	Genesis	Etats-Unis					R									AG	Oui
Styker	MAKO	Etats-Unis							AV	R						AG	Oui
Suzhou Kongduo Robot	Kangduo	Chine														TC	Non
Spiro Robotics		Etats-Unis															Non
Suzhou Kongduo Robotics	VSROBOT+	Chine							R							AG	Oui
Synaptive	MODUS V	Canada		R				R	R			R				AG	Oui
Tamar Robotics		Israël						AV								TC	Non
TheraSonic		France						AV								AG	Non
THINK Surgical	TMINI	Etats-Unis				R				R						AG	Oui
Think Surgical Inc.	TSolution One	Etats-Unis								R						AG	Oui
TINAVI Medical	TiRobot II	Chine							R							AG	Oui
Titan Medical	ENOS (SPORT)	Canada										AV	AV	AV	AV	TC	Non

Entreprise	Produit	Pays	Micro chir.	Chir. Pedia.	Chir. Plast.	Ophta.	Cardio.	Neuro chir.	Chir. Rachis	Ortho.	Chir. Thora.	ORL / Cervico	Viscéral e	Urologie	Gynéco.	Type de contrôle	Sur le marché
True Digital Surgery	OCCIPTA	Pays-Bas			AV	AV										AG	Non
Tuodao Medical	NP100	Chine									R					AG	Non
UAB Inovatyvi Medicina	Sentante	Lituanie					R									AG	Non
UNIK Orthopedics		Etats-Unis								AV						CM	Non
Vicarious Surgical	BETA 2 SYSTEM	Etats-Unis											AV		AV		Non
Virtual Incision	MIRA	Etats-Unis											AV		AV	TC	Oui
Virtuoso Surgical	VIRTUOSO SYSTEM	Etats-Unis						AV		AV	AV	AV		AV	AV		Non
WEGO	Microhand-S System	Chine											R	AV	AV	TC	Non
XACT Robotics	ACE	Israël									R		R	R		AG	Non
YakeBot Technology	Yakebot	Chine										R				AG	Non
Yuanhua Intelligent Technology / Yuanhua Technology	Yuanhua Surgical Robot	Chine								R						AG	Oui
Zimmer Biomet	ROSA KNEE	Etats-Unis								R						AG	Oui
Zimmer Biomet	ROSA ONE	Etats-Unis						R								AG	Oui

Annexe D

Classification de Bosniak 2019 des tumeurs kystiques du rein

Type I Kyste bénin avec une paroi fine; densité hydrique; pas de cloison, calcification ou composante tissulaire; absence de rehaussement à l'injection du produit de contraste	• Aucune surveillance	
Type II Kyste bénin pouvant contenir une fine cloison; calcifications pariétales fines, absence de rehaussement.	• Aucune surveillance	
Type IIF Cloisons nombreuses et fines; paroi légèrement épaissie; calcifications régulières de la paroi ou des cloisons; minime rehaussement de la paroi et/ou des cloisons; kyste hyperdense intra-rénal et ≥ 3 cm.	• Surveillance TDM tous les ans pendant 5 ans	
Type III Paroi épaisse et irrégulière; cloisons nombreuses et épaisses; calcifications épaisses et irrégulières, rehaussement significatif de la paroi et/ou des cloisons.	• Probabilité de malignité ≈ 50 % • Exérèse du dôme saillant strictement proscrite • Chirurgie d'exérèse selon les règles oncologiques • Surveillance possible dans des cas sélectionnés	
Type IV Végétations intra-kystiques se rehaussant; paroi épaisse et irrégulière se rehaussant.	• Kyste malin • Exérèse du dôme saillant strictement proscrite • Chirurgie d'exérèse selon les règles oncologiques	

Illustrations : https://radiologyassistant.nl/abdomen/kidney/bozniak-2019

Annexe E

Score de néphrométrie RENAL

	1 point	2 points	3 points
Diamètre (D) Mesure du plus grand diamètre transversal (IRM/CT) dans un seul plan.	D ≤ 4 cm	7 > D > 4 cm	D ≥ 7 cm
Exophytique/endophytique Évaluer le pourcentage de masse exophytique par rapport à la masse endophytique.	≥ 50 % exophytique	< 50 % exophytique	Entièrement endophytique
Proximité du sinus rénal ou du système collecteur Mesurer la distance la plus courte (d).	d ≥ 7 mm	7 > d > 4 mm	d ≤ 4 mm
Antérieur vs. postérieur Suffixe appliqué à la localisation primaire de la tumeur par rapport au plan coronal au niveau des vaisseaux hilaires.	A : Antérieur aux vaisseaux hilaires	P : Postérieur aux vaisseaux hilaires	X : Ni l'un ni l'autre
Lignes polaires Déterminer l'emplacement de la tumeur par rapport aux lignes polaires.	Entièrement au-dessus ou au-dessous des lignes polaires	Traverse une ligne polaire	Diamètre supérieur à 50\,\% de la ligne polaire, entièrement contenu dans les lignes polaires ou traversant la ligne médiane
Tumeur hilaire Suffixe	H : En cas de contact avec une artère ou une veine rénale		

Score RENAL	Complexité
4 - 6	Faible
7 - 9	Intermédiaire
10 - 12	Elevé

Annexe F

Score de néphrométrie PADUA

	1 point	2 points	3 points
Localisation longitudinale (polaire)	Supérieure ou inférieure	Moyenne	
Exophytique vs. endophytique	< 50 % exophytique	≥ 50 % exophytique	Entièrement endophytique
Bordure rénal	Latérale	Médiale	
Sinus rénal	Non impliqué	Impliqué	
Système collecteur urinaire	Non impliqué	Impliqué	
Taille de la tumeur	T ≤ 4 cm	4,1 - 7 cm	T > 7 cm

Score PADUA	Risque
6 - 7	Faible
8 - 9	Intermédiaire
≥ 10	Elevé

Annexe G

Classification Tumour Node Metastasis (TNM) 2017 du cancer du rein

T	Tx	Tumeur primitive non évaluable
	T0	Pas de signe de tumeur primitive
	T1	Tumeur limitée au rein, inférieure ou égale à 7 cm dans sa plus grande dimension
		T1a : Tumeur inférieure ou égale à 4 cm
		T1b : Tumeur inférieure ou égale à 7 cm et supérieure à 4 cm
	T2	Tumeur limitée au rein, supérieure à 7 cm dans sa plus grande dimension
		T2a : Tumeur inférieure ou égale à 10 cm et supérieure à 7 cm
		T2b : Tumeur limitée au rein et supérieure à 10 cm
	T3	La tumeur s'étend aux veines principales ou aux tissus périnéphriques, mais pas à la glande surrénale ni au fascia de Gerota
		T3a : Tumeur envahissant la veine rénale ou à ses branches segmentaires ou le système caliciel ou la graisse périrénale et/ou le sinus rénal
		T3b : Tumeur envahissant la veine rénale et/ou de la veine cave inférieure sous-diaphragmatique
		T3c : Tumeur envahissant la veine cave inférieure sus-diaphragmatique ou sa paroi
	T4	Tumeur envahissant les organes de voisinage ou la graisse périrénale à travers le rein (franchissement du fascia de Gerota)
N	Nx	Ganglions régionaux non évaluables
	N0	Pas d'atteinte des ganglions régionaux
	N1	Métastase(s) ganglionnaire(s)
M	Mx	Métastases à distance non évaluables
	M0	Absence de métastase à distance
	M1	Présence de métastase(s) à distance

pTNM			
Stade I	T1	N0	M0
Stade II	T2	N0	M0
Stade III	T3	N0	M0
	T1, T2, T3	N1	M0
Stade IV	T4	N0, N1	M0
	T1, T2, T3, T4	N0, N1	M1

Annexe H

Répartition des néphrectomies réalisées en 2023 selon la CCAM

Données issues de la base de données Open CCAM 2023.

Code CCAM	Intitulé CCAM	CO CMI	NT NP	NB Actes
JAFA031	Néphrectomie **totale élargie** à la loge rénale, par thoraco-phréno-**laparotomie**	CO	NT	0
JAFA030	Néphrectomie **partielle** sans dissection du pédicule vasculaire, par **laparotomie**	CO	NP	0
JAFC002	**Héminéphrectomie** avec urétérectomie partielle, par **coelioscopie** ou par rétropéritonéoscopie	CMI	NP	0
JAFA029	Néphrectomie **totale élargie** à la loge rénale avec surrénalectomie, par **laparotomie** ou par abord lomboabdominal	CO	NT	74
JAFA028	Néphrectomie **totale élargie** à la loge rénale avec surrénalectomie et thrombectomie par cavotomie, par **laparotomie** ou par abord lomboabdominal	CO	NT	0
JAFA024	Néphrectomie **partielle** sans dissection du pédicule vasculaire, par **lombotomie** ou par abord lombal vertical	CO	NP	0
JAFA025	Néphrectomie **totale élargie** à la loge rénale avec surrénalectomie et thrombectomie par **cavotomie**, par thoraco-phréno-laparotomie	CO	NT	0
JAFA026	**Héminéphrectomie** pour fusion rénale, par abord **direct**	CO	NP	0
JAFA027	**Binéphrectomie**, par abords **lombaux** verticaux	CO	NT	0
JAFC004	**Binéphrectomie**, par **coelioscopie** ou par rétropéritonéoscopie	CMI	NT	0
JAFC005	Néphrectomie **partielle**, par **coelioscopie** ou par rétropéritonéoscopie	CMI	NP	5077
JAFC006	Néphrectomie **totale unilatérale**, par **coelioscopie** ou par rétropéritonéoscopie	CMI	NT	490
JAFC007	**Héminéphrectomie** avec urétérectomie totale, par **coelioscopie** ou par rétropéritonéoscopie	CMI	NP	0
JCFA004	Exérèse d'une urétérocèle avec réimplantation urétérovésicale, **héminéphrectomie** et urétérectomie totale homolatérales, par abord direct	CO	NP	0
JAFA023	Néphrectomie **totale** unilatérale, par **laparotomie**	CO	NT	91

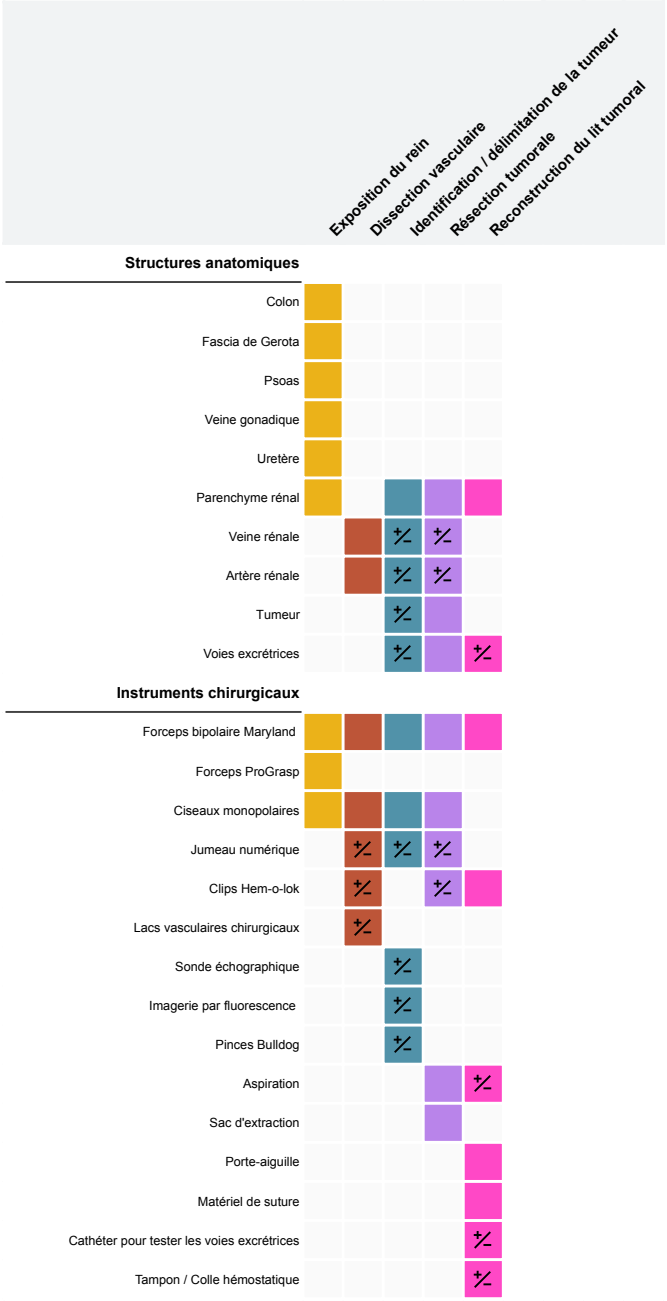
Code CCAM	Intitulé CCAM	CO CMI	NT NP	NB Actes
JAFA022	Néphrectomie **totale élargie** à la loge rénale avec thrombectomie par cavotomie, par thoraco-phréno-**laparotomie**	CO	NT	0
JAFA003	**Héminéphrectomie** avec urétérectomie totale, par abord **direct**	CO	NP	0
JAFA002	Néphrectomie **totale**, par **lombotomie**	CO	NT	0
JAFA005	Néphrectomie **totale** élargie à la loge rénale avec résection latérale de la veine cave inférieure, par abord **direct**	CO	NT	0
JAFA006	Néphrectomie **partielle** ex situ avec autotransplantation du rein, par abord **direct**	CO	NP	0
JAFA001	**Héminéphrectomie** avec urétérectomie partielle, par abord **direct**	CO	NP	0
JAFA008	Néphrectomie **partielle** avec dissection du pédicule vasculaire, par **lombotomie** ou par abord lombal postérieur	CO	NP	161
JAFA016	**Binéphrectomie**, par **laparotomie**	CO	NT	0
JAFA019	Néphrectomie **partielle** avec dissection du pédicule vasculaire, par **laparotomie**	CO	NP	62
JAFA021	Néphrectomie **totale élargie** à la loge rénale avec surrénalectomie, par thoraco-phréno-**laparotomie**	CO	NT	0
JAFA014	Néphrectomie **totale élargie** à la loge rénale avec thrombectomie par cavotomie, par **laparotomie** ou par abord lomboabdominal	CO	NT	0
JAFA009	Néphrectomie **totale élargie** à la loge rénale, par **laparotomie** ou par abord lomboabdominal	CO	NT	237
JAFA010	Néphrectomie **totale** unilatérale, par **abord lombal vertical**	CO	NT	0
JAFA011	Néphrectomie **totale élargie** à la loge rénale avec thrombectomie par cavotomie et contrôle intraatrial cardiaque, par thoraco-phréno-**laparotomie**	CO	NT	0
JAFA012	Néphrectomie **totale** par **laparotomie**, avec extraction endoscopique rétrograde de l'uretère [stripping de l'uretère]	CO	NT	0

CCAM : Classification Commune des Actes Médicaux
CO : Chirurgie Ouverte
CMI : Chirurgie Mini Invasive
NT : Néphrectomie Totale
NP : Néphrectomie Partielle
Note : les effectifs inférieurs à 11 ont été masqués afin de préserver l'anonymat des patients.

Annexe I

Structures anatomiques exposées et instrumentation employée lors des phases d'une NPRA

(Figure page suivante)


Exposition du rein
Dissection vasculaire
Identification / délimitation de la tumeur
Résection tumorale
Reconstruction du lit tumoral
Structures anatomiques
Colon
Fascia de Gerota
Psoas
Veine gonadique
Uretère
Parenchyme rénal
Veine rénale
Artère rénale
Tumeur
Voies excrétrices
Instruments chirurgicaux
Forceps bipolaire Maryland
Forceps ProGrasp
Ciseaux monopolaires
Jumeau numérique
Clips Hem-o-lok
Lacs vasculaires chirurgicaux
Sonde échographique
Imagerie par fluorescence
Pinces Bulldog
Aspiration
Sac d'extraction
Porte-aiguille
Matériel de suture
Cathéter pour tester les voies excrétrices
Tampon / Colle hémostatique

Annexe J

Prototypes de réalité augmentée en NPRA

Liste des fonctionnalités principales des différentes versions du prototype de réalité augmentée pour la néphrectomie partielle.

Date	Version	Fonctionnalités
Octobre 2022	V0.1	Import des données préopératoires (jumeau numérique) Acquisition du flux vidéo stéréoscopique de l'endoscope Étalonnage géométrique de l'endoscope stéréoscopique Reconstruction peropératoire de la surface rénale par stéréovision Recalage initial manuel rigide Affichage et enregistrement stéréoscopique de la réalité augmentée
Avril 2023	V0.2	**Nouveau** Suivi monoculaire de l'organe avec sélecteur automatique d'images clés
Janvier 2024	V0.3	**Nouveau** Nouveau back-end d'acquisition du flux vidéo stéréoscopique Nouveau back-end pour l'étalonnage géométrique monoculaire Nouveau back-end de **recalage manuel déformable** Nouveau back-end de suivi monoculaire avec ajout auto d'images clés Intégration du **modèle de segmentation** de la **surface rénale** Nouveau back-end d'affichage monoculaire de la réalité augmentée Nouveau back-end d'enregistrement vidéo Nouvelle interface utilisateur **Supprimé** Reconstruction peropératoire de la surface rénale par stéréovision Étalonnage géométrique de l'endoscope stéréoscopique Affichage stéréoscopique de la réalité augmentée
Février 2024	V0.3.1	**Nouveau** Page de configuration des sources vidéo Gestion du flux vidéo peropératoire Visualisation des images clés recalées Affichage de la réalité augmentée dans la console du chirurgien
Juillet 2024	V0.3.2	**Nouveau** Ajout de la page patient pour la création de chirurgies Ajout d'un outil de mesure en millimètres Ajout d'une méthode de recalage semi-automatique Affichage des instruments par dessus la réalité augmentée Sauvegarde et chargement de sessions

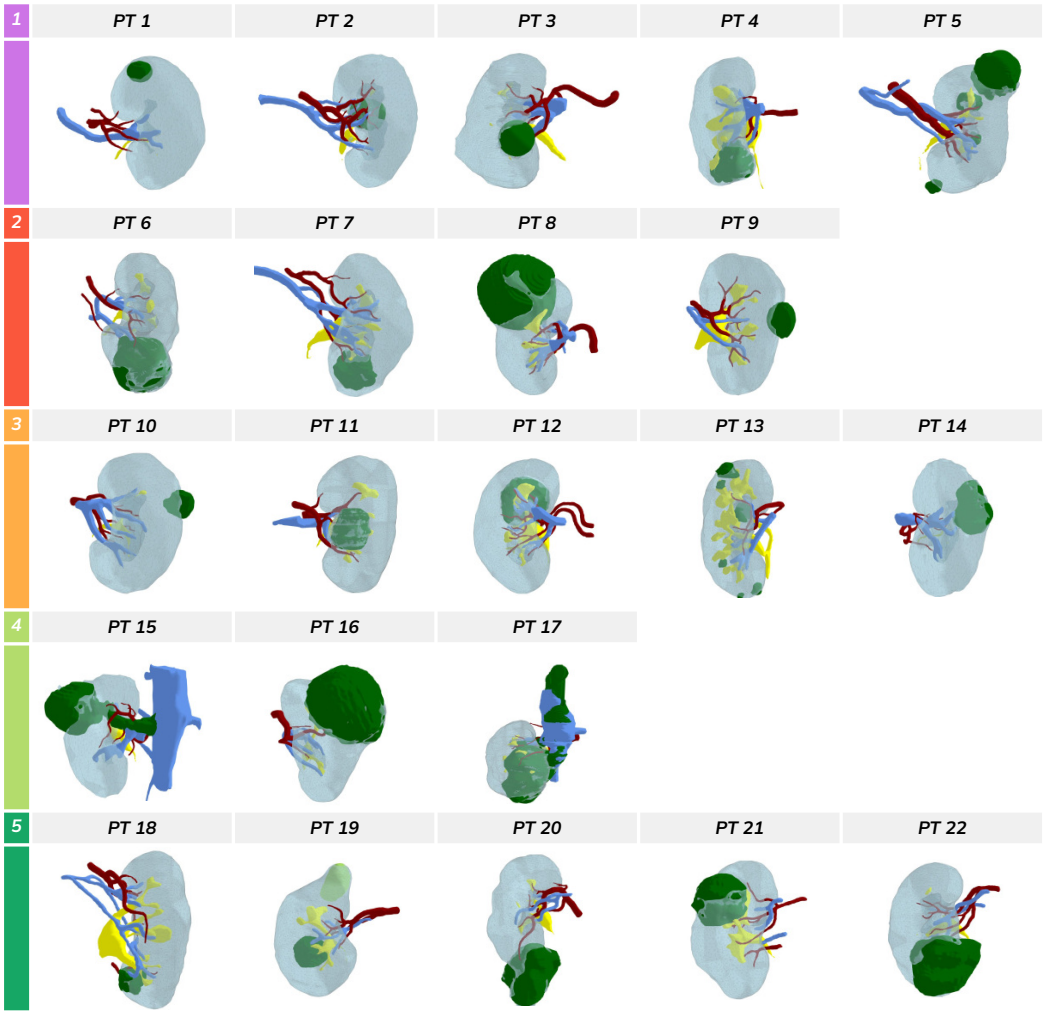
Date	Version	Fonctionnalités
Septembre 2024	V0.3.3	Nouveau Acquisition et étalonnage du flux vidéo stéréoscopique Région personnalisable de suivi Affichage en réalité virtuelle avec ou sans suivi Affichage stéréoscopique dans la console du chirurgien Enregistrement stéréoscopique
Octobre 2024	V0.3.4	Nouveau Ajout de la gestion d'un flux auxiliaire Réalité virtuelle dans la console du chirurgien Sauvegarde intégrale des données issues du suivi automatique
Février 2025	V0.3.5	Nouveau Ajout de statistiques sur la chirurgie et la réalité augmentée Ajout de la gestion d'un second flux auxiliaire Recalage manuel à partir de la console du chirurgien Contrôles basiques du logiciel dans la console du chirurgien Modifié Réorganisation du workflow pour l'obtention de la réalité augmentée Correction de l'étalonnage stéréoscopique (convergence) Amélioration de la gestion des flux vidéo Supprimé Recalage manuel déformable

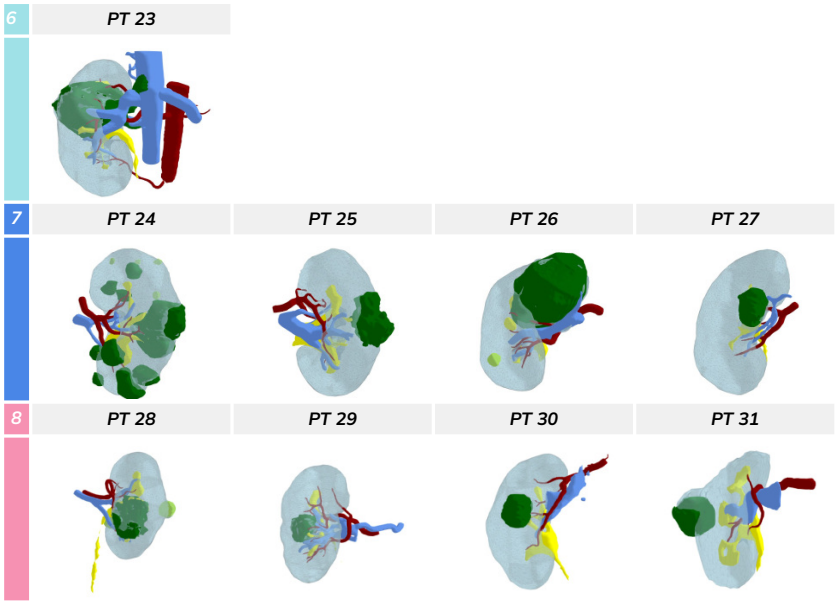
Annexe K

Cas cliniques de NPRA assistée par réalité augmentée

Cas cliniques réalisés lors des huit sessions d'essais de faisabilité au CHU de Bordeaux entre octobre 2022 et février 2025.

Chaque ligne correspond à une session d'essai, tandis que chaque cellule représente une intervention chirurgicale soit un patient.





PT : Patient Test

Annexe L

Liste non exhaustive de 123 méthodes de la littérature pour la détection et la description de points clés

Les méthodes en bleu sont celles utilisées pour la construction de Scores de Comparaison Gauche-Droite (SCGD).

Nom de la méthode	Type de méthode	Classe de méthode	Code disponible	Date
Affine-SIFT	DET-DESC	VPO	OUI	2009
AffNet	DET-DESC	APP	OUI	2018
AGAST	DET	VPO	OUI	2010
AKAZE	DET-DESC	VPO	OUI	2011
ALIKE	DET-DESC	APP	OUI	2021
ALIKED	DET-DESC	APP	OUI	2023
ASLFeat	DET-DESC	APP	OUI	2020
ASpanFormer	E-T-E	APP	OUI	2022
ASTR	E-T-E	APP	OUI	2023
BDCN	DET	APP	OUI	2019
BEBLID	DESC	APP	OUI	2020
BinBoost	DESC	VPO	OUI	2013
BoostDesc	DESC	VPO	OUI	2013
BRIEF	DESC	VPO	OUI	2010
BRISK	DET-DESC	VPO	OUI	2011
Canny	DET	VPO	OUI	1986
CED	DET	APP	OUI	2017
CEDN	DET	APP	OUI	2016
CenSurE	DET	VPO	OUI	2008
ClusterGNN	E-T-E	APP	NON	2022
COAM	E-T-E	APP	OUI	2021
COB	DET	APP	OUI	2016
ContextDesc	DESC	APP	OUI	2019
D-BRIEF	DESC	APP	NON	2012
D2-Net	DET-DESC	APP	OUI	2019

Nom de la méthode	Type de méthode	Classe de méthode	Code disponible	Date
DAISY	DESC	VPO	OUI	2011
DarkFeat	DET-DESC	APP	OUI	2023
DeDoDe	DET-DESC	APP	OUI	2023
DeepBit	DESC	APP	OUI	2016
DeepContour	DET	APP	OUI	2015
DeepCorner	DET-DESC	APP	NON	2023
DeepEdge	DET	APP	OUI	2015
DeepMatcher	E-T-E	APP	OUI	2023
DexiNed	DET	APP	OUI	2020
DiffusionEdge	DET	APP	OUI	2024
DISK	DET-DESC	APP	OUI	2020
DKM	E-T-E	APP	OUI	2023
DOAP	DESC	APP	NON	2018
DoG	DET	VPO	OUI	1999
DSCB	DET	APP	NON	2020
DSP-SIFT	DET-DESC	VPO	OUI	2015
DualRC-Net	E-T-E	APP	OUI	2020
EDTER	DET	APP	OUI	2022
Efficient LoFTR	E-T-E	APP	OUI	2024
FAST	DET	VPO	OUI	2006
FREAK	DESC	VPO	OUI	2012
GeoDesc	DESC	APP	OUI	2018
GFTT	DET	VPO	OUI	1994
GLOH	DESC	VPO	NON	?
GlueStick	E-T-E	APP	OUI	2023
GoodPoint	DET-DESC	APP	OUI	2020
HardNet	DESC	APP	OUI	2017
Harris Corner Measure	DET	VPO	OUI	1988
HED	DET	APP	OUI	2015
HessAff	DET	VPO	OUI	2004
Harris Laplace	DET	VPO	OUI	2004
HOG	DET-DESC	VPO	OUI	2005
HyNet	DESC	APP	OUI	2020
KAZE	DET-DESC	VPO	OUI	2012
KeyNet	DET	APP	OUI	2019
L2-Net	DESC	APP	OUI	2017
LATCH	DESC	VPO	OUI	2016
LDAHash	DESC	VPO	NON	2010
LDB	DESC	APP	OUI	2013
LF-Net	DET-DESC	APP	OUI	2018

Nom de la méthode	Type de méthode	Classe de méthode	Code disponible	Date
LiFF	DET-DESC	APP	OUI	2019
LIFT	DET-DESC	APP	OUI	2016
LightGlue	E-T-E	APP	OUI	2023
LoFTR	E-T-E	APP	OUI	2021
LogPolarDesc	DESC	APP	OUI	2019
LPCB	DET	APP	OUI	2018
LUCID	DESC	VPO	OUI	2012
MagicPoint	DET	APP	NON	2017
MambaGlue	E-T-E	APP	OUI	2025
MatchFormer	E-T-E	APP	OUI	2022
MROGH	DESC	VPO	OUI	2011
MSD	DET	VPO	OUI	2014
MSER	DET	VPO	OUI	2008
MuGE	DET	APP	OUI	2024
OEF	DET	APP	OUI	2015
ORB	DET-DESC	VPO	OUI	2011
Patch2Pix	E-T-E	APP	OUI	2021
PATS	E-T-E	APP	OUI	2023
PCA-SIFT	DET-DESC	VPO	OUI	2004
PDC-Net	E-T-E	APP	OUI	2021
PDC-Net+	E-T-E	APP	OUI	2023
PiDiNet	DET	APP	OUI	2021
PosFeat	DET-DESC	APP	OUI	2022
R2D2	DET-DESC	APP	OUI	2019
R2FD2	E-T-E	APP	NON	2022
RankED	DET	APP	OUI	2024
RCF	DET	APP	OUI	2017
RCN	DET	APP	OUI	2019
RIFT	E-T-E	APP	OUI	2019
RIFT2	E-T-E	APP	OUI	2023
RoMa	E-T-E	APP	OUI	2024
RootSift	DET-DESC	VPO	OUI	2012
S-AKAZE	DET-DESC	VPO	NON	2016
SCFEAT	DET	APP	OUI	2022
SDDH	DESC	APP	OUI	2023
SE	DET	VPO	OUI	2014
SEKD	DET-DESC	APP	OUI	2020
SEM	E-T-E	APP	OUI	2023
SGMNet	E-T-E	APP	OUI	2021
Shi-Tomasi	DET	VPO	OUI	1994
SIFT	DET-DESC	VPO	OUI	2004

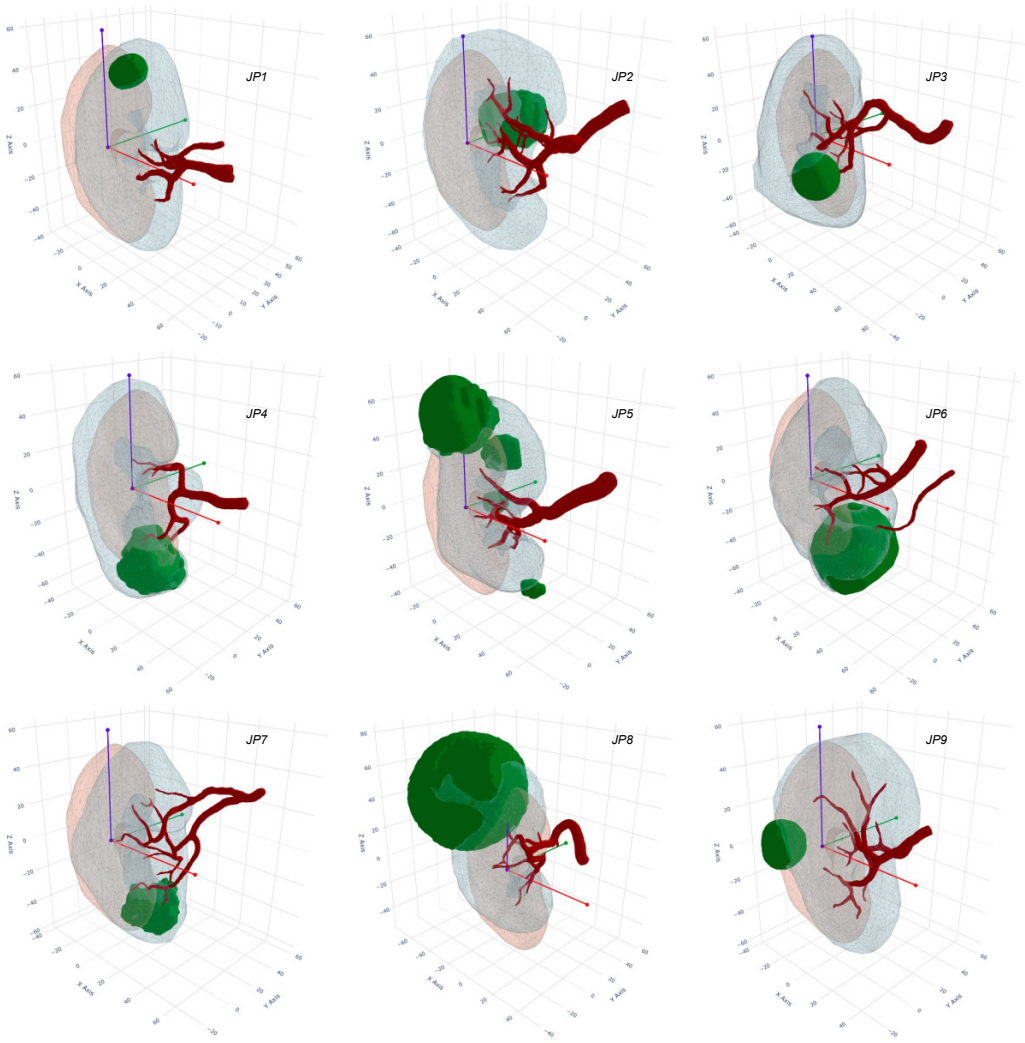
Nom de la méthode	Type de méthode	Classe de méthode	Code disponible	Date
SiLK	DET	APP	OUI	2023
SOSNet	DESC	APP	OUI	2019
SPHORB	DET-DESC	VPO	OUI	2015
STAR	DET	VPO	OUI	2008
SuperGlue	E-T-E	APP	OUI	2020
SuperPoint	DET-DESC	APP	OUI	2017
SURF	DET-DESC	VPO	OUI	2008
TCDET	DET	APP	OUI	2017
TFEAT	DESC	APP	OUI	2016
TILDE	DET	APP	OUI	2015
TopicFM	E-T-E	APP	OUI	2023
TopicFM+	E-T-E	APP	OUI	2024
UAED	DET	APP	OUI	2023
UCN	E-T-E	APP	OUI	2016
VGG	DESC	VPO	OUI	2014
XFEAT	E-T-E	APP	OUI	2024
ZippyPoint	DET-DESC	APP	OUI	2023

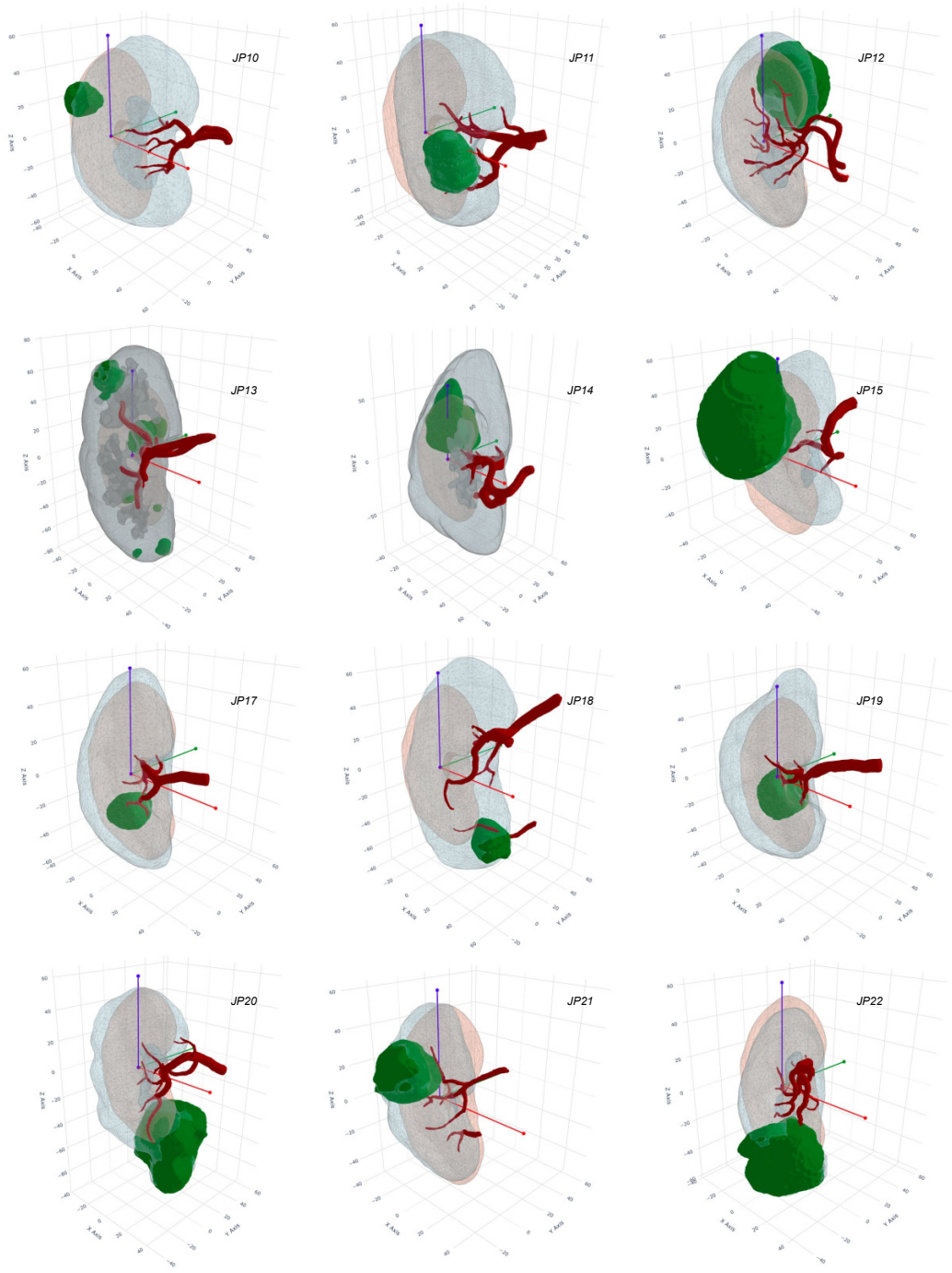
DET : détecteur pur
DESC : descripteur pur
DET-DESC : détecteur-descripteur
E-T-E : méthode end-to-end (détection, description et appariement)
VPO : méthode basée sur la vision par ordinateur
APP : méthode basée sur l'apprentissage (apprentissage machine ou apprentissage profond)

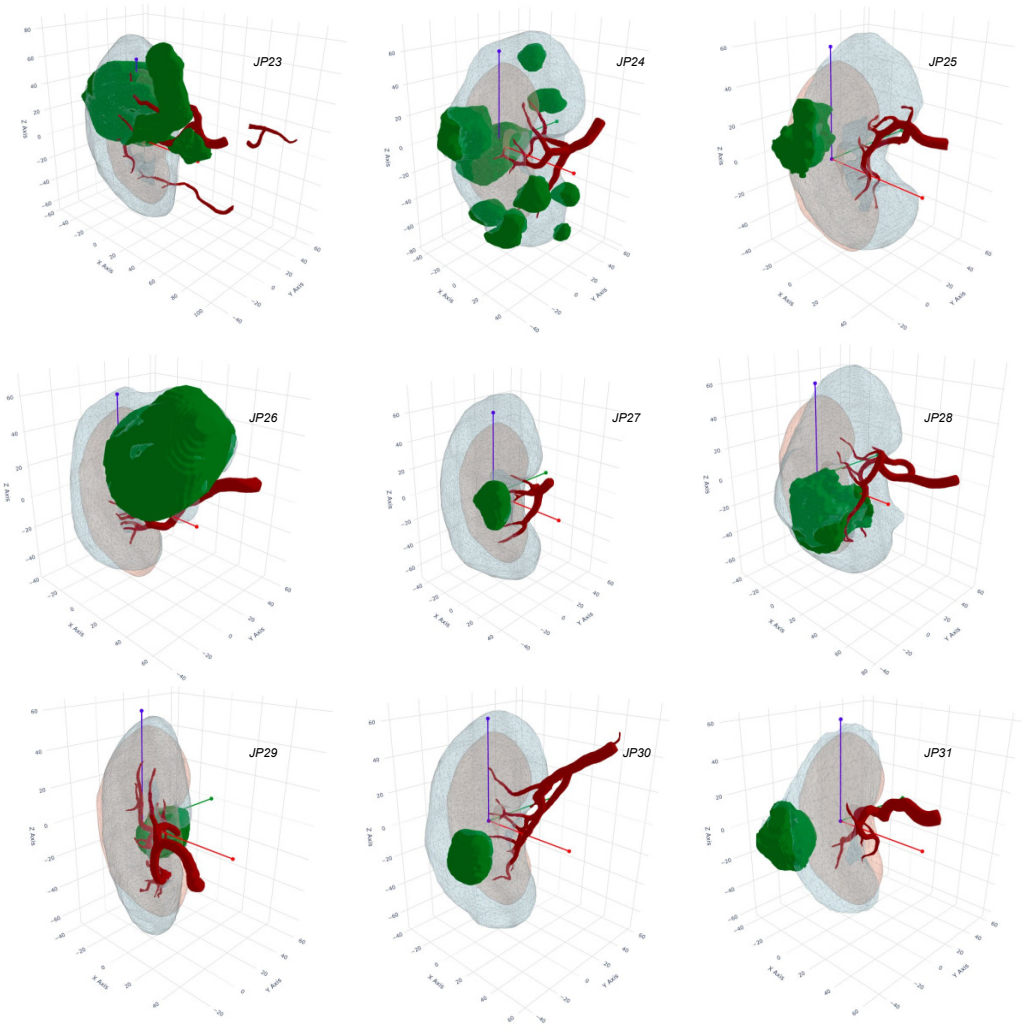
Annexe M

Visualisation des jumeaux numériques recalés sur un prototype de forme

Visualisation des trente Jumeaux numériques Patient (JP) recalés de manière rigide sur le prototype de forme visible en orange.




JP10
JP11
JP12
JP13
JP14
JP15
JP17
JP18
JP19
JP20
JP21
JP22


JP23
JP24
JP25
JP26
JP27
JP28
JP29
JP30
JP31

Bibliographie

[Abaza 2016] Ronney Abaza, Daniel D Eun, Michele Gallucci, Inderbir S Gill, Mani Menon, Alexandre Mottrie et Ahmad Shabsigh. *Robotic surgery for renal cell carcinoma with vena caval tumor thrombus*. European urology focus, vol. 2, no. 6, pages 601–607, 2016. (Cité en page 79.)

[Abel 2011] E Jason Abel, Stephen H Culp, Nizar M Tannir, Surena F Matin, Pheroze Tamboli, Eric Jonasch et Christopher G Wood. *Primary tumor response to targeted agents in patients with metastatic renal cell carcinoma*. European urology, vol. 59, no. 1, pages 10–15, 2011. (Cité en page 77.)

[Abou Youssif 2007] Tamer Abou Youssif, Wassim Kassouf, Jordan Steinberg, Armen G Aprikian, Micheal P Laplante et Simon Tanguay. *Active surveillance for selected patients with renal masses : updated results with long-term follow-up*. Cancer, vol. 110, no. 5, pages 1010–1014, 2007. (Cité en page 75.)

[Abu-Ghanem 2020] Yasmin Abu-Ghanem, Sergio Fernandez-Pello, Axel Bex, Börje Ljungberg, Laurence Albiges, Saeed Dabestani, Rachel H Giles, Fabian Hofmann, Milan Hora, Markus A Kuczyk*et al.* *Limitations of available studies prevent reliable comparison between tumour ablation and partial nephrectomy for patients with localised renal masses : a systematic review from the European Association of Urology Renal Cell Cancer Guideline Panel*. European urology oncology, vol. 3, no. 4, pages 433–452, 2020. (Cité en page 74.)

[Académie nationale de médecine 2024] Académie nationale de médecine. *Dictionnaire de l'Académie nationale de médecine*, 2024. (Cité en pages 4 et 8.)

[Acidi 2023] Belkacem Acidi, Mohammed Ghallab, Stéphane Cotin, Eric Vibert et Nicolas Golse. *Augmented reality in liver surgery*. Journal of Visceral Surgery, vol. 160, no. 2, pages 118–126, 2023. (Cité en page 171.)

[Adagolodjo 2017] Yinoussa Adagolodjo, Raffaella Trivisonne, Nazim Haouchine, Stéphane Cotin et Hadrien Courtecuisse. *Silhouette-based pose estimation for deformable organs application to surgical augmented reality*. In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 539–544. IEEE, 2017. (Cité en page 172.)

[Adam 2014] Kingma DP Ba J Adam*et al.* *A method for stochastic optimization*. arXiv preprint arXiv :1412.6980, vol. 1412, no. 6, 2014. (Cité en page 262.)

[Agence de la biomédecine 2022] Agence de la biomédecine. *Registre français des traitements de suppléance de l'insuffisance rénale chronique*. 2022. (Cité en page 67.)

[Agence Régional de Santé Auvergne-Rhône-Alpes 2018] Agence Régional de Santé Auvergne-Rhône-Alpes. *Blocs opératoires et locaux associés. Guide de bonnes pratiques de conception*, Octobre 2018. (Cité en page 13.)

[Agochukwu 2014] Nnenaya Agochukwu et Brian Shuch. *Clinical management of renal cell carcinoma with venous tumor thrombus*. World Journal of Urology, vol. 32, pages 581–589, 2014. (Cité en page 78.)

[Agrawal 2008] Motilal Agrawal, Kurt Konolige et Morten Rufus Blas. *Censure : Center surround extremas for realtime feature detection and matching*. In European conference on computer vision, pages 102–115. Springer, 2008. (Cité en pages 145 et 221.)

[Ahmed 2014] Muneeb Ahmed, Luigi Solbiati, Christopher L Brace, David J Breen, Matthew R Callstrom, J William Charboneau, Min-Hua Chen, Byung Ihn Choi, Thierry De Baère, Gerald D Dodd III*et al.* *Image-guided tumor ablation : standardization of terminology and reporting criteria—a 10-year update*. Radiology, vol. 273, no. 1, pages 241–260, 2014. (Cité en page 74.)

[Åkerlund 2021] John Åkerlund, Erik Holmberg, Per Lindblad, Maria Stendahl, Börje Ljungberg, Andreas Thorstenson et Sven Lundstam. *Increased risk for renal cell carcinoma in end stage renal disease–a population-based case-control study*. Scandinavian journal of urology, vol. 55, no. 3, pages 209–214, 2021. (Cité en pages 69 et 70.)

[Akhmetyanov 2021] Azat Akhmetyanov, Anastasiia Kornilova, Marsel Faizullin, David Pozo et Gonzalo Ferrer. *Sub-millisecond Video Synchronization of Multiple Android Smartphones*. In 2021 IEEE Sensors, pages 1–4. IEEE, 2021. (Cité en page 127.)

[Al-Bayati 2019] Osamah Al-Bayati, Aws Hasan, Deepak Pruthi, Dharam Kaushik et Michael A Liss. *Systematic review of modifiable risk factors for kidney cancer*. In Urologic Oncology : Seminars and Original Investigations, volume 37, pages 359–371. Elsevier, 2019. (Cité en page 69.)

[Al-Nuaimi 2012] Anas Al-Nuaimi, Burak Cizmeci, Florian Schweiger, Roman Katz, Sinan Taifour, Eckehard Steinbach et Michael Fahrmair. *ConCor+ : Robust and confident video synchronization using consensus-based Cross-correlation*. In 2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP), pages 83–88. IEEE, 2012. (Cité en page 129.)

[Alahi 2012] Alexandre Alahi, Raphael Ortiz et Pierre Vandergheynst. *Freak : Fast retina keypoint*. In 2012 IEEE conference on computer vision and pattern recognition, pages 510–517. Ieee, 2012. (Cité en page 145.)

[Alam 2019] Ridwan Alam, Hiten D Patel, Tijani Osumah, Arnav Srivastava, Michael A Gorin, Michael H Johnson, Bruce J Trock, Peter Chang, Andrew A Wagner, James M McKiernan*et al.* *Comparative effectiveness of management options for patients with small renal masses : a prospective cohort study*. BJU international, vol. 123, no. 1, pages 42–50, 2019. (Cité en page 75.)

[Alander 2012] Jarmo T Alander, Ilkka Kaartinen, Aki Laakso, Tommi Pätilä, Thomas Spillmann, Valery V Tuchin, Maarit Venermo et Petri Välisuo. *A review of indocyanine green fluorescent imaging in surgery*. International journal of biomedical imaging, vol. 2012, no. 1, page 940585, 2012. (Cité en page 49.)

[Albiges 2022] Laurence Albiges, Nizar M Tannir, Mauricio Burotto, David McDermott, Elizabeth R Plimack, Philippe Barthélémy, Camillo Porta, Thomas Powles, Frede Donskov, Saby George*et al. First-line nivolumab plus ipilimumab versus sunitinib in patients without nephrectomy and with an evaluable primary renal tumor in the CheckMate 214 trial*. European urology, vol. 81, no. 3, pages 266–271, 2022. (Cité en page 77.)

[Albl 2017] Cenek Albl, Zuzana Kukelova, Andrew Fitzgibbon, Jan Heller, Matej Smid et Tomas Pajdla. *On the two-view geometry of unsynchronized cameras*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4847–4856, 2017. (Cité en page 129.)

[Alcala 2022] Karine Alcala, Daniela Mariosa, Karl Smith-Byrne, Dariush Nasrollahzadeh Nesheli, Robert Carreras-Torres, Eva Ardanaz Aicua, Nicola P Bondonno, Catalina Bonet, Mattias Brunström, Bas Bueno-de Mesquita*et al. The relationship between blood pressure and risk of renal cell carcinoma*. International journal of epidemiology, vol. 51, no. 4, pages 1317–1327, 2022. (Cité en page 69.)

[Alcantarilla 2011] Pablo F Alcantarilla et T Solutions. *Fast explicit diffusion for accelerated features in nonlinear scale spaces*. IEEE Trans. Patt. Anal. Mach. Intell, vol. 34, no. 7, pages 1281–1298, 2011. (Cité en page 145.)

[Alcantarilla 2012] Pablo Fernández Alcantarilla, Adrien Bartoli et Andrew J Davison. *KAZE features*. In Computer Vision–ECCV 2012 : 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI 12, pages 214–227. Springer, 2012. (Cité en pages 146 et 217.)

[Alchin 2016] David R Alchin, Declan Murphy et Nathan Lawrentschuk. *Predicting the risk of positive surgical margins following robotic-assisted radical prostatectomy*. Minerva Urologica e Nefrologica= The Italian Journal of Urology and Nephrology, vol. 69, no. 1, pages 56–62, 2016. (Cité en page 170.)

[Ali 2022] Muhammad Ali, Jennifer Mooi, Nathan Lawrentschuk, Rana R McKay, Raquibul Hannan, Simon S Lo, William A Hall et Shankar Siva. *The role of stereotactic ablative body radiotherapy in renal cell carcinoma*. European urology, vol. 82, no. 6, pages 613–622, 2022. (Cité en page 74.)

[Allan 2021] Max Allan, Jonathan Mcleod, Congcong Wang, Jean Claude Rosenthal, Zhenglei Hu, Niklas Gard, Peter Eisert, Ke Xue Fu, Trevor Zeffiro, Wenyao Xia*et al. Stereo correspondence and reconstruction of endoscopic data challenge*. arXiv preprint arXiv :2101.01133, 2021. (Cité en page 226.)

[Altamar 2011] Hernan O Altamar, Rowena E Ong, Courtenay L Glisson, Davis P Viprakasit, Michael I Miga, Stanley Duke Herrell et Robert L Galloway. *Kidney deformation and intraprocedural registration : a study of elements of image-guided kidney surgery*. Journal of endourology, vol. 25, no. 3, pages 511–517, 2011. (Cité en pages 96 et 177.)

[Altemeier 1985] William A. Altemeier. *Manual on Control of Infection in Surgical Patients.* Annals of Internal Medicine, vol. 102, pages 881–881, 1985. (Cité en page 12.)

[Amanatides 1992] John Amanatides. *Algorithms for the detection and elimination of specular aliasing.* In Proc. Graphics Interface, volume 92, pages 86–93. Citeseer, 1992. (Cité en page 131.)

[Amano 2018] Kohtaro Amano. *Camera head for endoscope*, 2018. (Cité en page 21.)

[Amaral 2021] BS Amaral, P Macek, A Arora, CL Pazeto, AS Zugail, A Mombet, A Fregeville, M Lefevre, R Sanchez-Salas et X Cathelineau. *Renal tumor biopsy : rationale to avoid surgery in small renal masses.* Current Urology Reports, vol. 22, pages 1–7, 2021. (Cité en page 72.)

[Amberg 2007] Brian Amberg, Sami Romdhani et Thomas Vetter. *Optimal step nonrigid ICP algorithms for surface registration.* In 2007 IEEE conference on computer vision and pattern recognition, pages 1–8. IEEE, 2007. (Cité en page 175.)

[Amin 2017] Mahul B Amin, Frederick L Greene, Stephen B Edge, Carolyn C Compton, Jeffrey E Gershenwald, Robert K Brookland, Laura Meyer, Donna M Gress, David R Byrd et David P Winchester. *The eighth edition AJCC cancer staging manual : continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging.* CA : a cancer journal for clinicians, vol. 67, no. 2, pages 93–99, 2017. (Cité en page 73.)

[Amir-Khalili 2013] Alborz Amir-Khalili, Masoud S Nosrati, Jean-Marc Peyrat, Ghassan Hamarneh et Rafeef Abugharbieh. *Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy : A phantom study.* In International Workshop on Medical Imaging and Virtual Reality, pages 182–191. Springer, 2013. (Cité en page 97.)

[Amparore 2022a] Daniele Amparore, Enrico Checcucci, Pietro Piazzolla, Federico Piramide, Sabrina De Cillis, Alberto Piana, Paolo Verri, Matteo Manfredi, Cristian Fiori, Enrico Vezzetti*et al. Indocyanine green drives computer vision based 3D augmented reality robot assisted partial nephrectomy : the beginning of "automatic" overlapping era.* Urology, vol. 164, pages e312–e316, 2022. (Cité en pages 99 et 177.)

[Amparore 2022b] Daniele Amparore, Federico Piramide, Angela Pecoraro, Paolo Verri, Enrico Checcucci, Sabrina De Cillis, Alberto Piana, Giovanni Busacca, Matteo Manfredi, Cristian Fiori*et al. Identification of recurrent anatomical clusters using three-dimensional virtual models for complex renal tumors with an imperative indication for nephron-sparing surgery : new technological tools for driving decision-making.* European Urology Open Science, vol. 38, pages 60–66, 2022. (Cité en pages 94 et 176.)

[Amparore 2023a] Daniele Amparore, Federico Piramide, Enrico Checcucci, Paolo Verri, Sabrina De Cillis, Alberto Piana, Gabriele Volpi, Giovanni Busacca, Marco Colombo, Cristian Fiori*et al. Three-dimensional virtual models of*

the kidney with colored perfusion regions : a new algorithm-based tool for optimizing the clamping strategy during robot-assisted partial nephrectomy. European Urology, vol. 84, no. 4, pages 418–425, 2023. (Cité en page 85.)

[Amparore 2023b] Daniele Amparore, Federico Piramide, Paolo Verri, Enrico Checcucci, Sabrina De Cillis, Alberto Piana, Gabriele Volpi, Mariano Burgio, Giovanni Busacca, Marco Colombo*et al. New generation of 3D virtual models with perfusional zones : perioperative assistance for the best pedicle management during robotic partial nephrectomy.* Current Oncology, vol. 30, no. 4, pages 4021–4032, 2023. (Cité en page 85.)

[Amparore 2024a] Daniele Amparore, Enrico Checcucci, Federico Piramide, Giovanni Busacca, Gabriele Volpi, Sabrina De Cillis, Michele Sica, Paolo Verri, Alberto Piana, Michele Di Dio*et al. Robotic vena cava thrombectomy with three-dimensional augmented reality guidance.* European urology open science, vol. 62, pages 43–46, 2024. (Cité en page 95.)

[Amparore 2024b] Daniele Amparore, Michele Sica, Paolo Verri, Federico Piramide, Enrico Checcucci, Sabrina De Cillis, Alberto Piana, Davide Campobasso, Mariano Burgio, Edoardo Cisero*et al. Computer vision and machine-learning techniques for automatic 3D virtual images overlapping during augmented reality guided robotic partial nephrectomy.* Technology in Cancer Research & Treatment, vol. 23, page 15330338241229368, 2024. (Cité en pages 51, 52, 99 et 100.)

[Andrews 2019] Jack R Andrews, Thomas Atwell, Grant Schmit, Christine M Lohse, A Nicholas Kurup, Adam Weisbrod, Matthew R Callstrom, John C Cheville, Stephen A Boorjian, Bradley C Leibovich*et al. Oncologic outcomes following partial nephrectomy and percutaneous ablation for cT1 renal masses.* European urology, vol. 76, no. 2, pages 244–251, 2019. (Cité en page 74.)

[Andrews 2020] Christopher M Andrews, Alexander B Henry, Ignacio M Soriano, Michael K Southworth et Jonathan R Silva. *Registration techniques for clinical applications of three-dimensional augmented reality devices.* IEEE journal of translational engineering in health and medicine, vol. 9, pages 1–14, 2020. (Cité en page 94.)

[Anguelov 2005] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers et James Davis. *Scape : shape completion and animation of people.* ACM SIGGRAPH 2005 Papers, pages 408–416, 2005. (Cité en page 178.)

[Ansel 2024] Jason Ansel, Edward Yang, Horace He, Natalia Gimelshein, Animesh Jain, Michael Voznesensky, Bin Bao, Peter Bell, David Berard, Evgeni Burovski*et al. Pytorch 2 : Faster machine learning through dynamic python bytecode transformation and graph compilation.* In Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2, pages 929–947, 2024. (Cité en page 243.)

[Antonelli 2019] Alessandro Antonelli, Alessandro Veccia, Carlotta Palumbo, Angelo Peroni, Giuseppe Mirabella, Alberto Cozzoli, Paolo Martucci, Filippo Ferrari, Claudio Simeone et Walter Artibani. *Holographic reconstructions for preoperative planning before partial nephrectomy : a head-to-head comparison with standard CT scan.* Urologia internationalis, vol. 102, no. 2, pages 212–217, 2019. (Cité en page 50.)

[Aoki 2018] Takeshi Aoki, Masahiko Murakami, Tomotake Koizumi, Kazuhiro Matsuda, Akira Fujimori, Tomokazu Kusano, Yuta Enami, Satoru Goto, Makoto Watanabe et Koji Otsuka. *Determination of the surgical margin in laparoscopic liver resections using infrared indocyanine green fluorescence.* Langenbeck's Archives of Surgery, vol. 403, pages 671–680, 2018. (Cité en page 48.)

[Artusi 2011] Alessandro Artusi, Francesco Banterle et Dmitry Chetverikov. *A survey of specularity removal methods.* In Computer graphics forum, volume 30, pages 2208–2230. Wiley Online Library, 2011. (Cité en page 131.)

[Arulampalam 2009] T Arulampalam, S Paterson-Brown, AJ Morris et MC Parker. *Natural orifice transluminal endoscopic surgery.* The Annals of The Royal College of Surgeons of England, vol. 91, no. 6, pages 456–459, 2009. (Cité en page 22.)

[Atasoy 2008] Selen Atasoy, David P Noonan, Selim Benhimane, Nassir Navab et Guang-Zhong Yang. *A global approach for automatic fibroscopic video mosaicing in minimally invasive diagnosis.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 850–857. Springer, 2008. (Cité en page 221.)

[Atwell 2007] Thomas D Atwell, Michael A Farrell, Matthew R Callstrom, J William Charboneau, Bradley C Leibovich, Igor Frank et David E Patterson. *Percutaneous cryoablation of large renal masses : technical feasibility and short-term outcome.* American Journal of Roentgenology, vol. 188, no. 5, pages 1195–1200, 2007. (Cité en page 77.)

[Auffenberg 2020] Gregory B Auffenberg, Michael Curry, Renee Gennarelli, Kyle A Blum, Elena Elkin et Paul Russo. *Comparison of cancer specific outcomes following minimally invasive and open surgical resection of early stage kidney cancer from a national cancer registry.* The Journal of urology, vol. 203, no. 6, pages 1094–1100, 2020. (Cité en page 74.)

[Avinash 2020] Apeksha Avinash, Alaa Eldin Abdelaal et Septimiu E Salcudean. *Evaluation of increasing camera baseline on depth perception in surgical robotics.* In 2020 IEEE International Conference on Robotics and Automation (ICRA), pages 5509–5515. IEEE, 2020. (Cité en page 131.)

[Avrithis 2014] Yannis Avrithis et Giorgos Tolias. *Hough pyramid matching : Speeded-up geometry re-ranking for large scale image retrieval.* International journal of computer vision, vol. 107, pages 1–19, 2014. (Cité en page 218.)

[Aydemir 2024] Görkay Aydemir, Weidi Xie et Fatma Güney. *Can Visual Foundation Models Achieve Long-term Point Tracking ?* arXiv preprint arXiv :2408.13575, 2024. (Cité en page 225.)

[Azad 2024] Md Abulkalam Azad, Artem Chernyshov, John Nyberg, Ingrid Tveten, Lasse Lovstakken, Håvard Dalen, Bjørnar Grenne et Andreas Østvik. *EchoTracker : Advancing Myocardial Point Tracking in Echocardiography.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 645–655. Springer, 2024. (Cité en page 225.)

[Azagra 2023] Pablo Azagra, Carlos Sostres, Ángel Ferrández, Luis Riazuelo, Clara Tomasini, O León Barbed, Javier Morlana, David Recasens, Víctor M Batlle, Juan J Gómez-Rodríguez*et al. Endomapper dataset of complete calibrated endoscopy procedures.* Scientific Data, vol. 10, no. 1, page 671, 2023. (Cité en pages 219, 226 et 227.)

[Azagury 2015] Dan E Azagury, Monica M Dua, James C Barrese, Jaimie M Henderson, Nicolas C Buchs, Frederic Ris, Jordan M Cloyd, John B Martinie, Sharif Razzaque, Stéphane Nicolau*et al. Image-guided surgery.* Curr Probl Surg, vol. 52, no. 12, pages 476–520, 2015. (Cité en pages 38 et 39.)

[Azarby 2022] Sahand Azarby et Arthur Rice. *Understanding the effects of virtual reality system usage on spatial perception : The Potential impacts of immersive virtual reality on spatial design decisions.* Sustainability, vol. 14, no. 16, page 10326, 2022. (Cité en page 131.)

[Azhar 2023] Raed A Azhar. *The influence of 3D renal reconstruction on surgical planning for complex renal tumors : An interactive case-based survey.* International braz j urol, vol. 49, pages 372–382, 2023. (Cité en page 50.)

[Azizian 2015] Mahdi Azizian et Jonathan Sorger. *System and method for specular reflection detection and reduction*, 2015. (Cité en page 37.)

[Badino 2013] Hernán Badino, Akihiro Yamamoto et Takeo Kanade. *Visual odometry by multi-frame feature integration.* In Proceedings of the IEEE International Conference on Computer Vision Workshops, pages 222–229, 2013. (Cité en page 230.)

[Bajcsy 1996] Ruzena Bajcsy, Sang Wook Lee et Aleš Leonardis. *Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation.* International Journal of Computer Vision, vol. 17, no. 3, pages 241–272, 1996. (Cité en page 132.)

[Balntas 2016] Vassileios Balntas, Edgar Riba, Daniel Ponsa et Krystian Mikolajczyk. *Learning local feature descriptors with triplets and shallow convolutional neural networks.* In Bmvc, volume 1, page 3, 2016. (Cité en page 145.)

[Bano 2019] Sophia Bano, Francisco Vasconcelos, Marcel Tella Amo, George Dwyer, Caspar Gruijthuijsen, Jan Deprest, Sebastien Ourselin, Emmanuel Vander Poorten, Tom Vercauteren et Danail Stoyanov. *Deep sequential mosaicking of fetoscopic videos.* In International conference on medical image com-

puting and computer-assisted intervention, pages 311–319. Springer, 2019. (Cité en page 221.)

[Bano 2020] Sophia Bano, Francisco Vasconcelos, Luke M Shepherd, Emmanuel Vander Poorten, Tom Vercauteren, Sebastien Ourselin, Anna L David, Jan Deprest et Danail Stoyanov. *Deep placental vessel segmentation for fetoscopic mosaicking.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 763–773. Springer, 2020. (Cité en page 221.)

[Bano 2023] Sophia Bano, Francisco Vasconcelos, Anna L David, Jan Deprest et Danail Stoyanov. *Placental vessel-guided hybrid framework for fetoscopic mosaicking.* Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization, vol. 11, no. 4, pages 1166–1171, 2023. (Cité en page 221.)

[Baranoski 2010] Gladimir VG Baranoski et Aravind Krishnaswamy. Light and skin interactions : simulations for computer graphics applications. Morgan Kaufmann, 2010. (Cité en page 123.)

[Barath 2019] Daniel Barath, Jiri Matas et Jana Noskova. *MAGSAC : marginalizing sample consensus.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 10197–10205, 2019. (Cité en page 218.)

[Barbed 2023] O León Barbed, José MM Montiel, Pascal Fua et Ana C Murillo. *Tracking adaptation to improve superpoint for 3d reconstruction in endoscopy.* In International conference on medical image computing and computer-assisted intervention, pages 583–593. Springer, 2023. (Cité en page 220.)

[Barcali 2022] Eleonora Barcali, Ernesto Iadanza, Leonardo Manetti, Piergiorgio Francia, Cosimo Nardi et Leonardo Bocchi. *Augmented reality in surgery : a scoping review.* Applied Sciences, vol. 12, no. 14, page 6890, 2022. (Cité en page 51.)

[Barr 1984] Alan H Barr. *Global and local deformations of solid primitives.* ACM Siggraph Computer Graphics, vol. 18, no. 3, pages 21–30, 1984. (Cité en pages 95, 171 et 187.)

[Barreto 2008] Joao Barreto, JM Santos, P Menezes et F Fonseca. *Ray-based calibration of rigid medical endoscopes.* In The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras-OMNIVIS, 2008. (Cité en page 110.)

[Barreto 2009] Joao Barreto, Jose Roquette, Peter Sturm et Fernando Fonseca. *Automatic camera calibration applied to medical endoscopy.* In BMVC 2009-20th British Machine Vision Conference, pages 1–10. The British Machine Vision Association (BMVA), 2009. (Cité en page 106.)

[Baumhauer 2008] Matthias Baumhauer, Tobias Simpfendörfer, Beat P Müller-Stich, Dogu Teber, Carsten N Gutt, Jens Rassweiler, H P Meinzer et Ivo Wolf. *Soft tissue navigation for laparoscopic partial nephrectomy.* Interna-

tional Journal of Computer Assisted Radiology and Surgery, vol. 3, pages 307–314, 2008. (Cité en page 95.)

[Bay 2006] Herbert Bay, Tinne Tuytelaars et Luc Van Gool. *Surf : Speeded up robust features*. In Computer Vision–ECCV 2006 : 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9, pages 404–417. Springer, 2006. (Cité en pages 97, 129, 146 et 217.)

[Bazarevsky 2019] Valentin Bazarevsky, Yury Kartynnik, Andrey Vakunov, Karthik Raveendran et Matthias Grundmann. *Blazeface : Sub-millisecond neural face detection on mobile gpus*. arXiv preprint arXiv :1907.05047, 2019. (Cité en page 183.)

[Becher 2020] Ezequiel Becher, Dora Jericevic et William C Huang. *Minimally invasive surgery for patients with locally advanced and/or metastatic renal cell carcinoma*. Urologic Clinics, vol. 47, no. 3, pages 389–397, 2020. (Cité en page 77.)

[Behrens 2013] G Behrens et MF Leitzmann. *The association between physical activity and renal cancer : systematic review and meta-analysis*. British journal of cancer, vol. 108, no. 4, pages 798–811, 2013. (Cité en page 69.)

[Bekema 2013] Hendrika J Bekema, Steven MacLennan, Mari Imamura, Thomas BL Lam, Fiona Stewart, Neil Scott, Graeme MacLennan, Sam McClinton, TR Leyshon Griffiths, Andreas Skolarikos*et al.* *Systematic review of adrenalectomy and lymph node dissection in locally advanced renal cell carcinoma*. European urology, vol. 64, no. 5, pages 799–810, 2013. (Cité en page 76.)

[Belhaoua 2016] Abdelkrim Belhaoua, Johan Moreau, Alexandre Krebs, Julien Waechter, Jean-Pierre Radoux et Jacques Marescaux. *Multiple video sequences synchronization during minimally invasive surgery*. In Medical Imaging 2016 : Image-Guided Procedures, Robotic Interventions, and Modeling, volume 9786, pages 540–548. SPIE, 2016. (Cité en page 127.)

[Benabid 1992] AL Benabid, S Lavallee, D Hoffmann, P Cinquin, J Demongeot et F Danel. *Potential use of robots in endoscopic neurosurgery*. In Minimally Invasive Neurosurgery I, pages 93–97. Springer, 1992. (Cité en page 25.)

[Bender 2014] Jan Bender, Dan Koschier, Patrick Charrier et Daniel Weber. *Position-based simulation of continuous materials*. Computers & Graphics, vol. 44, pages 1–10, 2014. (Cité en page 175.)

[Bender 2025] Johannes Bender, Jeremy Kwe, Benedikt Hoeh, Katharina Boehm, Ivan Platzek, Angelika Borkowetz, Stefanie Speidel et Micha Pfeiffer. *An augmented reality overlay for navigated prostatectomy using fiducial-free 2D–3D registration*. International Journal of Computer Assisted Radiology and Surgery, pages 1–8, 2025. (Cité en page 226.)

[Benincasa 2008] Anne B Benincasa, Logan W Clements, S Duke Herrell et Robert L Galloway. *Feasibility study for image-guided kidney surgery : Assessment of required intraoperative surface for accurate physical to image space registra-*

tions. Medical Physics, vol. 35, no. 9, pages 4251–4261, 2008. (Cité en page 96.)

[Benrhaiem 2015] Rania Benrhaiem, Sébastien Roy et Jean Meunier. *Real-time software synchronisation of webcams for live 3D tracking.* In 2015 IEEE International Conference on Image Processing (ICIP), pages 3891–3895. IEEE, 2015. (Cité en page 129.)

[Bensalah 2010] Karim Bensalah, Allan J Pantuck, Nathalie Rioux-Leclercq, Rodolphe Thuret, Francesco Montorsi, Pierre I Karakiewicz, Nicolas Mottet, Laurent Zini, Roberto Bertini, Laurent Salomon*et al. Positive surgical margin appears to have negligible impact on survival of renal cell carcinomas treated by nephron-sparing surgery.* European urology, vol. 57, no. 3, pages 466–473, 2010. (Cité en page 73.)

[Bentahila 2024] Rita Bentahila, Karim Bensalah, Nicolas Benziane-Ouaritini, Philippe Barthelemy, Nathalie Rioux-Leclerc, Jean-Michel Correas, Sarah Belhomme, Pierre Bigot et Paul Sargos. *Stereotactic body radiation therapy for primary renal cell carcinoma : a review on behalf of the CC AFU.* The French Journal of Urology, page 102660, 2024. (Cité en page 74.)

[Berger 2002] Abi Berger. *How does it work ? : Magnetic resonance imaging.* BMJ : British Medical Journal, vol. 324, no. 7328, page 35, 2002. (Cité en page 43.)

[Berndt 1994] Donald J Berndt et James Clifford. *Using dynamic time warping to find patterns in time series.* In Proceedings of the 3rd international conference on knowledge discovery and data mining, pages 359–370, 1994. (Cité en page 129.)

[Bernhard 2010] Jean-Christophe Bernhard, Allan J Pantuck, Hervé Wallerand, Maxime Crepel, Jean-Marie Ferrière, Laurent Bellec, Sylvie Maurice-Tison, Grégoire Robert, Baptiste Albouy, Gilles Pasticier*et al. Predictive factors for ipsilateral recurrence after nephron-sparing surgery in renal cell carcinoma.* European urology, vol. 57, no. 6, pages 1080–1086, 2010. (Cité en page 73.)

[Bernhard 2016] Jean-Christophe Bernhard, Shuji Isotani, Toru Matsugasumi, Vinay Duddalwar, Andrew J Hung, Evren Suer, Eduard Baco, Raj Satkunasivam, Hooman Djaladat, Charles Metcalfe*et al. Personalized 3D printed model of kidney and tumor anatomy : a useful tool for patient education.* World journal of urology, vol. 34, pages 337–345, 2016. (Cité en page 46.)

[Bernhard 2022] Jean-Christophe Bernhard, Grégoire Robert, Solène Ricard, Clément Michiels, Grégoire Capon, Astrid Boulenger de Hautecloque, Henri Bensadoun, Joséphine Gay, Julien Rogier, Patrick Tauzin-Fin*et al. Day-case robotic-assisted partial nephrectomy : feasibility and preliminary results of a prospective evaluation (UroCCR-25 AMBU-REIN study).* World Journal of Urology, pages 1–7, 2022. (Cité en page 81.)

[Bernhard 2023] Jean-Christophe Bernhard, Grégoire Robert, Solène Ricard, Julien Rogier, Cécile Degryse, Clément Michiels, Gaëlle Margue, Peggy Blanc, Eric Alezra, Vincent Estrade*et al. Nurse-led coordinated surgical care pathways for*

cost optimization of robotic-assisted partial nephrectomy : Medico-economic analysis of the UroCCR-25 AMBU-REIN study. World Journal of Urology, vol. 41, no. 2, pages 325–333, 2023. (Cité en page 81.)

[Bertolo 2020] Riccardo Bertolo, Andrew Hung, Francesco Porpiglia, Pierluigi Bove, Mary Schleicher et Prokar Dasgupta. *Systematic review of augmented reality in urological interventions : the evidences of an impact on surgical outcomes are yet to come.* World journal of urology, vol. 38, pages 2167–2176, 2020. (Cité en page 94.)

[Bertolo 2023] Riccardo Bertolo, Alessio Pecoraro, Umberto Carbonara, Daniele Amparore, Pietro Diana, Stijn Muselaers, Michele Marchioni, Maria Carmen Mir, Alessandro Antonelli, Ketan Badani*et al. Resection techniques during robotic partial nephrectomy : a systematic review.* European Urology Open Science, vol. 52, pages 7–21, 2023. (Cité en page 79.)

[Bertrand 2020] Le Roy Bertrand, Mourad Abdallah, Yamid Espinel, Lilian Calvet, Bruno Pereira, Erol Ozgur, Denis Pezet, Emmanuel Buc et Adrien Bartoli. *A case series study of augmented reality in laparoscopic liver resection with a deformable preoperative model.* Surgical endoscopy, vol. 34, pages 5642–5648, 2020. (Cité en page 172.)

[Bex 2025] A Bex, L Albiges, J Bedke, U Capitanio, S Dabestani, M Hora, T Klatte, T Kuusk, L Lund, L Marconi*et al. Renal Cell Carcinoma.* In European Association of Urology guidelines, 2025. (Cité en page 169.)

[Bhat 2023] Shariq Farooq Bhat, Reiner Birkl, Diana Wofk, Peter Wonka et Matthias Müller. *Zoedepth : Zero-shot transfer by combining relative and metric depth.* arXiv preprint arXiv :2302.12288, 2023. (Cité en page 223.)

[Bhindi 2018] Bimal Bhindi, Christopher JD Wallis, Stephen A Boorjian, R Houston Thompson, Ann Farrell, Simon P Kim, Jose A Karam, Umberto Capitanio, Dragan Golijanin, Bradley C Leibovich*et al. The role of lymph node dissection in the management of renal cell carcinoma : a systematic review and meta-analysis.* BJU international, vol. 121, no. 5, pages 684–698, 2018. (Cité en page 77.)

[Bian 2017] JiaWang Bian, Wen-Yan Lin, Yasuyuki Matsushita, Sai-Kit Yeung, Tan-Dat Nguyen et Ming-Ming Cheng. *GMS : Grid-based motion statistics for fast, ultra-robust feature correspondence.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4181–4190, 2017. (Cité en pages 99, 218 et 256.)

[Bianchi 2021] Lorenzo Bianchi, Francesco Chessa, Andrea Angiolini, Laura Cercenelli, Simone Lodi, Barbara Bortolani, Enrico Molinaroli, Carlo Casablanca, Matteo Droghetti, Caterina Gaudiano*et al. The use of augmented reality to guide the intraoperative frozen section during robot-assisted radical prostatectomy.* European Urology, vol. 80, no. 4, pages 480–488, 2021. (Cité en page 171.)

[Bigot 2022] P Bigot, P Barthelemy, R Boissier, Z-E Khene, C Pettenati, J-C Bernhard, J-M Correas, N Doumerc, P Gimel, A Ingels*et al. French AFU Cancer committee guidelines-update 2022-2024 : management of kidney cancer.* Progrès en Urologie, vol. 32, pages 1195–1274, 2022. (Cité en page 78.)

[Birlo 2022] Manuel Birlo, PJ Eddie Edwards, Matthew Clarkson et Danail Stoyanov. *Utility of optical see-through head mounted displays in augmented reality-assisted surgery : A systematic review.* Medical Image Analysis, vol. 77, page 102361, 2022. (Cité en page 51.)

[Bishoff 2007] Jay T Bishoff et Louis R Kavoussi. *Laparoscopic surgery of the kidney.* Campbell-walsh urology, vol. 9, pages 1759–1809, 2007. (Cité en page 83.)

[Blinn 1977] James F Blinn. *Models of light reflection for computer synthesized pictures.* In Proceedings of the 4th annual conference on Computer graphics and interactive techniques, pages 192–198, 1977. (Cité en page 86.)

[Blom 2009] Jan HM Blom, Hein van Poppel, Jean M Maréchal, Didier Jacqmin, Fritz H Schröder, Linda de Prijck, Richard Sylvester*et al. Radical nephrectomy with and without lymph-node dissection : final results of European Organization for Research and Treatment of Cancer (EORTC) randomized phase 3 trial 30881.* European urology, vol. 55, no. 1, pages 28–34, 2009. (Cité en page 76.)

[Bobrow 2023] Taylor L Bobrow, Mayank Golhar, Rohan Vijayan, Venkata S Akshintala, Juan R Garcia et Nicholas J Durr. *Colonoscopy 3D video dataset with paired depth from 2D-3D registration.* Medical image analysis, vol. 90, page 102956, 2023. (Cité en page 226.)

[Bochkovskiy 2020] Alexey Bochkovskiy, Chien-Yao Wang et Hong-Yuan Mark Liao. *Yolov4 : Optimal speed and accuracy of object detection.* arXiv preprint arXiv :2004.10934, 2020. (Cité en page 132.)

[Boguet 2007] J Boguet, B Lepage et Y Rochais. *Mise au point en chirurgie endoscopique transluminale ou notes.* IRBM News, vol. 28, no. 5-6, pages 41–46, 2007. (Cité en page 22.)

[Bolelli 2019] Federico Bolelli, Stefano Allegretti, Lorenzo Baraldi et Costantino Grana. *Spaghetti labeling : Directed acyclic graphs for block-based connected components labeling.* IEEE Transactions on Image Processing, vol. 29, pages 1999–2012, 2019. (Cité en page 137.)

[Borghesi 2013] Marco Borghesi, Riccardo Schiavina, Melanie Gan, Giacomo Novara, Alexandre Mottrie et Vincenzo Ficarra. *Expanding utilization of robotic partial nephrectomy for clinical T1b and complex T1a renal masses.* World journal of urology, vol. 31, pages 499–504, 2013. (Cité en page 75.)

[Borregales 2016] Leonardo D Borregales, Dae Y Kim, Angie L Staller, Wei Qiao, Arun Z Thomas, Mehrad Adibi, Pheroze Tamboli, Kanishka Sircar, Eric Jonasch, Nizar M Tannir*et al. Prognosticators and outcomes of patients with renal cell carcinoma and adjacent organ invasion treated with radical*

nephrectomy. In Urologic Oncology : Seminars and Original Investigations, volume 34, pages 237–e19. Elsevier, 2016. (Cité en page 78.)

[Bosniak 2012] Morton A Bosniak. *The Bosniak renal cyst classification : 25 years later*. Radiology, vol. 262, no. 3, pages 781–785, 2012. (Cité en page 71.)

[Boudet 2017] Samuel Boudet. *Anatomie Uro-Génitale*, 2017. Consulté le 22 novembre 2024. (Cité en pages 57, 58, 59, 61, 62, 63 et 65.)

[Boul-Atarass 2025] Imán Laga Boul-Atarass, Carmen Cepeda Franco, José Domingo Sanmartín Sierra, Javier Castell Monsalve et Javier Padillo Ruiz. *Virtual 3D models, augmented reality systems and virtual laparoscopic simulations in complicated pancreatic surgeries : state of art, future perspectives, and challenges*. International Journal of Surgery, pages 10–1097, 2025. (Cité en page 171.)

[Boulenger De Hauteclocque 2023] Astrid Boulenger De Hauteclocque, Loïc Ferrer, Damien Ambrosetti, Solene Ricard, Pierre Bigot, Karim Bensalah, François Henon, Nicolas Doumerc, Arnaud Méjean, Virginie Verkarre*et al.* *Machine-learning approach for prediction of pT3a upstaging and outcomes of localized renal cell carcinoma (UroCCR-15)*. BJU international, vol. 132, no. 2, pages 160–169, 2023. (Cité en page 76.)

[Bourdel 2019] Nicolas Bourdel, Pauline Chauvet, Lilian Calvet, Benoit Magnin, Adrien Bartoli et Michel Canis. *Use of augmented reality in gynecologic surgery to visualize adenomyomas*. Journal of minimally invasive gynecology, vol. 26, no. 6, pages 1177–1180, 2019. (Cité en page 170.)

[Bradley 2009] Derek Bradley, Bradley Atcheson, Ivo Ihrke et Wolfgang Heidrich. *Synchronization and rolling shutter compensation for consumer video camera arrays*. In 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pages 1–8. IEEE, 2009. (Cité en page 129.)

[Bradshaw 2020] Aaron W Bradshaw, Riccardo Autorino, Giuseppe Simone, Bo Yang, Robert G Uzzo, Francesco Porpiglia, Umberto Capitanio, James Porter, Riccardo Bertolo, Andrea Minervini*et al.* *Robotic partial nephrectomy vs minimally invasive radical nephrectomy for clinical T2a renal mass : a propensity score-matched comparison from the ROSULA (Robotic Surgery for Large Renal Mass) Collaborative Group*. BJU international, vol. 126, no. 1, pages 114–123, 2020. (Cité en page 76.)

[Bragayrac 2016] Luciano A Nunez Bragayrac, Daniel Abbotoy, Kristopher Attwood, Fadi Darwiche, Jan Hoffmeyer, Eric C Kauffman et Thomas Schwaab. *Outcomes of minimal invasive vs open radical nephrectomy for the treatment of locally advanced renal-cell carcinoma*. Journal of Endourology, vol. 30, no. 8, pages 871–876, 2016. (Cité en page 77.)

[Branch 1982] Robert A Branch. *Drugs as indicators of hepatic function*. Hepatology, vol. 2, no. 1, pages 97S–105S, 1982. (Cité en page 48.)

[Brassier 2022] Marie Brassier, Zine-Eddine Khene, Jean-Christophe Bernhard, Idir Ouzaid, François Xavier Nouhaud, Jonathan Olivier, Cosmina Nedelcu, Nico-

las Grenier, Luc Beuzit, Nicolas Doumer*et al. Percutaneous Ablation Versus Surgical Resection for Local Recurrence Following Partial Nephrectomy for Renal Cell Cancer : A Propensity Score Analysis (REPART Study—UroCCR 71)*. European Urology Focus, vol. 8, no. 1, pages 210–216, 2022. (Cité en page 74.)

[Bravi 2021] Carlo Andrea Bravi, Alessandro Larcher, Umberto Capitanio, Andrea Mari, Alessandro Antonelli, Walter Artibani, Maurizio Barale, Roberto Bertini, Pierluigi Bove, Eugenio Brunocilla*et al. Perioperative outcomes of open, laparoscopic, and robotic partial nephrectomy : a prospective multicenter observational study (The RECORd 2 Project)*. European urology focus, vol. 7, no. 2, pages 390–396, 2021. (Cité en page 81.)

[Bricault 2017] Ivan Bricault. Principes de la tomodensitométrie. Université de Grenoble, 2017. (Cité en pages 40 et 41.)

[Brisson 2023] Gabriel F. Brisson, II Gregory W. Dachs, Niels Smaby et Melody Wu. *Systems and methods for instrument engagement*, 2023. (Cité en pages 37 et 38.)

[Brito 2011] Darlan N Brito, Flávio LC Pádua, Guilherme AS Pereira et Rodrigo L Carceroni. *Temporal synchronization of non-overlapping videos using known object motion*. Pattern recognition letters, vol. 32, no. 1, pages 38–46, 2011. (Cité en page 129.)

[Brown III 2013] William Brown III, Po-Yin Yen, Marlene Rojas et Rebecca Schnall. *Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology*. Journal of biomedical informatics, vol. 46, no. 6, pages 1080–1087, 2013. (Cité en page 92.)

[Brown 1996] Duane Brown. *Decentering distortion of lenses*. Photogrammetric engineering, vol. 32, no. 3, pages 444–462, 1996. (Cité en page 110.)

[Bucholz 1995] Richard D Bucholz. *Introduction to journal of image guided surgery*. Journal of image guided surgery, vol. 1, no. 1, pages 1–3, 1995. (Cité en page 38.)

[Budd 2024] Charlie Budd et Tom Vercauteren. *Transferring Relative Monocular Depth to Surgical Vision with Temporal Consistency*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 692–702. Springer, 2024. (Cité en page 109.)

[Buhmann 2000] Martin Dietrich Buhmann. *Radial basis functions*. Acta numerica, vol. 9, pages 1–38, 2000. (Cité en page 175.)

[Bukavina 2022] Laura Bukavina, Karim Bensalah, Freddie Bray, Maria Carlo, Ben Challacombe, Jose A Karam, Wassim Kassouf, Thomas Mitchell, Rodolfo Montironi, Tim O'Brien*et al. Epidemiology of renal cell carcinoma : 2022 update*. European urology, vol. 82, no. 5, pages 529–542, 2022. (Cité en pages 68 et 69.)

[Burschka 2004] Darius Burschka, Ming Li, Russell Taylor et Gregory D Hager. *Scale-invariant registration of monocular stereo images to 3d surface models*.

In 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566), volume 3, pages 2581–2586. IEEE, 2004. (Cité en page 222.)

[Burt 1983] Peter J Burt et Edward H Adelson. *A multiresolution spline with application to image mosaics*. ACM Transactions on Graphics (ToG), vol. 2, no. 4, pages 217–236, 1983. (Cité en page 221.)

[Calonder 2010] Michael Calonder, Vincent Lepetit, Christoph Strecha et Pascal Fua. *Brief : Binary robust independent elementary features*. In Computer Vision–ECCV 2010 : 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11, pages 778–792. Springer, 2010. (Cité en pages 145, 217 et 221.)

[Çamlıdağ 2019] İlkay Çamlıdağ, Mehmet Selim Nural, Murat Danacı et Ender Özden. *Usefulness of rapid kV-switching dual energy CT in renal tumor characterization*. Abdominal Radiology, vol. 44, pages 1841–1849, 2019. (Cité en page 71.)

[Campi 2021] Riccardo Campi, Grant D Stewart, Michael Staehler, Saeed Dabestani, Markus A Kuczyk, Brian M Shuch, Antonio Finelli, Axel Bex, Börje Ljungberg et Umberto Capitanio. *Novel liquid biomarkers and innovative imaging for kidney cancer diagnosis : what can be implemented in our practice today ? A systematic review of the literature*. European urology oncology, vol. 4, no. 1, pages 22–41, 2021. (Cité en page 70.)

[Campi 2023] Riccardo Campi, Giacomo Rebez, Tobias Klatte, Eduard Roussel, Idir Ouizad, Alexander Ingels, Nicola Pavan, Onder Kara, Selcuk Erdem, Riccardo Bertolo*et al.* *Effect of smoking, hypertension and lifestyle factors on kidney cancer—perspectives for prevention and screening programmes*. Nature Reviews Urology, vol. 20, no. 11, pages 669–681, 2023. (Cité en page 70.)

[Cannon 2023] Piper C Cannon, James M Ferguson, E Bryn Pitt, Jason A Shrand, Shaan A Setia, Naren Nimmagadda, Eric J Barth, Nicholas L Kavoussi, Robert L Galloway, S Duke Herrell*et al.* *A safe framework for quantitative in vivo human evaluation of image guidance*. IEEE Open Journal of Engineering in Medicine and Biology, vol. 5, pages 133–139, 2023. (Cité en page 100.)

[Canny 2009] John Canny. *A computational approach to edge detection*. IEEE Transactions on pattern analysis and machine intelligence, pages 679–698, 2009. (Cité en page 145.)

[Cao 2006] Xiaochun Cao, Jiangjian Xiao et Hassan Foroosh. *Camera motion quantification and alignment*. In 18th International Conference on Pattern Recognition (ICPR'06), volume 2, pages 13–16. IEEE, 2006. (Cité en page 129.)

[Cao 2009] Yanpeng Cao et John McDonald. *Viewpoint invariant features from single images using 3D geometry*. In 2009 Workshop on Applications of Computer Vision (WACV), pages 1–6. IEEE, 2009. (Cité en page 230.)

[Cao 2010] Xiaochun Cao, Lin Wu, Jiangjian Xiao, Hassan Foroosh, Jigui Zhu et Xiaohong Li. *Video synchronization and its application to object transfer*.

Image and Vision Computing, vol. 28, no. 1, pages 92–100, 2010. (Cité en page 129.)

[Cao 2017] De-Hong Cao, Liang-Ren Liu, Yu Fang, Ping Tang, Tao Li, YunJin Bai, Jia Wang et Qiang Wei. *Simple tumor enucleation may not decrease oncologic outcomes for T1 renal cell carcinoma : A systematic review and meta-analysis*. In Urologic Oncology : Seminars and Original Investigations, volume 35, pages 661–e15. Elsevier, 2017. (Cité en page 79.)

[Capitanio 2009] Umberto Capitanio, Paul Perrotte, Laurent Zini, Claudio Jeldres, Shahrokh F Shariat, Hendrik Isbarn, Philippe Arjane, François Peloquin, Daniel Pharand, Francesco Montorsi*et al. Nephrectomy improves survival in patients with invasion of adjacent viscera and absence of nodal metastases (stage T4N0 renal cell carcinoma)*. BJU international, vol. 104, no. 6, pages 795–799, 2009. (Cité en page 78.)

[Capitanio 2014] Umberto Capitanio, Nazareno Suardi, Rayan Matloob, Marco Roscigno, Firas Abdollah, Ettore Di Trapani, Marco Moschini, Andrea Gallina, Andrea Salonia, Alberto Brigantiet *al. Extent of lymph node dissection at nephrectomy affects cancer-specific survival and metastatic progression in specific sub-categories of patients with renal cell carcinoma (RCC)*. BJU international, vol. 114, no. 2, 2014. (Cité en page 76.)

[Capitanio 2015] Umberto Capitanio, Carlo Terrone, Alessandro Antonelli, Andrea Minervini, Alessandro Volpe, Maria Furlan, Rayan Matloob, Federica Regis, Cristian Fiori, Francesco Porpigliaet *al. Nephron-sparing techniques independently decrease the risk of cardiovascular events relative to radical nephrectomy in patients with a T1a–T1b renal mass and normal preoperative renal function*. European urology, vol. 67, no. 4, pages 683–689, 2015. (Cité en pages 75 et 79.)

[Capitanio 2019] Umberto Capitanio, Karim Bensalah, Axel Bex, Stephen A Boorjian, Freddie Bray, Jonathan Coleman, John L Gore, Maxine Sun, Christopher Wood et Paul Russo. *Epidemiology of renal cell carcinoma*. European urology, vol. 75, no. 1, pages 74–84, 2019. (Cité en pages 68 et 69.)

[Carlo 2022] Maria I Carlo, Kyrollis Attalla, Yousef Mazaheri, Sounak Gupta, Onur Yildirim, Samuel J Murray, Devyn T Coskey, Ritesh Kotecha, Chung-Han Lee, Darren R Feldmanet *al. Phase II study of neoadjuvant nivolumab in patients with locally advanced clear cell renal cell carcinoma undergoing nephrectomy*. European urology, vol. 81, no. 6, pages 570–573, 2022. (Cité en page 77.)

[Caron 2021] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski et Armand Joulin. *Emerging properties in self-supervised vision transformers*. In Proceedings of the IEEE/CVF international conference on computer vision, pages 9650–9660, 2021. (Cité en page 225.)

[Carr 2019] Brendan M Carr, Jennifer A Lyon, Jamie Romeiser, Mark Talamini et A Laurie W Shroyer. *Laparoscopic versus open surgery : a systematic review evaluating Cochrane systematic reviews.* Surgical endoscopy, vol. 33, pages 1693–1709, 2019. (Cité en page 24.)

[Carreira 2016] Joao Carreira, Pulkit Agrawal, Katerina Fragkiadaki et Jitendra Malik. *Human pose estimation with iterative error feedback.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4733–4742, 2016. (Cité en page 202.)

[Cartucho 2021] João Cartucho, Samyakh Tukra, Yunpeng Li, Daniel S. Elson et Stamatia Giannarou. *VisionBlender : a tool to efficiently generate computer vision datasets for robotic surgery.* Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization, vol. 9, no. 4, pages 331–338, 2021. (Cité en page 226.)

[Cartucho 2024] João Cartucho, Alistair Weld, Samyakh Tukra, Haozheng Xu, Hiroki Matsuzaki, Taiyo Ishikawa, Minjun Kwon, Yong Eun Jang, Kwang-Ju Kim, Gwang Lee*et al.* *Surgt challenge : Benchmark of soft-tissue trackers for robotic surgery.* Medical image analysis, vol. 91, page 102985, 2024. (Cité en pages 222 et 226.)

[Caspi 2000] Yaron Caspi et Michal Irani. *A step towards sequence-to-sequence alignment.* In Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662), volume 2, pages 682–689. IEEE, 2000. (Cité en page 129.)

[Caspi 2002a] Yaron Caspi et Michal Irani. *Aligning non-overlapping sequences.* International Journal of Computer Vision, vol. 48, pages 39–51, 2002. (Cité en pages 129 et 147.)

[Caspi 2002b] Yaron Caspi et Michal Irani. *Spatio-temporal alignment of sequences.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 11, pages 1409–1424, 2002. (Cité en page 129.)

[Caspi 2006] Yaron Caspi, Denis Simakov et Michal Irani. *Feature-based sequence-to-sequence matching.* International Journal of Computer Vision, vol. 68, pages 53–64, 2006. (Cité en page 129.)

[Cavalli 2020] Luca Cavalli, Viktor Larsson, Martin Ralf Oswald, Torsten Sattler et Marc Pollefeys. *Adalam : Revisiting handcrafted outlier detection.* arXiv preprint arXiv :2006.04250, 2020. (Cité en page 218.)

[Cazalas 2023] Grégoire Cazalas, Clément Klein, Gilles Piana, Eric De Kerviler, Afshin Gangi, Philippe Puech, Cosmina Nedelcu, Remi Grange, Xavier Buy, Marc-Antoine Jegonday*et al.* *A multicenter comparative matched-pair analysis of percutaneous tumor ablation and robotic-assisted partial nephrectomy of T1b renal cell carcinoma (AblatT1b study—UroCCR 80).* European Radiology, vol. 33, no. 9, pages 6513–6521, 2023. (Cité en page 77.)

[Ceccariglia 2022] Francesco Ceccariglia, Laura Cercenelli, Giovanni Badiali, Emanuela Marcelli et Achille Tarsitano. *Application of augmented reality to maxil-*

lary resections : a three-dimensional approach to maxillofacial oncologic surgery. Journal of Personalized Medicine, vol. 12, no. 12, page 2047, 2022. (Cité en pages 51 et 52.)

[Centre Hospitalier Universitaire de Clermont-Ferrand 2024] Centre Hospitalier Universitaire de Clermont-Ferrand. *Rapport d'activité 2023*, Octobre 2024. (Cité en page 11.)

[Chandelon 2022] Kilian Chandelon et Adrien Bartoli. *Tracking better, tracking longer : automatic keyframe selection in model-based laparoscopic augmented reality.* International Journal of Computer Assisted Radiology and Surgery, vol. 17, no. 8, pages 1507–1511, 2022. (Cité en pages 54, 197 et 206.)

[Chandelon 2023] Kilian Chandelon, Rasoul Sharifian, Salomé Marchand, Abderrahmane Khaddad, Nicolas Bourdel, Nicolas Mottet, Jean-Christophe Bernhard et Adrien Bartoli. *Kidney tracking for live augmented reality in stereoscopic mini-invasive partial nephrectomy.* Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization, vol. 11, no. 4, pages 1251–1260, 2023. (Cité en page 54.)

[Chandelon 2024] Kilian Chandelon et Adrien Bartoli. *Synchronising a stereoscopic surgical video stream using specular reflection.* International Journal of Computer Assisted Radiology and Surgery, pages 1–11, 2024. (Cité en page 53.)

[Chandelon 2025] Kilian Chandelon, Alice Pitout, Mathieu Souchaud, Julie Desternes, Gaëlle Margue, Julien Peyras, Nicolas Bourdel, Jean-Christophe Bernhard et Adrien Bartoli. *Landmark-free automatic digital twin registration in robot-assisted partial nephrectomy using a generic end-to-end model.* International Journal of Computer Assisted Radiology and Surgery, pages 1–10, 2025. (Cité en page 54.)

[Chang 2015] Xiaofeng Chang, Fan Zhang, Tieshi Liu, Changwei Ji, Xiaozhi Zhao, Rong Yang, Xiang Yan, Wei Wang et Hongqian Guo. *Radio frequency ablation versus partial nephrectomy for clinical T1b renal cell carcinoma : long-term clinical and oncologic outcomes.* The Journal of urology, vol. 193, no. 2, pages 430–435, 2015. (Cité en page 77.)

[Chang 2018] Ki Don Chang, Ali Abdel Raheem, Kwang Hyun Kim, Cheol Kyu Oh, Sung Yul Park, Young Sik Kim, Won Sik Ham, Woong Kyu Han, Young Deuk Choi, Byung Ha Chung*et al. Functional and oncological outcomes of open, laparoscopic and robot-assisted partial nephrectomy : a multicentre comparative matched-pair analyses with a median of 5 years' follow-up.* BJU international, vol. 122, no. 4, pages 618–626, 2018. (Cité en page 81.)

[Chauvet 2018] Pauline Chauvet, Toby Collins, Clement Debize, Lorraine Novais-Gameiro, Bruno Pereira, Adrien Bartoli, Michel Canis et Nicolas Bourdel. *Augmented reality in a tumor resection model.* Surgical endoscopy, vol. 32, pages 1192–1201, 2018. (Cité en page 216.)

[Chauvet 2020] Pauline Chauvet, Nicolas Bourdel, Lilian Calvet, Benoit Magnin, Guillaume Teluob, Michel Canis et Adrien Bartoli. *Augmented reality with*

diffusion tensor imaging and tractography during laparoscopic myomectomies. Journal of minimally invasive gynecology, vol. 27, no. 4, pages 973–976, 2020. (Cité en pages 51, 52 et 170.)

[Checcucci 2023] Enrico Checcucci, Alberto Piana, Gabriele Volpi, Pietro Piazzolla, Daniele Amparore, Sabrina De Cillis, Federico Piramide, Cecilia Gatti, Ilaria Stura, Enrico Bollito, Federica Massa, Michele Di Dio, Cristian Fiori et Francesco Porpiglia. *Three-dimensional automatic artificial intelligence driven augmented-reality selective biopsy during nerve-sparing robot-assisted radical prostatectomy : A feasibility and accuracy study.* Asian Journal of Urology, vol. 10, no. 4, pages 407–415, 2023. (Cité en pages 176 et 177.)

[Cheema 2018] Muhammad Nadeem Cheema, Anam Nazir, Bin Sheng, Ping Li, Jing Qin, Jinman Kim et David Dagan Feng. *Image-aligned dynamic liver reconstruction using intra-operative field of views for minimal invasive surgery.* IEEE Transactions on Biomedical Engineering, vol. 66, no. 8, pages 2163–2173, 2018. (Cité en page 222.)

[Chen 2014] Yuanbo Chen, Hulin Li, Dingtao Wu, Keming Bi et Chunxiao Liu. *Surgical planning and manual image fusion based on 3D model facilitate laparoscopic partial nephrectomy for intrarenal tumors.* World journal of urology, vol. 32, pages 1493–1499, 2014. (Cité en pages 94, 100 et 176.)

[Chen 2016] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever et Pieter Abbeel. *Infogan : Interpretable representation learning by information maximizing generative adversarial nets.* Advances in neural information processing systems, vol. 29, 2016. (Cité en page 223.)

[Chen 2017] Xin-Zu Chen, Shao-Yong Wang, Yin-Su Wang, Zi-Han Jiang, Wei-Han Zhang, Kai Liu, Kun Yang, Xiao-Long Chen, Lin-Yong Zhao, Meng Qiu*et al. Comparisons of short-term and survival outcomes of laparoscopy-assisted versus open total gastrectomy for gastric cancer patients.* Oncotarget, vol. 8, no. 32, page 52366, 2017. (Cité en page 24.)

[Chen 2020] Xin Chen, Xiangfei Wu, Shiqi Gao, Xiaomei Xie et Ya Huang. *Synchronization and calibration of a stereo vision system.* In Global Oceans 2020 : Singapore–US Gulf Coast, pages 1–6. IEEE, 2020. (Cité en page 127.)

[Chen 2021a] Jiaxuan Chen, Shuang Chen, Xiaoxian Chen, Yang Yang, Linjie Xing, Xiaoyan Fan et Yujing Rao. *LSV-ANet : Deep learning on local structure visualization for feature matching.* IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pages 1–18, 2021. (Cité en page 219.)

[Chen 2021b] Lin Chen, Franz Rottensteiner et Christian Heipke. *Feature detection and description for image matching : from hand-crafted design to deep learning.* Geo-spatial Information Science, vol. 24, no. 1, pages 58–74, 2021. (Cité en page 219.)

[Chen 2022] Hongkai Chen, Zixin Luo, Lei Zhou, Yurun Tian, Mingmin Zhen, Tian Fang, David Mckinnon, Yanghai Tsin et Long Quan. *Aspanformer : Detector-*

free image matching with adaptive span transformer. In European Conference on Computer Vision, pages 20–36. Springer, 2022. (Cité en page 146.)

[Chen 2024a] Junyu Chen, Yihao Liu, Shuwen Wei, Zhangxing Bian, Shalini Subramanian, Aaron Carass, Jerry L Prince et Yong Du. *A survey on deep learning in medical image registration : New technologies, uncertainty, evaluation metrics, and beyond*. Medical Image Analysis, page 103385, 2024. (Cité en page 44.)

[Chen 2024b] Weirong Chen, Le Chen, Rui Wang et Marc Pollefeys. *Leap-vo : Long-term effective any point tracking for visual odometry*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 19844–19853, 2024. (Cité en page 225.)

[Chen 2025] Sili Chen, Hengkai Guo, Shengnan Zhu, Feihu Zhang, Zilong Huang, Jiashi Feng et Bingyi Kang. *Video Depth Anything : Consistent Depth Estimation for Super-Long Videos*. arXiv preprint arXiv :2501.12375, 2025. (Cité en page 109.)

[Chernyshov 2025a] Artem Chernyshov. *Automated analysis of the right ventricle with efficient deep learning methods in 2D echocardiography*. 2025. (Cité en page 225.)

[Chernyshov 2025b] Artem Chernyshov, John Nyberg, Vegard Holmstrøm, Md Abulkalam Azad, Bjørnar Grenne, Håvard Dalen, Svein Arne Aase, Lasse Lovstakken et Andreas Østvik. *Low Complexity Point Tracking of the Myocardium in 2D Echocardiography*. arXiv preprint arXiv :2503.10431, 2025. (Cité en page 225.)

[Ching 1993] W-S Ching, P-S Toh et M-H Er. *A new specular highlights detection algorithm using multiple views*. In Proceedings of IEEE Singapore International Conference on Networks/International Conference on Information Engineering'93, volume 1, pages 474–478. IEEE, 1993. (Cité en page 131.)

[Cho 2024] Seokju Cho, Jiahui Huang, Seungryong Kim et Joon-Young Lee. *Flowtrack : Revisiting optical flow for long-range dense tracking*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 19268–19277, 2024. (Cité en page 225.)

[Choi 2015] Ji Eun Choi, Ji Hye You, Dae Keun Kim, Koon Ho Rha et Seon Heui Lee. *Comparison of perioperative outcomes between robotic and laparoscopic partial nephrectomy : a systematic review and meta-analysis*. European urology, vol. 67, no. 5, pages 891–901, 2015. (Cité en pages 76 et 81.)

[Christensson 2013] Anders Christensson, Caroline Savage, Daniel D Sjoberg, Angel M Cronin, M Frank O'Brien, William Lowrance, Peter M Nilsson, Andrew J Vickers, Paul Russo et Hans Lilja. *Association of cancer with moderately impaired renal function at baseline in a large, representative, population-based cohort followed for up to 30 years*. International journal of cancer, vol. 133, no. 6, pages 1452–1458, 2013. (Cité en page 69.)

[Chum 2003] Ondřej Chum, Jiří Matas et Josef Kittler. *Locally optimized RANSAC*. In Joint pattern recognition symposium, pages 236–243. Springer, 2003. (Cité en page 218.)

[Chum 2005] Ondrej Chum et Jiri Matas. *Matching with PROSAC-progressive sample consensus*. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), volume 1, pages 220–226. IEEE, 2005. (Cité en pages 218 et 219.)

[Cignoli 2023] Daniele Cignoli, Giuseppe Basile, Giuseppe Fallara, Giuseppe Rosiello, Federico Belladelli, Francesco Cei, Giacomo Musso, Chiara Re, Roberto Bertini, Pierre Karakiewicz*et al.* *Risks and benefits of partial nephrectomy performed with limited or with zero ischaemia time*. BJU international, vol. 132, no. 3, pages 283–290, 2023. (Cité en page 81.)

[Collins 2014] Toby Collins, Daniel Pizarro, Adrien Bartoli, Michel Canis et Nicolas Bourdel. *Computer-assisted laparoscopic myomectomy by augmenting the uterus with pre-operative MRI data*. In 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 243–248. IEEE, 2014. (Cité en pages 246 et 247.)

[Collins 2016] Toby Collins, Adrien Bartoli, Nicolas Bourdel et Michel Canis. *Robust, real-time, dense and deformable 3D organ tracking in laparoscopic videos*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 404–412. Springer, 2016. (Cité en page 221.)

[Collins 2020] Toby Collins, Daniel Pizarro, Simone Gasparini, Nicolas Bourdel, Pauline Chauvet, Michel Canis, Lilian Calvet et Adrien Bartoli. *Augmented reality guided laparoscopic surgery of the uterus*. IEEE Transactions on Medical Imaging, vol. 40, no. 1, pages 371–380, 2020. (Cité en pages 51, 52, 172, 173, 174, 197, 206, 216, 243 et 246.)

[Conrady 1919] Alexander Eugen Conrady. *Decentred lens-systems*. Monthly notices of the royal astronomical society, vol. 79, no. 5, pages 384–390, 1919. (Cité en page 110.)

[Conseil National de l'Ordre des Médecins 2024a] Conseil National de l'Ordre des Médecins. *Atlas de la Démographie Médicale en France 2024 - Tome 1*, 2024. (Cité en pages 4, 5 et 6.)

[Conseil National de l'Ordre des Médecins 2024b] Conseil National de l'Ordre des Médecins. *Atlas de la Démographie Médicale en France 2024 - Tome 2*, 2024. (Cité en pages 6 et 7.)

[Cootes 1992] Timothy F Cootes, Christopher J Taylor, David H Cooper et Jim Graham. *Training models of shape from sets of examples*. In BMVC92 : Proceedings of the British Machine Vision Conference, organised by the British Machine Vision Association 22–24 September 1992 Leeds, pages 9–18. Springer, 1992. (Cité en page 175.)

[Cormier 2021] Luc Cormier, Antoine Valeri*et al.* Reins et voies urinaires-appareil génital masculin : Enseignement intégré. Elsevier Health Sciences, 2021. (Cité en page 66.)

[Cornelis 2017] Francois Cornelis et Nicolas Grenier. *Multiparametric magnetic resonance imaging of solid renal tumors : a practical algorithm.* In Seminars in Ultrasound, CT and MRI, volume 38, pages 47–58. Elsevier, 2017. (Cité en page 71.)

[Correa 2019] Rohann JM Correa, Alexander V Louie, Nicholas G Zaorsky, Eric J Lehrer, Rodney Ellis, Lee Ponsky, Irving Kaplan, Anand Mahadevan, William Chu, Anand Swaminath*et al. The emerging role of stereotactic ablative radiotherapy for primary renal cell carcinoma : a systematic review and meta-analysis.* European urology focus, vol. 5, no. 6, pages 958–969, 2019. (Cité en page 74.)

[Correa 2021] Andres F Correa, Opeyemi A Jegede, Naomi B Haas, Keith T Flaherty, Michael R Pins, Adebowale Adeniran, Edward M Messing, Judith Manola, Christopher G Wood, Christopher J Kane*et al. Predicting disease recurrence, early progression, and overall survival following surgical resection for high-risk localized and locally advanced renal cell carcinoma.* European urology, vol. 80, no. 1, pages 20–31, 2021. (Cité en page 77.)

[Couapel 2014] J-P Couapel, K Bensalah, J-C Bernhard, G Pignot, L Zini, H Lang, J Rigaud, L Salomon, L Bellec, M Souliét *al. Is there a volume–outcome relationship for partial nephrectomy ?* World journal of urology, vol. 32, pages 1323–1329, 2014. (Cité en page 82.)

[Couture 2021] Félix Couture, Tony Finelli, Rodney H Breau, Ranjeeta Mallick, Bimal Bhindi, Simon Tanguay, Anil Kapoor, Ricardo A Rendon, Frédéric Pouliot, Luke Lavalléet *al. The increasing use of renal tumor biopsy amongst Canadian urologists : When is biopsy most utilized ?* In Urologic Oncology : Seminars and Original Investigations, volume 39, pages 499–e15. Elsevier, 2021. (Cité en page 72.)

[Crispen 2009] Paul L Crispen, Rosalia Viterbo, Stephen A Boorjian, Richard E Greenberg, David YT Chen et Robert G Uzzo. *Natural history, growth kinetics, and outcomes of untreated clinically localized renal tumors under active surveillance.* Cancer, vol. 115, no. 13, pages 2844–2852, 2009. (Cité en page 75.)

[Cui 2023] Zejian Cui, João Cartucho, Stamatia Giannarou et Ferdinando Rodriguez y Baena. *Caveats on the first-generation da Vinci research kit : Latent technical constraints and essential calibrations.* IEEE Robotics & Automation Magazine, 2023. (Cité en page 106.)

[Cui 2024] Beilei Cui, Mobarakol Islam, Long Bai, An Wang et Hongliang Ren. *Endodac : Efficient adapting foundation model for self-supervised depth estimation from any endoscopic camera.* In International Conference on Medi-

cal Image Computing and Computer-Assisted Intervention, pages 208–218. Springer, 2024. (Cité en pages 109 et 171.)

[Cunningham 2000] Ian A Cunningham et Philip F Judy. *Computed tomography.* The biomedical engineering handbook, vol. 1, pages 62–61, 2000. (Cité en page 40.)

[Daher 2023] Rema Daher, Francisco Vasconcelos et Danail Stoyanov. *A temporal learning approach to inpainting endoscopic specularities and its effect on image correspondence.* Medical Image Analysis, vol. 90, page 102994, 2023. (Cité en page 132.)

[Dahmen 2014] Jan Dahmen et Walter Vogel. *Optics tube for an endoscope, endoscope, and method for mounting at least one rod lens in an optics tube*, 2014. (Cité en page 21.)

[Dai 2006a] Congxia Dai, Yunfei Zheng et Xin Li. *Accurate video alignment using phase correlation.* IEEE Signal Processing Letters, vol. 13, no. 12, pages 737–740, 2006. (Cité en page 128.)

[Dai 2006b] Congxia Dai, Yunfei Zheng et Xin Li. *Subframe video synchronization via 3d phase correlation.* In 2006 International Conference on Image Processing, pages 501–504. IEEE, 2006. (Cité en page 128.)

[Dai 2018] Chenchen Dai, Yingli Cao, Yan Jia, Yuqin Ding, Ruofan Sheng, Mengsu Zeng et Jianjun Zhou. *Differentiation of renal cell carcinoma subtypes with different iodine quantification methods using single-phase contrast-enhanced dual-energy CT : areal vs. volumetric analyses.* Abdominal Radiology, vol. 43, pages 672–678, 2018. (Cité en page 71.)

[Daniels 2023] Anthony B Daniels, Amit Tirosh, Kristin Huntoon, Gautam U Mehta, Philippe E Spiess, Debra L Friedman, Steven G Waguespack, Jill E Kilkelly, Surya Rednam, Sumit Pruthi*et al. Guidelines for surveillance of patients with von Hippel-Lindau disease : Consensus statement of the International VHL Surveillance Guidelines Consortium and VHL Alliance.* Cancer (0008543X), vol. 129, no. 19, 2023. (Cité en page 70.)

[Dappa 2016] Evelyn Dappa, Kai Higashigaito, Jürgen Fornaro, Sebastian Leschka, Simon Wildermuth et Hatem Alkadhi. *Cinematic rendering–an alternative to volume rendering for 3D computed tomography imaging.* Insights into imaging, vol. 7, pages 849–856, 2016. (Cité en page 44.)

[Darcet 2023] Timothée Darcet, Maxime Oquab, Julien Mairal et Piotr Bojanowski. *Vision transformers need registers.* arXiv preprint arXiv :2309.16588, 2023. (Cité en page 225.)

[Datta 2008] Ritendra Datta, Dhiraj Joshi, Jia Li et James Z Wang. *Image retrieval : Ideas, influences, and trends of the new age.* ACM Computing Surveys (Csur), vol. 40, no. 2, pages 1–60, 2008. (Cité en page 216.)

[De Backer 2023] Pieter De Backer, Maria Peraire Lores, Meret Demuynck, Federico Piramide, Jente Simoens, Tim Oosterlinck, Wouter Bogaert, Chi Victor

Shan, Karel Van Regemorter, Aube Wastyn*et al.* *Surgical phase duration in robot-assisted partial nephrectomy : a surgical data science exploration for clinical relevance.* Diagnostics, vol. 13, no. 21, page 3386, 2023. (Cité en page 84.)

[De Buck 2007] Stijn De Buck, Frederik Maes, André D'Hoore et Paul Suetens. *Evaluation of a novel calibration technique for optically tracked oblique laparoscopes.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 467–474. Springer, 2007. (Cité en page 110.)

[De Hauteclocque 2020] Astrid De Hauteclocque, Clément Michiels, Joffrey Sarrazin, Matthieu Faessel, Luca Mosillo, Mélanie Percot, Peggy Blanc, Henri Bensadoun, Franck Bladou, Grégoire Robert*et al.* *PD30-09 AN ACCURATE TOOL TO INFORM PATIENTS BEFORE RAPN : 3D-PRINTED KIDNEY MODEL.* Journal of Urology, vol. 203, page e628, 2020. (Cité en page 46.)

[De Leon 2019] Alberto Diaz De Leon, Payal Kapur et Ivan Pedrosa. *Radiomics in kidney cancer : MR imaging.* Magnetic Resonance Imaging Clinics, vol. 27, no. 1, pages 1–13, 2019. (Cité en page 71.)

[Defossez 2019] Gautier Defossez, Sandra Le Guyader-Peyrou, Zoé Uhry, Pascale Grosclaude, Marc Colonna, Emmanuelle Dantony, Patricia Delafosse, Florence Molinié, Anne-Sophie Woronoff, Anne-Marie Bouvier*et al.* *Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018.* Étude à partir des registres des cancers du réseau Francim, vol. 1, 2019. (Cité en pages 68 et 69.)

[Delahunt 2013] Brett Delahunt, John C Cheville, Guido Martignoni, Peter A Humphrey, Cristina Magi-Galluzzi, Jesse McKenney, Lars Egevad, Ferran Algaba, Holger Moch, David J Grignon*et al.* *The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters.* The American journal of surgical pathology, vol. 37, no. 10, pages 1490–1504, 2013. (Cité en page 73.)

[Della Corte 2024] Marcello Della Corte, Alberto Quarà, Sabrina De Cillis, Gabriele Volpi, Daniele Amparore, Federico Piramide, Alberto Piana, Michele Sica, Michele Di Dio, Stefano Alba, Francesco Porpiglia, Enrico Checcucci et Cristian Fiori. *3D virtual models and augmented reality for radical prostatectomy : a narrative review.* Chinese Clinical Oncology, vol. 13, no. 4, pages 56–56, 2024. (Cité en pages 171 et 176.)

[DelPozo 2007] Andrey DelPozo et Silvio Savarese. *Detecting specular surfaces on natural images.* In 2007 IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8. IEEE, 2007. (Cité en page 132.)

[Deng 2009] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li et Li Fei-Fei. *Imagenet : A large-scale hierarchical image database.* In 2009 IEEE confe-

rence on computer vision and pattern recognition, pages 248–255. Ieee, 2009. (Cité en pages 146, 202, 240 et 262.)

[Deng 2023] Zhengyi Deng, Morteza Hajihosseini, Justin X Moore, Saira Khan, Rebecca E Graff, Melissa L Bondy, Benjamin I Chung et Marvin E Langston. *Lifetime body weight trajectories and risk of renal cell cancer : a large US prospective cohort study.* Cancer Epidemiology, Biomarkers & Prevention, vol. 32, no. 11, pages 1651–1659, 2023. (Cité en page 69.)

[Deniau 2020] Nicolas Deniau, Sophan Soing-Altrach, Sylvie Maugat et Anne Berger-Carbonne. *Décès liés aux infections nosocomiales : Bilan 2008-2017 des signalements externes en France – Focus sur les bactériémies à Staphylococcus aureus.* Bulletin Épidémiologique Hebdomadaire (BEH), no. 15, pages 305–313, 2020. (Cité en page 12.)

[Desai 2005] Mihir M Desai, Brenda Strzempkowski, Surena F Matin, Andrew P Steinberg, Christopher Ng, Anoop M Meraney, Jihad H Kaouk et Inderbir S Gill. *Prospective randomized comparison of transperitoneal versus retroperitoneal laparoscopic radical nephrectomy.* The Journal of urology, vol. 173, no. 1, pages 38–41, 2005. (Cité en page 74.)

[Desai 2011] Pankaj Desai et Purvi Patel. *Fibroids, infertility and laparoscopic myomectomy.* Journal of gynecological endoscopy and surgery, vol. 2, no. 1, page 36, 2011. (Cité en page 169.)

[Detmer 2017] Felicitas J Detmer, Julian Hettig, Daniel Schindele, Martin Schostak et Christian Hansen. *Virtual and augmented reality systems for renal interventions : A systematic review.* IEEE reviews in biomedical engineering, vol. 10, pages 78–94, 2017. (Cité en page 94.)

[DeTone 2018] Daniel DeTone, Tomasz Malisiewicz et Andrew Rabinovich. *Superpoint : Self-supervised interest point detection and description.* In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pages 224–236, 2018. (Cité en pages 146, 217 et 220.)

[Dexter 2009] Emilie Dexter, Patrick Pérez, Ivan Laptev et Imran N Junejo. *View-independent video synchronization from temporal self-similarities.* In International Conference on Computer Vision Theory and Applications, volume 1, pages 383–391. SCITEPRESS, 2009. (Cité en page 128.)

[Deyanov 2020] Rumen Deyanov. *Endoscopic instrument with compliant thermal interface*, 2020. (Cité en page 37.)

[Dhaya 2020] R Dhaya. *CCTV surveillance for unprecedented violence and traffic monitoring.* Journal of Innovative Image Processing (JIIP), vol. 2, no. 01, pages 25–34, 2020. (Cité en page 129.)

[Diana 2023] Pietro Diana, Tobias Klatte, Daniele Amparore, Riccardo Bertolo, Umberto Carbonara, Selcuk Erdem, Alexandre Ingels, Onder Kara, Laura Marandino, Michele Marchioni*et al. Screening programs for renal cell carcinoma : a systematic review by the EAU young academic urologists renal*

cancer working group. World journal of urology, vol. 41, no. 4, pages 929–940, 2023. (Cité en page 70.)

[Dicarlo 2015] Jeffrey Dicarlo, Stephen J. Blumenkranz, Brian D. Hoffman, Geoff Richmond et Tao Zhao. *Surgical system including a non-white light general illuminator*, 2015. (Cité en pages 36 et 37.)

[Diego 2013] Ferran Diego, Joan Serrat et Antonio M López. *Joint spatio-temporal alignment of sequences*. IEEE Transactions on Multimedia, vol. 15, no. 6, pages 1377–1387, 2013. (Cité en page 128.)

[Ding 2016] Xu Ding et Pin Tao. *Synchronization detection of multicamera system based on LED matrix*. In 2016 13th International Conference on Embedded Software and Systems (ICESS), pages 18–23. IEEE, 2016. (Cité en page 127.)

[Doddamani 2018] Rajalaxmi Doddamani, Srikantaiah Chandrasekharaiah Hiremath, Zameer Ahmed et Lahari Surapaneni. *Complications of laparoscopic surgery in general surgical practice and their management*. International Surgery Journal, vol. 5, no. 4, pages 1233–1239, 2018. (Cité en page 25.)

[Doersch 2022] Carl Doersch, Ankush Gupta, Larisa Markeeva, Adria Recasens, Lucas Smaira, Yusuf Aytar, Joao Carreira, Andrew Zisserman et Yi Yang. *Tap-vid : A benchmark for tracking any point in a video*. Advances in Neural Information Processing Systems, vol. 35, pages 13610–13626, 2022. (Cité en page 224.)

[Doersch 2023] Carl Doersch, Yi Yang, Mel Vecerik, Dilara Gokay, Ankush Gupta, Yusuf Aytar, Joao Carreira et Andrew Zisserman. *Tapir : Tracking any point with per-frame initialization and temporal refinement*. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 10061–10072, 2023. (Cité en page 225.)

[Dong 2009] Zilong Dong, Guofeng Zhang, Jiaya Jia et Hujun Bao. *Keyframe-based real-time camera tracking*. In 2009 IEEE 12th international conference on computer vision, pages 1538–1545. IEEE, 2009. (Cité en page 247.)

[Dong 2014] Zilong Dong, Guofeng Zhang, Jiaya Jia et Hujun Bao. *Efficient keyframe-based real-time camera tracking*. Computer Vision and Image Understanding, vol. 118, pages 97–110, 2014. (Cité en page 247.)

[Dong 2018] Wen Dong, Jitao Wu, Chalairat Suk-Ouichai, Elvis Caraballo Antonio, Erick M Remer, Jianbo Li, Joseph Zabell, Sudhir Isharwal et Steven C Campbell. *Ischemia and functional recovery from partial nephrectomy : refined perspectives*. European Urology Focus, vol. 4, no. 4, pages 572–578, 2018. (Cité en page 81.)

[Dorsey 2006] Julie Dorsey et Pat Hanrahany. *Modeling and rendering of metallic patinas*. In ACM SIGGRAPH 2006 Courses, pages 2–es. 2006. (Cité en page 123.)

[Dowrick 2023] Thomas Dowrick, Guofang Xiao, Daniil Nikitichev, Eren Dursun, Niels van Berkel, Moustafa Allam, Bongjin Koo, Joao Ramalhinho, Stephen

Thompson, Kurinchi Gurusamy *et al. Evaluation of a calibration rig for stereo laparoscopes.* Medical Physics, vol. 50, no. 5, pages 2695–2704, 2023. (Cité en page 110.)

[Dreyer 2009] Francis Dreyer. *Histoire de l'hôpital. L'éclairage des salles d'opération aux XIXe et XXe siècles : l'apparition du scialytique.* In Situ. Revue des patrimoines, no. 10, 2009. (Cité en page 11.)

[Driessen 2017] Sara RC Driessen, Evelien M Sandberg, Sharon P Rodrigues, Erik W van Zwet et Frank Willem Jansen. *Identification of risk factors in minimally invasive surgery : a prospective multicenter study.* Surgical endoscopy, vol. 31, pages 2467–2473, 2017. (Cité en page 25.)

[Du 2015] Xiaofei Du, Neil Clancy, Shobhit Arya, George B Hanna, John Kelly, Daniel S Elson et Danail Stoyanov. *Robust surface tracking combining features, intensity and illumination compensation.* International journal of computer assisted radiology and surgery, vol. 10, no. 12, pages 1915–1926, 2015. (Cité en page 221.)

[Du 2022] Wurong Du, Kaibo Guo, Huimin Jin, Leitao Sun, Shanming Ruan et Qiaoling Song. *Association between metabolic syndrome and risk of renal cell cancer : A meta-analysis.* Frontiers in Oncology, vol. 12, page 928619, 2022. (Cité en page 69.)

[Duisterhof 2024] Bardienus Duisterhof, Lojze Zust, Philippe Weinzaepfel, Vincent Leroy, Yohann Cabon et Jerome Revaud. *MASt3R-SfM : a Fully-Integrated Solution for Unconstrained Structure-from-Motion.* arXiv preprint arXiv :2409.19152, 2024. (Cité en page 190.)

[Dunnhofer 2022] Matteo Dunnhofer et Christian Micheloni. *Cocolot : Combining complementary trackers in long-term visual tracking.* In 2022 26th International Conference on Pattern Recognition (ICPR), pages 5132–5139. IEEE, 2022. (Cité en page 229.)

[Dusmanu 2019] Mihai Dusmanu, Ignacio Rocco, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii et Torsten Sattler. *D2-net : A trainable cnn for joint description and detection of local features.* In Proceedings of the ieee/cvf conference on computer vision and pattern recognition, pages 8092–8101, 2019. (Cité en pages 146 et 217.)

[Dwivedi 2024] Sai Kumar Dwivedi, Yu Sun, Priyanka Patel, Yao Feng et Michael J Black. *Tokenhmr : Advancing human mesh recovery with a tokenized pose representation.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1323–1333, 2024. (Cité en page 208.)

[Ebbing 2015] J Ebbing, T Wiebach, C Kempkensteffen, K Miller, A Bachmann, K Günzel, M Rieken, F Friedersdorff, D Baumunk et S Weikert. *Evaluation of perioperative complications in open and laparoscopic surgery for renal cell cancer with tumor thrombus involvement using the Clavien–Dindo classification.* European Journal of Surgical Oncology (EJSO), vol. 41, no. 7, pages 941–952, 2015. (Cité en pages 74 et 76.)

[Edgcumbe 2018] Philip Edgcumbe, Rohit Singla, Philip Pratt, Caitlin Schneider, Christopher Nguan et Robert Rohling. *Follow the light : projector-based augmented reality intracorporeal system for laparoscopic surgery.* Journal of Medical Imaging, vol. 5, no. 2, pages 021216–021216, 2018. (Cité en page 96.)

[Edstedt 2024a] Johan Edstedt, Georg Bökman, Mårten Wadenbäck et Michael Felsberg. *DeDoDe : Detect, don't describe—Describe, don't detect for local feature matching.* In 2024 International Conference on 3D Vision (3DV), pages 148–157. IEEE, 2024. (Cité en page 146.)

[Edstedt 2024b] Johan Edstedt, Qiyu Sun, Georg Bökman, Mårten Wadenbäck et Michael Felsberg. *RoMa : Robust dense feature matching.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 19790–19800, 2024. (Cité en page 146.)

[Edwards 2022] PJ Eddie Edwards, Dimitris Psychogyios, Stefanie Speidel, Lena Maier-Hein et Danail Stoyanov. *SERV-CT : A disparity dataset from cone-beam CT for validation of endoscopic 3D reconstruction.* Medical image analysis, vol. 76, page 102302, 2022. (Cité en page 226.)

[Eichler 2022] Nadav Eichler, Hagit Hel-Or et Ilan Shimshoni. *Spatio-temporal calibration of multiple kinect cameras using 3d human pose.* Sensors, vol. 22, no. 22, page 8900, 2022. (Cité en page 128.)

[Eid 2017] Marwen Eid, Carlo N De Cecco, John W Nance Jr, Damiano Caruso, Moritz H Albrecht, Adam J Spandorfer, Domenico De Santis, Akos Varga-Szemes et U Joseph Schoepf. *Cinematic rendering in CT : a novel, lifelike 3D visualization technique.* American Journal of Roentgenology, vol. 209, no. 2, pages 370–379, 2017. (Cité en page 44.)

[El Meslouhi 2011] Othmane El Meslouhi, Mustapha Kardouchi, Hakim Allali, Taoufiq Gadi et Yassir Ait Benkaddour. *Automatic detection and inpainting of specular reflections for colposcopic images.* Central European Journal of Computer Science, vol. 1, pages 341–354, 2011. (Cité en page 132.)

[Engel 2017] Jakob Engel, Vladlen Koltun et Daniel Cremers. *Direct sparse odometry.* IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 3, pages 611–625, 2017. (Cité en page 223.)

[Espinel 2020] Yamid Espinel, Erol Özgür, Lilian Calvet, Bertrand Le Roy, Emmanuel Buc et Adrien Bartoli. *Combining visual cues with interactions for 3D–2D registration in liver laparoscopy.* Annals of Biomedical Engineering, vol. 48, pages 1712–1727, 2020. (Cité en page 172.)

[Espinel 2021] Yamid Espinel, Lilian Calvet, Karim Botros, Emmanuel Buc, Christophe Tilmant et Adrien Bartoli. *Using multiple images and contours for deformable 3d-2d registration of a preoperative ct in laparoscopic liver surgery.* In Medical Image Computing and Computer Assisted Intervention–MICCAI 2021 : 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part IV 24, pages 657–666. Springer, 2021. (Cité en page 172.)

[Espinel 2024a] Yamid Espinel, Nalick Lombion, Luce Compagnone, Nicolas Saroul et Adrien Bartoli. *Preliminary trials of trackerless augmented reality in endoscopic endonasal surgery.* International Journal of Computer Assisted Radiology and Surgery, pages 1–5, 2024. (Cité en pages 51 et 52.)

[Espinel 2024b] Yamid Espinel, Navid Rabbani, Thien Bao Bui, Mathieu Ribeiro, Emmanuel Buc et Adrien Bartoli. *Keyhole-aware laparoscopic augmented reality.* Medical Image Analysis, vol. 94, page 103161, 2024. (Cité en pages 51 et 52.)

[Evangelidis 2008] Georgios D Evangelidis et Emmanouil Z Psarakis. *Parametric image alignment using enhanced correlation coefficient maximization.* IEEE transactions on pattern analysis and machine intelligence, vol. 30, no. 10, pages 1858–1865, 2008. (Cité en page 128.)

[Evangelidis 2011] Georgios D Evangelidis et Christian Bauckhage. *Efficient and robust alignment of unsynchronized video sequences.* In Joint Pattern Recognition Symposium, pages 286–295. Springer, 2011. (Cité en page 128.)

[Fan 2024] Junyu Fan et Shunping Ji. *Adaptive and anti-drift motion constraints for object tracking in satellite videos.* Remote Sensing, vol. 16, page 1347, 2024. (Cité en page 230.)

[Farinha 2023] Rui Farinha, Alberto Breda, James Porter, Alexandre Mottrie, Ben Van Cleynenbreugel, Jozef Vander Sloten, Angelo Mottaran, Anthony G Gallagher*et al. International expert consensus on metric-based characterization of robot-assisted partial nephrectomy.* European Urology Focus, vol. 9, no. 2, pages 388–395, 2023. (Cité en page 84.)

[Fedorov 2012] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka*et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network.* Magnetic resonance imaging, vol. 30, no. 9, pages 1323–1341, 2012. (Cité en page 44.)

[Ferguson 2020] James M Ferguson, E Bryn Pitt, Andria A Remirez, Michael A Siebold, Alan Kuntz, Nicholas L Kavoussi, Eric J Barth, S Duke Herrell et Robert J Webster. *Toward practical and accurate touch-based image guidance for robotic partial nephrectomy.* IEEE transactions on medical robotics and bionics, vol. 2, no. 2, pages 196–205, 2020. (Cité en page 96.)

[Feris 2006] Rogerio Feris, Ramesh Raskar, Kar-Han Tan et Matthew Turk. *Specular highlights detection and reduction with multi-flash photography.* Journal of the Brazilian Computer Society, vol. 12, pages 35–42, 2006. (Cité en page 132.)

[Ferlay 2024] Jacques Ferlay, M Ervik, F Lam, M Laversanne, M Colombet, L Mery, M Piñeros, A Znaor, I Soerjomataram et F Bray. *Global cancer observatory : cancer today (version 1.1).* Lyon, France : International Agency for Research on Cancer, 2024. (Cité en page 68.)

[Ferrari 2021] Matteo Ferrari, Roberto Cartolari, Jessica Barizzi, Ricardo Pereira Mestre, Eugenia D'Antonio et Julien Renard. *Percutaneous biopsy of small*

renal mass : can diagnostic accuracy be affected by hospital volume ? Central European Journal of Urology, vol. 74, no. 3, page 334, 2021. (Cité en page 72.)

[Ficarra 2009] Vincenzo Ficarra, Giacomo Novara, Silvia Secco, Veronica Macchi, Andrea Porzionato, Raffaele De Caro et Walter Artibani. *Preoperative aspects and dimensions used for an anatomical (PADUA) classification of renal tumours in patients who are candidates for nephron-sparing surgery.* European urology, vol. 56, no. 5, pages 786–793, 2009. (Cité en page 72.)

[Finelli 2020] Antonio Finelli, Douglas C Cheung, Ashraf Al-Matar, Andrew J Evans, Christopher G Morash, Stephen E Pautler, D Robert Siemens, Simon Tanguay, Ricardo A Rendon, Martin E Gleave*et al. Small renal mass surveillance : histology-specific growth rates in a biopsy-characterized cohort.* European urology, vol. 78, no. 3, pages 460–467, 2020. (Cité en page 75.)

[Fischler 1981] Martin A Fischler et Robert C Bolles. *Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography.* Communications of the ACM, vol. 24, no. 6, pages 381–395, 1981. (Cité en pages 97, 168 et 218.)

[Flores-Mangas 2012] Fernando Flores-Mangas, Allan D Jepson et Masoom A Haider. *Shape-Based Registration of Kidneys Across Differently Contrasted CT Scans.* In 2012 Ninth Conference on Computer and Robot Vision, pages 244–251. IEEE, 2012. (Cité en page 44.)

[Fong 2024] Khi Yung Fong, Valerie Huei Li Gan, Benjamin Jia Han Lim, Yiong Huak Chan, Daniele Castellani, Kenneth Chen, Kae Jack Tay, Henry Sun Sien Ho, John Shyi Peng Yuen, Edwin Aslim*et al. Off-clamp vs on-clamp robot-assisted partial nephrectomy : a systematic review and meta-analysis.* BJU international, vol. 133, no. 4, pages 375–386, 2024. (Cité en page 81.)

[Forde 2020] Claire Forde, Derek HK Lim, Yousef Alwan, George Burghel, Laura Butland, Ruth Cleaver, Abhijit Dixit, D Gareth Evans, Helen Hanson, Fiona Lalloo*et al. Hereditary leiomyomatosis and renal cell cancer : clinical, molecular, and screening features in a cohort of 185 affected individuals.* European urology oncology, vol. 3, no. 6, pages 764–772, 2020. (Cité en page 70.)

[Forsyth 2002] David A Forsyth et Jean Ponce. Computer vision : a modern approach. prentice hall professional technical reference, 2002. (Cité en page 215.)

[Fragoso 2013] Victor Fragoso, Pradeep Sen, Sergio Rodriguez et Matthew Turk. *EVSAC : accelerating hypotheses generation by modeling matching scores with extreme value theory.* In Proceedings of the IEEE international conference on computer vision, pages 2472–2479, 2013. (Cité en page 218.)

[François 2020] Tom François, Lilian Calvet, Sabrina Madad Zadeh, Damien Saboul, Simone Gasparini, Prasad Samarakoon, Nicolas Bourdel et Adrien Bartoli. *Detecting the occluding contours of the uterus to automatise augmented laparoscopy : score, loss, dataset, evaluation and user study.* International journal

of computer assisted radiology and surgery, vol. 15, pages 1177–1186, 2020. (Cité en pages 172, 244 et 245.)

[Freschi 2013] Cinzia Freschi, Vincenzo Ferrari, Franca Melfi, Mauro Ferrari, Franco Mosca et Alfred Cuschieri. *Technical review of the da Vinci surgical telemanipulator*. The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 9, no. 4, pages 396–406, 2013. (Cité en pages 31, 32, 33 et 34.)

[Froner 2016] Diego S Froner, José LS Pio et Walter CSS Simões. *Spatiotemporal alignment of sequential images : Utilization of GPS data as reference to alignment of multiple video cameras*. In 2016 XLII Latin American Computing Conference (CLEI), pages 1–7. IEEE, 2016. (Cité en page 127.)

[Fu 2020] Gang Fu, Qing Zhang, Qifeng Lin, Lei Zhu et Chunxia Xiao. *Learning to detect specular highlights from real-world images*. In Proceedings of the 28th ACM International Conference on Multimedia, pages 1873–1881, 2020. (Cité en page 131.)

[Fuentes-Jimenez 2022] David Fuentes-Jimenez, Daniel Pizarro, David Casillas-Pérez, Toby Collins et Adrien Bartoli. *Deep shape-from-template : Single-image quasi-isometric deformable registration and reconstruction*. Image and Vision Computing, vol. 127, page 104531, 2022. (Cité en page 201.)

[Gálvez-López 2012] Dorian Gálvez-López et Juan D Tardos. *Bags of binary words for fast place recognition in image sequences*. IEEE Transactions on robotics, vol. 28, no. 5, pages 1188–1197, 2012. (Cité en page 222.)

[Gandjbakhch 2009] Iradj Gandjbakhch. *Rapport Séance du 28 avril 2009 : Bloc opératoire : de la salle d'opération à la plate-forme interventionnelle*, 2009. (Cité en pages 12 et 15.)

[Gao 2017] Yefei Gao, Jian Yang, Shaodong Ma, Danni Ai, Tong Lin, Songyuan Tang et Yongtian Wang. *Dynamic searching and classification for highlight removal on endoscopic image*. Procedia Computer Science, vol. 107, pages 762–767, 2017. (Cité en page 132.)

[Garbens 2020] Alaina Garbens, Christopher JD Wallis, Zachary Klaassen, Refik Saskin, Lesley Plumptre, Ronald Kodama, Sender Herschorn et Robert K Nam. *Comprehensive assessment of the morbidity of renal mass biopsy : A population-based assessment of biopsy-related complications*. Canadian Urological Association Journal, vol. 15, no. 2, page 42, 2020. (Cité en page 72.)

[Garg 2022] Harshit Garg, Sarah P Psutka, Abraham Ari Hakimi, Hyung L Kim, Ahmed M Mansour, Deepak K Pruthi, Michael A Liss, Hanzhang Wang, Christine S Gaspard, Chethan Ramamurthy*et al. A decade of robotic-assisted radical nephrectomy with inferior vena cava thrombectomy : a systematic review and meta-analysis of perioperative outcomes*. The Journal of urology, vol. 208, no. 3, pages 542–560, 2022. (Cité en page 79.)

[Garisto 2018] Juan Garisto, Riccardo Bertolo, Julien Dagenais, Daniel Sagalovich, Khaled Fareed, Amr Fergany, Robert Stein et Jihad Kaouk. *Robotic versus*

open partial nephrectomy for highly complex renal masses : Comparison of perioperative, functional, and oncological outcomes. In Urologic Oncology : Seminars and Original Investigations, volume 36, pages 471–e1. Elsevier, 2018. (Cité en page 76.)

[Garud 2012] SS Garud et FF Willingham. *Natural orifice transluminal endoscopic surgery.* Endoscopy, vol. 44, no. 09, pages 865–868, 2012. (Cité en page 22.)

[Gaspar 2016] Tiago Gaspar, Paulo Oliveira et Paolo Favaro. *Synchronization of independently moving cameras via motion recovery.* SIAM Journal on Imaging Sciences, vol. 9, no. 3, pages 869–900, 2016. (Cité en page 129.)

[Gawaikar 2023] Ankita Gawaikar, Swati Sorte, Payal Ghutke, Wani Patil et Sonali Joshi. *Review of Da Vinci Surgical System and Haptic Feedback Device for Telesurgery.* In 2023 8th International Conference on Communication and Electronics Systems (ICCES), pages 1596–1600. IEEE, 2023. (Cité en page 37.)

[Geist 2024] A René Geist, Jonas Frey, Mikel Zobro, Anna Levina et Georg Martius. *Learning with 3D rotations, a hitchhiker's guide to SO (3).* arXiv preprint arXiv :2404.11735, 2024. (Cité en page 202.)

[Gennisson 2012] Jean-Luc Gennisson, Nicolas Grenier, Christian Combe et Mickaël Tanter. *Supersonic shear wave elastography of in vivo pig kidney : influence of blood pressure, urinary pressure and tissue anisotropy.* Ultrasound in medicine & biology, vol. 38, no. 9, pages 1559–1567, 2012. (Cité en page 58.)

[Gershman 2017] Boris Gershman, R Houston Thompson, Daniel M Moreira, Stephen A Boorjian, Matthew K Tollefson, Christine M Lohse, Brian A Costello, John C Cheville et Bradley C Leibovich. *Radical nephrectomy with or without lymph node dissection for nonmetastatic renal cell carcinoma : a propensity score-based analysis.* European urology, vol. 71, no. 4, pages 560–567, 2017. (Cité en page 77.)

[Gershman 2018] Boris Gershman, R Houston Thompson, Stephen A Boorjian, Alessandro Larcher, Umberto Capitanio, Francesco Montorsi, Cristina Carenzi, Roberto Bertini, Alberto Briganti, Christine M Lohse*et al. Radical nephrectomy with or without lymph node dissection for high risk nonmetastatic renal cell carcinoma : a multi-institutional analysis.* The Journal of Urology, vol. 199, no. 5, pages 1143–1148, 2018. (Cité en page 76.)

[Gholami 2022] Amir Gholami, Sehoon Kim, Zhen Dong, Zhewei Yao, Michael W Mahoney et Kurt Keutzer. *A survey of quantization methods for efficient neural network inference.* In Low-power computer vision, pages 291–326. Chapman and Hall/CRC, 2022. (Cité en page 243.)

[Giannarou 2012] Stamatia Giannarou, Marco Visentini-Scarzanella et Guang-Zhong Yang. *Probabilistic tracking of affine-invariant anisotropic regions.* IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pages 130–143, 2012. (Cité en pages 221 et 226.)

[Giannone 2021] Fabio Giannone, Emanuele Felli, Zineb Cherkaoui, Pietro Mascagni et Patrick Pessaux. *Augmented reality and image-guided robotic liver surgery.* Cancers, vol. 13, no. 24, page 6268, 2021. (Cité en page 51.)

[Gibson 2018] Eli Gibson, Francesco Giganti, Yipeng Hu, Ester Bonmati, Steve Bandula, Kurinchi Gurusamy, Brian Davidson, Stephen P Pereira, Matthew J Clarkson et Dean C Barratt. *Automatic multi-organ segmentation on abdominal CT with dense V-networks.* IEEE transactions on medical imaging, vol. 37, no. 8, pages 1822–1834, 2018. (Cité en page 45.)

[Gill 2007] Inderbir S Gill, Louis R Kavoussi, Brian R Lane, Michael L Blute, Denise Babineau, J Roberto Colombo, Igor Frank, Sompol Permpongkosol, Christopher J Weight, Jihad H Kaouk*et al. Comparison of 1,800 laparoscopic and open partial nephrectomies for single renal tumors.* The Journal of urology, vol. 178, no. 1, pages 41–46, 2007. (Cité en page 81.)

[Gill 2015] Inderbir S Gill, Charles Metcalfe, Andre Abreu, Vinay Duddalwar, Sameer Chopra, Mark Cunningham, Duraiyah Thangathurai, Osamu Ukimura, Raj Satkunasivam, Andrew Hung*et al. Robotic level III inferior vena cava tumor thrombectomy : initial series.* The Journal of urology, vol. 194, no. 4, pages 929–938, 2015. (Cité en page 79.)

[Giorgio 2018] Rossi Giorgio, Tarasconi Antonio, Baiocchi Gianluca, Luigi de'Angelis Gian, Gaiani Federica, Catena Fausto*et al. Fluorescence guided surgery in liver tumors : applications and advantages.* Acta Bio Medica : Atenei Parmensis, vol. 89, no. Suppl 9, page 135, 2018. (Cité en page 48.)

[Gleize 2023] Pierre Gleize, Weiyao Wang et Matt Feiszli. *Silk : Simple learned keypoints.* In Proceedings of the IEEE/CVF international conference on computer vision, pages 22499–22508, 2023. (Cité en page 145.)

[Goel 2023] Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa et Jitendra Malik. *Humans in 4D : Reconstructing and tracking humans with transformers.* In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 14783–14794, 2023. (Cité en page 208.)

[Goldberg 2020] Hanan Goldberg, Rami Ajaj, Jaime Omar Herrera Cáceres, Alejandro Berlin, Thenappan Chandrasekar, Zachary Klaassen, Christopher JD Wallis, Ardalan E Ahmad, Ricardo Leao, Anika R Petrella*et al. Psychological distress associated with active surveillance in patients younger than 70 with a small renal mass.* In Urologic Oncology : Seminars and Original Investigations, volume 38, pages 603–e17. Elsevier, 2020. (Cité en page 75.)

[Goldmann 2010] Lutz Goldmann, Jong-Seok Lee et Touradj Ebrahimi. *Temporal synchronization in stereoscopic video : Influence on quality of experience and automatic asynchrony detection.* In 2010 IEEE International Conference on Image Processing, pages 3241–3244. IEEE, 2010. (Cité en pages 129 et 130.)

[Gomel 1974] Victor Gomel. *Laparoscopy.* Canadian Medical Association Journal, vol. 111, no. 2, page 167, 1974. (Cité en page 23.)

[Gómez-Rodríguez 2021] Juan J Gómez-Rodríguez, José Lamarca, Javier Morlana, Juan D Tardós et José MM Montiel. *SD-DefSLAM : Semi-direct monocular SLAM for deformable and intracorporeal scenes.* In 2021 IEEE international conference on robotics and automation (ICRA), pages 5170–5177. IEEE, 2021. (Cité en page 223.)

[Gorin 2022] Michael A Gorin, Hiten D Patel, Steven P Rowe, Noah M Hahn, Hans J Hammers, Alice Pons, Bruce J Trock, Phillip M Pierorazio, Thomas R Nirschl, Daniela C Salles*et al. Neoadjuvant nivolumab in patients with high-risk nonmetastatic renal cell carcinoma.* European urology oncology, vol. 5, no. 1, pages 113–117, 2022. (Cité en page 77.)

[Goujon 2021] A Goujon, G Verhoest, F Sallusto, J Branchereau, JM Boutin, T Bessede, N Terrier, G Karam, L Badet, P Bigot*et al. Renal cell carcinoma in candidates for renal transplantation and recipients of a kidney transplant : The French guidelines from CTAFU.* Progres en Urologie : Journal de L'association Francaise D'urologie et de la Societe Francaise D'urologie, vol. 31, no. 1, pages 18–23, 2021. (Cité en page 70.)

[Graff 2022] Rebecca E Graff, Kathryn M Wilson, Alejandro Sanchez, Steven L Chang, David F McDermott, Toni K Choueiri, Eunyoung Cho, Sabina Signoretti, Edward L Giovannucci et Mark A Preston. *Obesity in relation to renal cell carcinoma incidence and survival in three prospective studies.* European urology, vol. 82, no. 3, pages 247–251, 2022. (Cité en page 69.)

[Gratzke 2009] Christian Gratzke, Michael Seitz, Florian Bayrle, Boris Schlenker, Patrick J Bastian, Niko Haseke, Markus Bader, Derya Tilki, Alexander Roosen, Alexander Karl*et al. Quality of life and perioperative outcomes after retroperitoneoscopic radical nephrectomy (RN), open RN and nephron-sparing surgery in patients with renal cell carcinoma.* BJU international, vol. 104, no. 4, pages 470–475, 2009. (Cité en pages 76 et 79.)

[Greco 2019] Francesco Greco, Riccardo Autorino, Vincenzo Altieri, Steven Campbell, Vincenzo Ficarra, Inderbir Gill, Alexander Kutikov, Alex Mottrie, Vincenzo Mirone et Hendrik van Poppel. *Ischemia techniques in nephron-sparing surgery : a systematic review and meta-analysis of surgical, oncological, and functional outcomes.* European urology, vol. 75, no. 3, pages 477–491, 2019. (Cité en page 81.)

[Grenier 2013] N Grenier, J-L Gennisson, F Cornelis, Y Le Bras et L Couzi. *Renal ultrasound elastography.* Diagnostic and interventional imaging, vol. 94, no. 5, pages 545–550, 2013. (Cité en page 58.)

[Grishchenko 2020] Ivan Grishchenko, Artsiom Ablavatski, Yury Kartynnik, Karthik Raveendran et Matthias Grundmann. *Attention mesh : High-fidelity face mesh prediction in real-time.* arXiv preprint arXiv :2006.10962, 2020. (Cité en page 183.)

[Griwodz 2018] Carsten Griwodz, Lilian Calvet et Pål Halvorsen. *Popsift : A faithful SIFT implementation for real-time applications.* In Proceedings of the

9th ACM Multimedia Systems Conference, pages 415–420, 2018. (Cité en pages 245, 251 et 256.)

[Griwodz 2021] Carsten Griwodz, Simone Gasparini, Lilian Calvet, Pierre Gurdjos, Fabien Castan, Benoit Maujean, Gregoire De Lillo et Yann Lanthony. *AliceVision Meshroom : An open-source 3D reconstruction pipeline*. In Proceedings of the 12th ACM multimedia systems conference, pages 241–247, 2021. (Cité en pages 244 et 270.)

[Grossberg 2005] Michael D Grossberg et Shree K Nayar. *The raxel imaging model and ray-based calibration*. International Journal of Computer Vision, vol. 61, pages 119–137, 2005. (Cité en page 110.)

[Gu 2024] Liangyou Gu, Cheng Peng, Huaikang Li, Tongyu Jia, Xinran Chen, Hanfeng Wang, Songliang Du, Lu Tang, Qiyang Liang, Baojun Wang*et al. Neoadjuvant therapy in renal cell carcinoma with tumor thrombus : a systematic review and meta-analysis*. Critical Reviews in Oncology/Hematology, page 104316, 2024. (Cité en page 77.)

[Guan 2009] Peng Guan, Alexander Weiss, Alexandru O Balan et Michael J Black. *Estimating human shape and pose from a single image*. In 2009 IEEE 12th International Conference on Computer Vision, pages 1381–1388. IEEE, 2009. (Cité en page 178.)

[Gudi 2020] Amogh Gudi, Xin Li et Jan van Gemert. *Efficiency in real-time webcam gaze tracking*. In Computer Vision–ECCV 2020 Workshops : Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16, pages 529–543. Springer, 2020. (Cité en page 183.)

[Guglielmetti 2022] Giuliano B Guglielmetti, Gabriel C Dos Anjos, Guilherme Sawczyn, Gilberto Rodrigues, Leonardo Cardili, Maurício D Cordeiro, Luiz CO Neves, José Pontes Junior, Arnaldo Fazoli, Rafael F Coelho*et al. A prospective, randomized trial comparing the outcomes of open vs laparoscopic partial nephrectomy*. The Journal of Urology, vol. 208, no. 2, pages 259–267, 2022. (Cité en page 81.)

[Guo 2025] Yuliang Guo, Sparsh Garg, S Mahdi H Miangoleh, Xinyu Huang et Liu Ren. *Depth Any Camera : Zero-Shot Metric Depth Estimation from Any Camera*. arXiv preprint arXiv :2501.02464, 2025. (Cité en page 109.)

[Guy 2022] Sylvain Guy, Jean-Loup Haberbusch, Emmanuel Promayon, Stéphane Mancini et Sandrine Voros. *Qualitative comparison of image stitching algorithms for multi-camera systems in laparoscopy*. Journal of imaging, vol. 8, no. 3, page 52, 2022. (Cité en pages 226 et 227.)

[Guzzo 2009] Thomas J Guzzo, Phillip M Pierorazio, Edward M Schaeffer, Elliot K Fishman et Mohamad E Allaf. *The accuracy of multidetector computerized tomography for evaluating tumor thrombus in patients with renal cell carcinoma*. The Journal of urology, vol. 181, no. 2, pages 486–491, 2009. (Cité en page 71.)

[Haas 2014] Naomi B Haas et Katherine L Nathanson. *Hereditary kidney cancer syndromes*. Advances in chronic kidney disease, vol. 21, no. 1, pages 81–90, 2014. (Cité en page 69.)

[Haase 1992] Chet S Haase et Gary W Meyer. *Modeling pigmented materials for realistic image synthesis*. ACM Transactions on Graphics (TOG), vol. 11, no. 4, pages 305–335, 1992. (Cité en page 123.)

[Hageman 2008] Danny Hageman, Victoria Caillet, Jeff Kostohryz et Sue Madick. *Laparoscopic-assisted colon surgery*. AORN journal, vol. 88, no. 3, pages 403–416, 2008. (Cité en page 24.)

[Haidegger 2022] Tamas Haidegger, Stefanie Speidel, Danail Stoyanov et Richard M Satava. *Robot-assisted minimally invasive surgery—Surgical robotics in the data age*. Proceedings of the IEEE, vol. 110, no. 7, pages 835–846, 2022. (Cité en page 29.)

[Halabi 2024] Mouhammad Halabi, Kayanne Khoury, Abdulrahman Alomar, Joseph El Dahdah, Obai Hassan, Khadija Hayyan, Engy Bishara et Hatem Moussa. *Operative efficiency : a comparative analysis of Versius and da Vinci robotic systems in abdominal surgery*. Journal of Robotic Surgery, vol. 18, no. 1, page 132, 2024. (Cité en page 33.)

[Halderman 2023] Jonathan Halderman, John A. Barton, Mathew Clopp, Rumen Deyanov, Anqi FanYves Lacroix, Derek C. Liou et Theodore W. Rogers. *Image capture device with reduced fogging*, 2023. (Cité en page 37.)

[Hallscheidt 2005] Peter J Hallscheidt, Christian Fink, Axel Haferkamp, Michael Bock, Ante Luburic, Ivan Zuna, Gerd Noeldge et Guenter Kauffmann. *Preoperative staging of renal cell carcinoma with inferior vena cava thrombus using multidetector CT and MRI : prospective study with histopathological correlation*. Journal of computer assisted tomography, vol. 29, no. 1, pages 64–68, 2005. (Cité en page 71.)

[Hamarneh 2014] Ghassan Hamarneh, Alborz Amir-Khalili, Masoud S Nosrati, Ivan Figueroa, Jeremy Kawahara, Osama Al-Alao, Jean-Marc Peyrat, Julien Abi-Nahed, Abdulla Al-Ansari et Rafeef Abugharbieh. *Towards multi-modal image-guided tumour identification in robot-assisted partial nephrectomy*. In 2nd Middle East Conference on Biomedical Engineering, pages 159–162. IEEE, 2014. (Cité en page 97.)

[Hamzah 2016] Rostam Affendi Hamzah et Haidi Ibrahim. *Literature survey on stereo vision disparity map algorithms*. Journal of Sensors, vol. 2016, no. 1, page 8742920, 2016. (Cité en page 109.)

[Haouchine 2014] Nazim Haouchine, Stephane Cotin, Igor Peterlik, Jeremie Dequidt, Mario Sanz Lopez, Erwan Kerrien et Marie-Odile Berger. *Impact of soft tissue heterogeneity on augmented reality for liver surgery*. IEEE transactions on visualization and computer graphics, vol. 21, no. 5, pages 584–597, 2014. (Cité en page 172.)

[Haouchine 2016] Nazim Haouchine, Frederick Roy, Lionel Untereiner et Stéphane Cotin. *Using contours as boundary conditions for elastic registration during minimally invasive hepatic surgery.* In 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 495–500. IEEE, 2016. (Cité en page 172.)

[Hardner 2022] M Hardner, R Docea et D Schneider. *Guided calibration of medical stereo endoscopes.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 43, pages 679–686, 2022. (Cité en page 110.)

[Haresh 2021] Sanjay Haresh, Sateesh Kumar, Huseyin Coskun, Shahram N Syed, Andrey Konin, Zeeshan Zia et Quoc-Huy Tran. *Learning by aligning videos in time.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5548–5558, 2021. (Cité en page 129.)

[Harley 2022] Adam W Harley, Zhaoyuan Fang et Katerina Fragkiadaki. *Particle video revisited : Tracking through occlusions using point trajectories.* In European Conference on Computer Vision, pages 59–75. Springer, 2022. (Cité en page 225.)

[Harley 2025] Adam W Harley, Yang You, Xinglong Sun, Yang Zheng, Nikhil Raghuraman, Yunqi Gu, Sheldon Liang, Wen-Hsuan Chu, Achal Dave, Pavel Tokmakov*et al. AllTracker : Efficient Dense Point Tracking at High Resolution.* arXiv preprint arXiv :2506.07310, 2025. (Cité en page 224.)

[Harrison 2023] Hannah Harrison, Angela Wood, Lisa Pennells, Sabrina H Rossi, Matthew Callister, Jon Cartledge, Grant D Stewart et Juliet A Usher-Smith. *Estimating the effectiveness of kidney cancer screening within lung cancer screening programmes : a validation in UK biobank.* European Urology Oncology, vol. 6, no. 3, pages 351–353, 2023. (Cité en page 70.)

[Hart 1988] Sandra G Hart et Lowell E Staveland. *Development of NASA-TLX (Task Load Index) : Results of empirical and theoretical research.* In Advances in psychology, volume 52, pages 139–183. Elsevier, 1988. (Cité en pages 92 et 273.)

[Hartley 2003] Richard Hartley et Andrew Zisserman. Multiple view geometry in computer vision. Cambridge university press, 2003. (Cité en page 104.)

[Hartwig 2022] Regine Hartwig, Daniel Ostler, Jean-Claude Rosenthal, Hubertus Feußner, Dirk Wilhelm et Dirk Wollherr. *Miti : Slam benchmark for laparoscopic surgery.* arXiv preprint arXiv :2202.11496, 2022. (Cité en page 226.)

[Hasegawa 2018] Takaaki Hasegawa, Takashi Yamanaka, Hideo Gobara, Masaya Miyazaki, Haruyuki Takaki, Yozo Sato, Yoshitaka Inaba et Koichiro Yamakado. *Radiofrequency ablation versus cryoablation for T1b renal cell carcinoma : a multi-center study.* Japanese Journal of Radiology, vol. 36, pages 551–558, 2018. (Cité en page 77.)

[Haute Autorité de Santé 2023] Haute Autorité de Santé. *Guide du parcours de soins – Maladie rénale chronique de l'adulte (MRC).* 2023. (Cité en page 67.)

[He 2016] Kaiming He, Xiangyu Zhang, Shaoqing Ren et Jian Sun. *Deep residual learning for image recognition*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016. (Cité en pages 99, 146, 240 et 262.)

[He 2017] Kaiming He, Georgia Gkioxari, Piotr Dollár et Ross Girshick. *Mask r-cnn*. In Proceedings of the IEEE international conference on computer vision, pages 2961–2969, 2017. (Cité en pages 97 et 133.)

[He 2022a] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár et Ross Girshick. *Masked autoencoders are scalable vision learners*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 16000–16009, 2022. (Cité en page 225.)

[He 2022b] Zhen He. *Depth-based auto-exposure management*, 2022. (Cité en page 37.)

[He 2022c] Zhen He, Jeffrey M. Dicarlo et Max J. Trejo. *Auto-exposure management of multi-component images*, 2022. (Cité en page 37.)

[He 2025] Xiankang He, Dongyan Guo, Hongji Li, Ruibo Li, Ying Cui et Chi Zhang. *Distill Any Depth : Distillation Creates a Stronger Monocular Depth Estimator*. arXiv preprint arXiv :2502.19204, 2025. (Cité en page 109.)

[Hemal 2007] AK Hemal, A Kumar, R Kumar, P Wadhwa, A Seth et NP Gupta. *Laparoscopic versus open radical nephrectomy for large renal tumors : a long-term prospective comparison*. The Journal of urology, vol. 177, no. 3, pages 862–866, 2007. (Cité en pages 74 et 76.)

[Hemal 2009] Ashok K Hemal et Anup Kumar. *A prospective comparison of laparoscopic and robotic radical nephrectomy for T1-2N0M0 renal cell carcinoma*. World journal of urology, vol. 27, pages 89–94, 2009. (Cité en page 74.)

[Heng 2009] Daniel YC Heng, Wanling Xie, Meredith M Regan, Mark A Warren, Ali Reza Golshayan, Chakshu Sahi, Bernhard J Eigl, J Dean Ruether, Tina Cheng, Scott North*et al.* *Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor–targeted agents : results from a large, multicenter study*. Journal of clinical oncology, vol. 27, no. 34, pages 5794–5799, 2009. (Cité en pages 72 et 73.)

[Henrich 2024] Pit Henrich, Jiawei Liu, Jiawei Ge, Samuel Schmidgall, Lauren Shepard, Ahmed Ezzat Ghazi, Franziska Mathis-Ullrich et Axel Krieger. *Tracking Tumors under Deformation from Partial Point Clouds using Occupancy Networks*. In 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 7227–7234. IEEE, 2024. (Cité en page 98.)

[Herrell 2009] Stanley D Herrell, David Morgan Kwartowitz, Paul M Milhoua et Robert L Galloway. *Toward image guided robotic surgery : system validation*. The Journal of urology, vol. 181, no. 2, pages 783–790, 2009. (Cité en page 96.)

[Herrlinger 1991] A Herrlinger, KM Schrott, G Schott et A Sigel. *What are the benefits of extended dissection of the regional renal lymph nodes in the therapy of renal cell carcinoma ?* The Journal of urology, vol. 146, no. 5, pages 1224–1227, 1991. (Cité en page 76.)

[Heryan 2016] Katarzyna Heryan, Andrzej Skalski, Janusz Gajda, Tomasz Drewniak et Jacek Jakubowski. *Registration of different phases of contrast-enhanced CT for facilitation of partial nephrectomy.* In 2016 IEEE International Conference on Imaging Systems and Techniques (IST), pages 255–260. IEEE, 2016. (Cité en page 44.)

[Himal 2002] HS Himal. *Minimally invasive (laparoscopic) surgery.* Surgical Endoscopy And Other Interventional Techniques, vol. 16, pages 1647–1652, 2002. (Cité en page 24.)

[Hinata 2021] Nobuyuki Hinata, Ryoichi Shiroki, Kazunari Tanabe, Masatoshi Eto, Atsushi Takenaka, Mutsushi Kawakita, Isao Hara, Fumiya Hongo, Naokazu Ibuki, Yasutomo Nasu*et al. Robot-assisted partial nephrectomy versus standard laparoscopic partial nephrectomy for renal hilar tumor : a prospective multi-institutional study.* International Journal of Urology, vol. 28, no. 4, pages 382–389, 2021. (Cité en page 76.)

[Hirschmuller 2005] Heiko Hirschmuller. *Accurate and efficient stereo processing by semi-global matching and mutual information.* In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), volume 2, pages 807–814. IEEE, 2005. (Cité en pages 157, 190 et 253.)

[Hirschmuller 2007] Heiko Hirschmuller. *Stereo processing by semiglobal matching and mutual information.* IEEE Transactions on pattern analysis and machine intelligence, vol. 30, no. 2, pages 328–341, 2007. (Cité en pages 97, 109, 157, 190 et 253.)

[Hirzle 2022] Teresa Hirzle, Fabian Fischbach, Julian Karlbauer, Pascal Jansen, Jan Gugenheimer, Enrico Rukzio et Andreas Bulling. *Understanding, addressing, and analysing digital eye strain in virtual reality head-mounted displays.* ACM Transactions on Computer-Human Interaction (TOCHI), vol. 29, no. 4, pages 1–80, 2022. (Cité en page 130.)

[Hora 2008] M Hora, O Hes, T Reischig, T Ürge, J Klecka, J Ferda, M Michal et V Eret. *Tumours in end-stage kidney.* In Transplantation proceedings, volume 40, pages 3354–3358. Elsevier, 2008. (Cité en page 69.)

[Hormann 2007] Kai Hormann, Bruno Lévy et Alla Sheffer. *Mesh parameterization : Theory and practice.* 2007. (Cité en page 253.)

[Hou 2010] Lei Hou, Shingo Kagami et Koichi Hashimoto. *Illumination-based synchronization of high-speed vision sensors.* Sensors, vol. 10, no. 6, pages 5530–5547, 2010. (Cité en page 127.)

[Hou 2013] Lei Hou, Shingo Kagami et Koichi Hashimoto. *Frame synchronization of high-speed vision sensors with respect to temporally encoded illumination*

in highly dynamic environments. Sensors, vol. 13, no. 4, pages 4102–4121, 2013. (Cité en page 127.)

[Hou 2023] Zhiqiang Hou, Jiaxin Zhao, Zhuo Wang, Sugang Ma, Wangsheng Yu et Jiulun Fan. *Object drift determination network based on dual-template joint decision-making in long-term visual tracking.* Journal of Visual Communication and Image Representation, vol. 97, page 103974, 2023. (Cité en page 230.)

[Hoznek 2007] A Hoznek, S Larré, L Salomon, A De La Taille et C-C Abbou. *Néphrectomie partielle par voie laparoscopique.* In Annales d'urologie, volume 41, pages 134–144. Elsevier, 2007. (Cité en page 82.)

[Hu 2007] Mingxing Hu, Graeme Penney, Philip Edwards, Michael Figl et David J Hawkes. *3D reconstruction of internal organ surfaces for minimal invasive surgery.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 68–77. Springer, 2007. (Cité en page 222.)

[Hu 2009] Mingxing Hu, Graeme P Penney, Daniel Rueckert, Philip J Edwards, Fernando Bello, Roberto Casula, Michael Figl et David J Hawkes. *Non-rigid reconstruction of the beating heart surface for minimally invasive cardiac surgery.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 34–42. Springer, 2009. (Cité en page 222.)

[Hu 2012] Mingxing Hu, Graeme Penney, Michael Figl, Philip Edwards, Fernando Bello, Roberto Casula, Daniel Rueckert et David Hawkes. *Reconstruction of a 3D surface from video that is robust to missing data and outliers : Application to minimally invasive surgery using stereo and mono endoscopes.* Medical image analysis, vol. 16, no. 3, pages 597–611, 2012. (Cité en page 222.)

[Hu 2016] Zexi Hu, Yuefang Gao, Dong Wang et Xuhong Tian. *A universal update-pacing framework for visual tracking.* In 2016 IEEE International Conference on Image Processing (ICIP), pages 1704–1708. IEEE, 2016. (Cité en page 229.)

[Hu 2018] Jie Hu, Li Shen et Gang Sun. *Squeeze-and-excitation networks.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7132–7141, 2018. (Cité en page 240.)

[Huang 2006] William C Huang, Andrew S Levey, Angel M Serio, Mark Snyder, Andrew J Vickers, Ganesh V Raj, Peter T Scardino et Paul Russo. *Chronic kidney disease after nephrectomy in patients with renal cortical tumours : a retrospective cohort study.* The lancet oncology, vol. 7, no. 9, pages 735–740, 2006. (Cité en page 79.)

[Huang 2019] Bo-Song Huang, Day-Fann Shen, Guo-Shiang Lin et Sin-Kuo Daniel Chai. *Multi-camera video synchronization based on feature point matching and refinement.* In 2019 IEEE/ACIS 18th International Conference on Com-

puter and Information Science (ICIS), pages 136–139. IEEE, 2019. (Cité en page 128.)

[Huang 2020] William C Huang, Nicholas M Donin, Andrew S Levey et Steven C Campbell. *Chronic kidney disease and kidney cancer surgery : new perspectives*. Journal of Urology, vol. 203, no. 3, pages 475–485, 2020. (Cité en pages 79 et 81.)

[Huang 2021] Ruizhen Huang, Chiyu Zhang, Xing Wang et Honglin Hu. *Partial nephrectomy versus radical nephrectomy for clinical T2 or higher stage renal tumors : A systematic review and meta-analysis*. Frontiers in oncology, vol. 11, page 680842, 2021. (Cité en page 75.)

[Huang 2022] Junjie Huang, David Ka-Wai Leung, Erica On-Ting Chan, Veeleah Lok, Sophia Leung, Iris Wong, Xiang-Qian Lao, Zhi-Jie Zheng, Peter Ka-Fung Chiu, Chi-Fai Ng*et al. A global trend analysis of kidney cancer incidence and mortality and their associations with smoking, alcohol consumption, and metabolic syndrome*. European urology focus, vol. 8, no. 1, pages 200–209, 2022. (Cité en pages 68 et 69.)

[Huang 2024] Qian Huang, Xiaotong Guo, Yiming Wang, Huashan Sun et Lijie Yang. *A survey of feature matching methods*. IET Image Processing, vol. 18, no. 6, pages 1385–1410, 2024. (Cité en page 217.)

[Hughes-Hallett 2014a] Archie Hughes-Hallett, Erik K Mayer, Hani J Marcus, Thomas P Cundy, Philip J Pratt, Ara W Darzi et Justin A Vale. *Augmented reality partial nephrectomy : examining the current status and future perspectives*. Urology, vol. 83, no. 2, pages 266–273, 2014. (Cité en page 94.)

[Hughes-Hallett 2014b] Archie Hughes-Hallett, Philip Pratt, Erik Mayer, Shirley Martin, Ara Darzi et Justin Vale. *Image guidance for all—TilePro display of 3-dimensionally reconstructed images in robotic partial nephrectomy*. Urology, vol. 84, no. 1, pages 237–243, 2014. (Cité en pages 95 et 100.)

[Huo 2023] Jiayi Huo, Changjiang Zhou, Bo Yuan, Qing Yang et Liqiang Wang. *Real-time dense reconstruction with binocular endoscopy based on StereoNet and ORB-SLAM*. Sensors, vol. 23, no. 4, page 2074, 2023. (Cité en page 223.)

[Ihler 2020] Sontje Ihler, Felix Kuhnke, Max-Heinrich Laves et Tobias Ortmaier. *Self-supervised domain adaptation for patient-specific, real-time tissue tracking*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 54–64. Springer, 2020. (Cité en page 222.)

[Ijsselsteijn 2005] Wijnand A Ijsselsteijn, Pieter JH Seuntiëns et Lydia MJ Meesters. *Human factors of 3D displays*. 3D Videocommunication : Algorithms, concepts and real-time systems in human centred communication, pages 217–233, 2005. (Cité en page 130.)

[Imre 2012] Evren Imre, Jean-Yves Guillemaut et Adrian Hilton. *Through-the-Lens multi-camera synchronisation and frame-drop detection for 3D reconstruction*. In 2012 Second International Conference on 3D Imaging, Modeling,

Processing, Visualization & Transmission, pages 395–402. IEEE, 2012. (Cité en page 129.)

[Itkowitz 2015] Brandon D. Itkowitz, Paul G. Griffiths, Nitish Swarup et Kamyar Ziaei. *System and method for integrated surgical table motion*, 2015. (Cité en page 36.)

[Izquierdo 2025] Sergio Izquierdo, Mohamed Sayed, Michael Firman, Guillermo Garcia-Hernando, Daniyar Turmukhambetov, Javier Civera, Oisin Mac Aodha, Gabriel Brostow et Jamie Watson. *MVSAnywhere : Zero-Shot Multi-View Stereo.* arXiv preprint arXiv :2503.22430, 2025. (Cité en page 109.)

[Jackson 2020] Peter Jackson, Richard Simon et Cristian Linte. *Surgical tracking, registration, and navigation characterization for image-guided renal interventions.* In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pages 5081–5084. IEEE, 2020. (Cité en page 96.)

[Jacobs 1991] M Jacobs, JC Verdeja et HS1688289 Goldstein. *Minimally invasive colon resection (laparoscopic colectomy)*, 1991. (Cité en page 23.)

[Jamil 2020] Marcus L Jamil, Jacob Keeley, Akshay Sood, Deepansh Dalela, Sohrab Arora, James O Peabody, Quoc-Dien Trinh, Mani Menon, Craig G Rogers et Firas Abdollah. *Long-term risk of recurrence in surgically treated renal cell carcinoma : a post hoc analysis of the Eastern Cooperative Oncology Group—American College Of Radiology Imaging Network E2805 trial cohort.* European urology, vol. 77, no. 2, pages 277–281, 2020. (Cité en page 73.)

[Janssen 2018] Martin Wilhelm Walter Janssen, Johannes Linxweiler, S Terwey, Sven Rugge, C-H Ohlmann, Frank Becker, Ch Thomas, A Neisius, JW Thüroff, Stefan Siemer*et al. Survival outcomes in patients with large clear cell renal cell carcinomas treated with nephron-sparing surgery versus radical nephrectomy : Results of a multicenter cohort with long-term follow-up.* PloS one, vol. 13, no. 5, page e0196427, 2018. (Cité en page 75.)

[Janus 1991] CL Janus et DS Mendelson. *Comparison of MRI and CT for study of renal and perirenal masses.* Critical reviews in diagnostic imaging, vol. 32, no. 2, pages 69–118, 1991. (Cité en page 71.)

[Jasinski 2022] Milosz Jasinski, Jerzy Siekiera et Maciej Tworkiewicz. *Ultrasound-guided renal mass biopsy and its clinical utility : a single-centre experience.* Urologia Internationalis, vol. 106, no. 6, pages 560–566, 2022. (Cité en page 72.)

[Javaheri 2025] Hamraz Javaheri, Omid Ghamarnejad, Ragnar Bade, Paul Lukowicz, Jakob Karolus et Gregor Alexander Stavrou. *Beyond the visible : preliminary evaluation of the first wearable augmented reality assistance system for pancreatic surgery.* International Journal of Computer Assisted Radiology and Surgery, vol. 20, no. 1, pages 117–129, 2025. (Cité en page 175.)

[Jayson 1998] Mary Jayson et Holt Sanders. *Increased incidence of serendipitously discovered renal cell carcinoma*. Urology, vol. 51, no. 2, pages 203–205, 1998. (Cité en page 70.)

[Jegou 2008] Herve Jegou, Matthijs Douze et Cordelia Schmid. *Hamming embedding and weak geometric consistency for large scale image search*. In Computer Vision–ECCV 2008 : 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part I 10, pages 304–317. Springer, 2008. (Cité en page 218.)

[Jeong 2017] In Gab Jeong, Yash S Khandwala, Jae Heon Kim, Deok Hyun Han, Shufeng Li, Ye Wang, Steven L Chang et Benjamin I Chung. *Association of robotic-assisted vs laparoscopic radical nephrectomy with perioperative outcomes and health care costs, 2003 to 2015*. Jama, vol. 318, no. 16, pages 1561–1568, 2017. (Cité en page 76.)

[Jewett 2011] Michael AS Jewett, Kamal Mattar, Joan Basiuk, Christopher G Morash, Stephen E Pautler, D Robert Siemens, Simon Tanguay, Ricardo A Rendon, Martin E Gleave, Darrel E Drachenberg*et al.* *Active surveillance of small renal masses : progression patterns of early stage kidney cancer*. European urology, vol. 60, no. 1, pages 39–44, 2011. (Cité en page 75.)

[Jia 2015] Li Jia, Qingling Jia, Yonggang Shang, Xingyou Dong et Longkun Li. *Vitamin C intake and risk of renal cell carcinoma : a meta-analysis*. Scientific reports, vol. 5, no. 1, page 17921, 2015. (Cité en page 69.)

[Jia 2021] Tingting Jia, Zeike A Taylor et Xiaojun Chen. *Long term and robust 6DoF motion tracking for highly dynamic stereo endoscopy videos*. Computerized Medical Imaging and Graphics, vol. 94, page 101995, 2021. (Cité en pages 99 et 223.)

[Jiang 2015] Long Jiang, Shaoting Zhang, Jie Yang, Xiahai Zhuang, Lixia Zhang et Lixu Gu. *A robust automated markerless registration framework for neurosurgery navigation*. The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 11, no. 4, pages 436–447, 2015. (Cité en page 220.)

[Jiang 2016] Jue Jiang, Yoshikazu Nakajima, Yoshio Sohma, Toki Saito, Taichi Kin, Horoshi Oyama et Nobuhito Saito. *Marker-less tracking of brain surface deformations by non-rigid registration integrating surface and vessel/sulci features*. International journal of computer assisted radiology and surgery, vol. 11, no. 9, pages 1687–1701, 2016. (Cité en page 220.)

[Jiang 2019a] Xingyu Jiang, Junjun Jiang, Aoxiang Fan, Zhongyuan Wang et Jiayi Ma. *Multiscale locality and rank preservation for robust feature matching of remote sensing images*. IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 9, pages 6462–6472, 2019. (Cité en page 218.)

[Jiang 2019b] Xingyu Jiang, Jiayi Ma, Junjun Jiang et Xiaojie Guo. *Robust feature matching using spatial clustering with heavy outliers*. IEEE Transactions on Image Processing, vol. 29, pages 736–746, 2019. (Cité en page 218.)

[Jiang 2022a] Pengbo Jiang, Raphael B Arada, Zhamshid Okhunov, Andrew S Afyouni, Akhil Peta, Andrew Brevik, Lillian Xie, Maged Ayad, Perry Xu, Kalon Morgan*et al.* *Multidisciplinary approach and outcomes of pretreatment small (cT1a) renal mass biopsy : single-center experience.* Journal of Endourology, vol. 36, no. 5, pages 703–711, 2022. (Cité en page 72.)

[Jiang 2022b] Xingyu Jiang, Yang Wang, Aoxiang Fan et Jiayi Ma. *Learning for mismatch removal via graph attention networks.* ISPRS Journal of Photogrammetry and Remote Sensing, vol. 190, pages 181–195, 2022. (Cité en page 219.)

[Jiao 2021] Licheng Jiao, Ruohan Zhang, Fang Liu, Shuyuan Yang, Biao Hou, Lingling Li et Xu Tang. *New generation deep learning for video object detection : A survey.* IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 8, pages 3195–3215, 2021. (Cité en pages 131 et 133.)

[Jing 2025] Junpeng Jing, Weixun Luo, Ye Mao et Krystian Mikolajczyk. *Stereo Any Video : Temporally Consistent Stereo Matching.* arXiv preprint arXiv :2503.05549, 2025. (Cité en page 109.)

[Joeres 2019] Fabian Joeres, Daniel Schindele, Maria Luz, Simon Blaschke, Nele Russwinkel, Martin Schostak et Christian Hansen. *How well do software assistants for minimally invasive partial nephrectomy meet surgeon information needs ? A cognitive task analysis and literature review study.* PloS one, vol. 14, no. 7, page e0219920, 2019. (Cité en pages 87 et 89.)

[Joeres 2021a] Fabian Joeres, Florian Heinrich, Danny Schott et Christian Hansen. *Towards natural 3D interaction for laparoscopic augmented reality registration.* Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization, vol. 9, no. 4, pages 384–391, 2021. (Cité en page 95.)

[Joeres 2021b] Fabian Joeres, Tonia Mielke et Christian Hansen. *Laparoscopic augmented reality registration for oncological resection site repair.* International Journal of Computer Assisted Radiology and Surgery, vol. 16, pages 1577–1586, 2021. (Cité en page 96.)

[John 2019] Nirmal Thampi John, Kyle A Blum et A Ari Hakimi. *Role of lymph node dissection in renal cell cancer.* In Urologic Oncology : Seminars and Original Investigations, volume 37, pages 187–192. Elsevier, 2019. (Cité en page 76.)

[Joseph 2023] Jithin Joseph, Sudhish N George et Kiran Raja. *Parameter-free matrix decomposition for specular reflections removal in endoscopic images.* IEEE Journal of Translational Engineering in Health and Medicine, vol. 11, pages 360–374, 2023. (Cité en page 132.)

[Jung 2021] Haijo Jung. *Basic physical principles and clinical applications of computed tomography.* Progress in Medical Physics, vol. 32, no. 1, pages 1–17, 2021. (Cité en page 40.)

[Kabala 1991] JE Kabala, DA Gillatt, RA Persad, JB Penry, JC Gingell et D Chadwick. *Magnetic resonance imaging in the staging of renal cell carcinoma.* The

British journal of radiology, vol. 64, no. 764, pages 683–689, 1991. (Cité en page 71.)

[Kall 1985] Bruce A Kall, Patrick J Kelly et Stephan J Goerss. *Interactive stereotactic surgical system for the removal of intracranial tumors utilizing the CO2 laser and CT-derived database*. IEEE transactions on biomedical engineering, no. 2, pages 112–116, 1985. (Cité en page 25.)

[Kalloo 2007] Anthony N Kalloo. *Natural orifice transluminal endoscopic surgery*. Gastroenterology & Hepatology, vol. 3, no. 3, page 183, 2007. (Cité en page 22.)

[Kamina 2009] Pierre Kamina. *Anatomie clinique, Tome 4 : organes urinaires et génitaux, pelvis, coupes du tronc*. VIII-383 p, 2009. (Cité en pages 58, 59, 61, 62, 63, 65, 66 et 67.)

[Kanazawa 2018] Angjoo Kanazawa, Michael J Black, David W Jacobs et Jitendra Malik. *End-to-end recovery of human shape and pose*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7122–7131, 2018. (Cité en pages 178 et 179.)

[Kang 2002] Jinman Kang, Isaac Cohen et Gerard Medioni. *Continuous multi-views tracking using tensor voting*. In Workshop on Motion and Video Computing, 2002. Proceedings., pages 181–186. IEEE, 2002. (Cité en page 129.)

[Kang 2005] Jinman Kang, Isaac Cohen et Gerard Medioni. *Persistent objects tracking across multiple non overlapping cameras*. In 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1, volume 2, pages 112–119. IEEE, 2005. (Cité en page 129.)

[Kang 2014] Xin Kang, Mahdi Azizian, Emmanuel Wilson, Kyle Wu, Aaron D Martin, Timothy D Kane, Craig A Peters, Kevin Cleary et Raj Shekhar. *Stereoscopic augmented reality for laparoscopic surgery*. Surgical endoscopy, vol. 28, pages 2227–2235, 2014. (Cité en page 172.)

[Kaplan 2016] Joshua R Kaplan, Ziho Lee, Daniel D Eun et Adam C Reese. *Complications of minimally invasive surgery and their management*. Current urology reports, vol. 17, pages 1–10, 2016. (Cité en page 25.)

[Kaplan 2023] Nicolas Kaplan, Mitchell Marques, Isabel Scharf, Kevin Yang, Lee Alkureishi, Chad Purnell, Pravin Patel et Linping Zhao. *Virtual reality and augmented reality in plastic and craniomaxillofacial surgery : a scoping review*. Bioengineering, vol. 10, no. 4, page 480, 2023. (Cité en page 51.)

[Karaev 2024a] Nikita Karaev, Iurii Makarov, Jianyuan Wang, Natalia Neverova, Andrea Vedaldi et Christian Rupprecht. *CoTracker3 : Simpler and better point tracking by pseudo-labelling real videos*. arXiv preprint arXiv :2410.11831, 2024. (Cité en pages 224, 226 et 257.)

[Karaev 2024b] Nikita Karaev, Ignacio Rocco, Benjamin Graham, Natalia Neverova, Andrea Vedaldi et Christian Rupprecht. *Cotracker : It is better to track together*. In European Conference on Computer Vision, pages 18–35. Springer, 2024. (Cité en pages 224, 225 et 226.)

[Karam 2014] Jose A Karam, Catherine E Devine, Diana L Urbauer, Marisa Lozano, Tapati Maity, Kamran Ahrar, Pheroze Tamboli, Nizar M Tannir et Christopher G Wood. *Phase 2 trial of neoadjuvant axitinib in patients with locally advanced nonmetastatic clear cell renal cell carcinoma.* European urology, vol. 66, no. 5, pages 874–880, 2014. (Cité en page 77.)

[Kartynnik 2019] Yury Kartynnik, Artsiom Ablavatski, Ivan Grishchenko et Matthias Grundmann. *Real-time facial surface geometry from monocular video on mobile GPUs.* arXiv preprint arXiv :1907.06724, 2019. (Cité en page 183.)

[Kato 2020] Hiroharu Kato, Deniz Beker, Mihai Morariu, Takahiro Ando, Toru Matsuoka, Wadim Kehl et Adrien Gaidon. *Differentiable rendering : A survey.* arXiv preprint arXiv :2006.12057, 2020. (Cité en page 226.)

[Kavoussi 2021] Nicholas L Kavoussi, Bryn Pitt, James M Ferguson, Josephine Granna, Andria Remirez, Naren Nimmagadda, Rachel Melnyk, Ahmed Ghazi, Eric J Barth, Robert J Webster*et al.* *Accuracy of touch-based registration during robotic image-guided partial nephrectomy before and after tumor resection in validated phantoms.* Journal of Endourology, vol. 35, no. 3, pages 362–368, 2021. (Cité en pages 96 et 100.)

[Ke 2023] Lei Ke, Mingqiao Ye, Martin Danelljan, Yu-Wing Tai, Chi-Keung Tang, Fisher Yu*et al.* *Segment anything in high quality.* Advances in Neural Information Processing Systems, vol. 36, pages 29914–29934, 2023. (Cité en page 238.)

[Kelly 2000] Patrick J Kelly. *Stereotactic surgery : What is past is prologue.* Neurosurgery, vol. 46, no. 1, page 16, 2000. (Cité en page 39.)

[Khaddad 2023] Abderrahmane Khaddad, Jean-Christophe Bernhard, Gaëlle Margue, Clément Michiels, Solène Ricard, Kilian Chandelon, Franck Bladou, Nicolas Bourdel et Adrien Bartoli. *A survey of augmented reality methods to guide minimally invasive partial nephrectomy.* World journal of urology, vol. 41, no. 2, pages 335–343, 2023. (Cité en pages 51, 94 et 176.)

[Khan 2025] Ufaq Khan, Umair Nawaz, Adnan Qayyum, Shazad Ashraf, Muhammad Bilal et Junaid Qadir. *Surgical Scene Understanding in the Era of Foundation AI Models : A Comprehensive Review.* arXiv preprint arXiv :2502.14886, 2025. (Cité en page 225.)

[Khene 2020] Zine-Eddine Khene, Benoit Peyronnet, Lucas Freton, Vivien Graffeille, Benjamin Pradere, Corentin Robert, Solène-Florence Kammerer-Jacquet, Grégory Verhoest, Nathalie Rioux-Leclercq, Shahrokh Shariat*et al.* *What is better for predicting morbidity of robotic partial nephrectomy—a score or your clinical judgement ?* European Urology Focus, vol. 6, no. 2, pages 313–319, 2020. (Cité en page 81.)

[Khene 2022] Zine-Eddine Khene, Clément Mazouin, Alessandro Larcher, Benoit Peyronnet, Anis Gasmi, Mathieu Roumiguié, Grégory Verhoest, Umberto Capitanio, Romain Mathieu, Nicolas Doumerc*et al.* *Predicting complications*

after robotic partial nephrectomy : back to simplicity. European Urology Focus, vol. 8, no. 3, pages 777–783, 2022. (Cité en page 81.)

[Khojasteh 2025] Samad Barri Khojasteh, David Fuentes-Jimenez, Daniel Pizarro, Yamid Espinel et Adrien Bartoli. *MIS-NeRF : Neural Radiance Fields in Minimally-Invasive Surgery.* International Journal of Computer Assisted Radiology and Surgery, 2025. (Cité en page 169.)

[Kim 2009] Jung Im Kim, Jeong Yeon Cho, Kyung Chul Moon, Hak Jong Lee et Seung Hyup Kim. *Segmental enhancement inversion at biphasic multidetector CT : characteristic finding of small renal oncocytoma.* Radiology, vol. 252, no. 2, pages 441–448, 2009. (Cité en page 71.)

[Kim 2013] Hyeongwoo Kim, Hailin Jin, Sunil Hadap et Inso Kweon. *Specular reflection separation using dark channel prior.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1460–1467, 2013. (Cité en page 132.)

[Kim 2021] Hyuno Kim et Masatoshi Ishikawa. *Sub-frame evaluation of frame synchronization for camera network using linearly oscillating light spot.* Sensors, vol. 21, no. 18, page 6148, 2021. (Cité en page 127.)

[Kirillov 2023] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo*et al.* *Segment anything.* In Proceedings of the IEEE/CVF international conference on computer vision, pages 4015–4026, 2023. (Cité en pages 225 et 238.)

[Klatte 2021] Tobias Klatte, Alessandro Berni, Sergio Serni et Riccardo Campi. *Intermediate-and long-term oncological outcomes of active surveillance for localized renal masses : a systematic review and quantitative analysis.* BJU international, vol. 128, no. 2, pages 131–143, 2021. (Cité en page 75.)

[Klein 2009] Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever et Josien PW Pluim. *Elastix : a toolbox for intensity-based medical image registration.* IEEE transactions on medical imaging, vol. 29, no. 1, pages 196–205, 2009. (Cité en page 44.)

[Klett 2021] Dane E Klett, Matvey Tsivian, Vignesh T Packiam, Christine M Lohse, Mohamed E Ahmed, Theodora A Potretzke, Ajay Gopalakrishna, Stephen A Boorjian, R Houston Thompson, Bradley C Leibovich*et al.* *Partial versus radical nephrectomy in clinical T2 renal masses.* International Journal of Urology, vol. 28, no. 11, pages 1149–1154, 2021. (Cité en page 75.)

[Knorr 2017] Sebastian Knorr, Simone Croci et Aljosa Smolic. *A modular scheme for artifact detection in stereoscopic omni-directional images.* In Proceedings of the Irish Machine Vision and Image Processing Conference, 2017. (Cité en page 131.)

[Kobayashi 2020] Satoshi Kobayashi, Byunghyun Cho, Jun Mutaguchi, Junichi Inokuchi, Katsunori Tatsugami, Makoto Hashizume et Masatoshi Eto. *Surgical*

navigation improves renal parenchyma volume preservation in robot-assisted partial nephrectomy : a propensity score matched comparative analysis. The Journal of Urology, vol. 204, no. 1, pages 149–156, 2020. (Cité en page 94.)

[Kojo 2024] Kosuke Kojo, Jaejong Kim, Tsukasa Saida, Tomoyuki Ohta, Keisuke Sano, Shuya Kandori, Akio Hoshi, Hiromitsu Negoro et Hiroyuki Nishiyama. *Practical step-by-step SYNAPSE VINCENT rendering of three-dimensional graphics in horseshoe kidney with bilateral varicoceles*. JMA journal, vol. 7, no. 4, pages 471–486, 2024. (Cité en page 44.)

[Kolotouros 2019] Nikos Kolotouros, Georgios Pavlakos, Michael J Black et Kostas Daniilidis. *Learning to reconstruct 3D human pose and shape via model-fitting in the loop*. In Proceedings of the IEEE/CVF international conference on computer vision, pages 2252–2261, 2019. (Cité en page 178.)

[Koo 2017] Bongjin Koo, Erol Özgür, Bertrand Le Roy, Emmanuel Buc et Adrien Bartoli. *Deformable registration of a preoperative 3D liver volume to a laparoscopy image using contour and shading cues*. In International conference on medical image computing and computer-assisted intervention, pages 326–334. Springer, 2017. (Cité en pages 172 et 175.)

[Koser 2007] Kevin Koser et Reinhard Koch. *Perspectively invariant normal features*. In 2007 IEEE 11th International Conference on Computer Vision, pages 1–8. IEEE, 2007. (Cité en page 230.)

[Krane 2012] L Spencer Krane, Theodore B Manny et Ashok K Hemal. *Is near infrared fluorescence imaging using indocyanine green dye useful in robotic partial nephrectomy : a prospective comparative study of 94 patients*. Urology, vol. 80, no. 1, pages 110–118, 2012. (Cité en page 48.)

[Kriegmair 2018] Maximilian C Kriegmair, Riccardo Bertolo, Pierre I Karakiewicz, Bradley C Leibovich, Borje Ljungberg, Maria C Mir, Idir Ouzaid, Maciej Salagierski, Michael Staehler, Hendrik van Poppel*et al.* *Systematic review of the management of local kidney cancer relapse*. European Urology Oncology, vol. 1, no. 6, pages 512–523, 2018. (Cité en page 74.)

[Krokidis 2017] Miltiadis E Krokidis, Franco Orsi, Konstantinos Katsanos, Thomas Helmberger et Andy Adam. *CIRSE guidelines on percutaneous ablation of small renal cell carcinoma*. Cardiovascular and interventional radiology, vol. 40, pages 177–191, 2017. (Cité en page 74.)

[Kulawik 2021] Joanna Kulawik et Mariusz Kubanek. *Detection of false synchronization of stereo image transmission using a convolutional neural network*. Symmetry, vol. 13, no. 1, page 78, 2021. (Cité en page 129.)

[Kunath 2017] Frank Kunath, Stefanie Schmidt, Laura-Maria Krabbe, Arkadiusz Miernik, Philipp Dahm, Anne Cleves, Mario Walther et Nils Kroeger. *Partial nephrectomy versus radical nephrectomy for clinical localised renal masses*. Cochrane Database of Systematic Reviews, no. 5, 2017. (Cité en pages 73 et 75.)

[Kuo 2013] Thomas Kuo, Santhoshkumar Sunderrajan et BS Manjunath. *Camera alignment using trajectory intersections in unsynchronized videos.* In Proceedings of the IEEE International Conference on Computer Vision, pages 1121–1128, 2013. (Cité en page 129.)

[Kupas 2019] David Kupas, Peter Torok, Andras Hajdu et Balazs Harangi. *Visualization of fibroid in laparoscopy videos using ultrasound image segmentation and augmented reality.* In 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pages 60–63. IEEE, 2019. (Cité en page 171.)

[Kurzhunov 2017] Dmitry Kurzhunov. *Novel reconstruction and quantification methods for Oxygen-17 magnetic resonance imaging at clinical field strengths.* PhD thesis, Albert-Ludwigs-Universität Freiburg im Breisgau, 2017. (Cité en page 42.)

[Kutikov 2009] Alexander Kutikov et Robert G Uzzo. *The RENAL nephrometry score : a comprehensive standardized system for quantitating renal tumor size, location and depth.* The Journal of urology, vol. 182, no. 3, pages 844–853, 2009. (Cité en page 72.)

[Kwoh 1988] Yik San Kwoh, Joahin Hou, Edmond A Jonckheere et Samad Hayati. *A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery.* IEEE transactions on biomedical engineering, vol. 35, no. 2, pages 153–160, 1988. (Cité en page 25.)

[Labrunie 2023] Mathieu Labrunie, Daniel Pizarro, Christophe Tilmant et Adrien Bartoli. *Automatic 3D/2D Deformable Registration in Minimally Invasive Liver Resection using a Mesh Recovery Network.* In MIDL, pages 1104–1123, 2023. (Cité en pages 51, 52 et 174.)

[Labrunie 2024] Mathieu Labrunie, Daniel Pizarro, Christophe Tilmant et Adrien Bartoli. *Generic Liver Modelling with Application to Mini-invasive Surgery Guidance.* In International Conference on Medical Imaging and Computer-Aided Diagnosis, pages 219–229. Springer, 2024. (Cité en pages 175 et 206.)

[Lamarca 2020] Jose Lamarca, Shaifali Parashar, Adrien Bartoli et JMM Montiel. *Defslam : Tracking and mapping of deforming scenes from monocular sequences.* IEEE Transactions on robotics, vol. 37, no. 1, pages 291–303, 2020. (Cité en page 223.)

[Lamarca 2022] Jose Lamarca, Juan J Gómez Rodríguez, Juan D Tardos et José MM Montiel. *Direct and sparse deformable tracking.* IEEE Robotics and Automation Letters, vol. 7, no. 4, pages 11450–11457, 2022. (Cité en page 223.)

[Lane 2009] Brian R Lane, Ho-Yee Tiong, Steven C Campbell, Amr F Fergany, Christopher J Weight, Benjamin T Larson, Andrew C Novick et Stuart M Flechner. *Management of the adrenal gland during partial nephrectomy.* The Journal of urology, vol. 181, no. 6, pages 2430–2437, 2009. (Cité en page 76.)

[Lane 2011] Brian R Lane, Paul Russo, Robert G Uzzo, Adrian V Hernandez, Stephen A Boorjian, R Houston Thompson, Amr F Fergany, Thomas E Love et

Steven C Campbell. *Comparison of cold and warm ischemia during partial nephrectomy in 660 solitary kidneys reveals predominant role of nonmodifiable factors in determining ultimate renal function.* The Journal of urology, vol. 185, no. 2, pages 421–427, 2011. (Cité en page 79.)

[Lapôtre-Ledoux 2023] Bénédicte Lapôtre-Ledoux, Laurent Remontet, Zoé Uhry, Emmanuelle Dantony, Pascale Grosclaude, Florence Molinié, Anne-Sophie Woronoff, Camille Lecoffre, Lafay Lionel, Gautier Defossez*et al. Incidence des principaux cancers en France métropolitaine en 2023 et tendances depuis 1990.* Bulletin Epidémiologique Hebdomadaire-BEH, 2023. (Cité en page 68.)

[Larcher 2015] Alessandro Larcher, Nicola Fossati, Francesco Mistretta, Giovanni Lughezzani, Giuliana Lista, Paolo Dell'Oglio, Alberto Abrate, Maxine Sun, Pierre Karakiewicz, Nazareno Suardi*et al. Long-term oncologic outcomes of laparoscopic renal cryoablation as primary treatment for small renal masses.* In Urologic Oncology : Seminars and Original Investigations, volume 33, pages 22–e1. Elsevier, 2015. (Cité en page 74.)

[Larousse 2024] Larousse. *Cœliochirurgie.* Encyclopédie Larousse en ligne, 2024. (Cité en page 24.)

[Lasser 2012] Michael S Lasser, Matthew Doscher, Aryeh Keehn, Victoria Chernyak, Evan Garfein et Reza Ghavamian. *Virtual surgical planning : a novel aid to robot-assisted laparoscopic partial nephrectomy.* Journal of endourology, vol. 26, no. 10, pages 1372–1379, 2012. (Cité en page 50.)

[Lavallée 1989] S Lavallée, Ph Cinquin, J Demongeot, AL Benabid, I Marque et M Djaid. *Computer assisted driving of a needle into the brain.* In Computer assisted radiology, Proceedings CAR, volume 89, pages 416–420, 1989. (Cité en page 25.)

[Le Roy 2019] B Le Roy, Erol Ozgur, Bongjin Koo, Emmanuel Buc et Adrien Bartoli. *Augmented reality guidance in laparoscopic hepatectomy with deformable semi-automatic computed tomography alignment (with video).* Journal of Visceral Surgery, vol. 156, no. 3, pages 261–262, 2019. (Cité en page 172.)

[Lebacle 2019] Cedric Lebacle, Karim Bensalah, Jean-Christophe Bernhard, Laurence Albiges, Brigitte Laguerre, Marine Gross-Goupil, Herve Baumert, Herve Lang, Thibault Tricard, Brigitte Duclos*et al. Evaluation of axitinib to downstage cT 2a renal tumours and allow partial nephrectomy : a phase II study.* BJU international, vol. 123, no. 5, pages 804–810, 2019. (Cité en page 77.)

[Lebeda 2012] Karel Lebeda, Jirı Matas et Ondrej Chum. *Fixing the locally optimized ransac–full experimental evaluation.* In British machine vision conference, volume 2. Citeseer Princeton, NJ, USA, 2012. (Cité en page 218.)

[Lee 2013] Soon-Young Lee, Jae-Young Sim, Chang-Su Kim et Sang-Uk Lee. *Correspondence matching of multi-view video sequences using mutual information*

based similarity measure. IEEE transactions on multimedia, vol. 15, no. 8, pages 1719–1731, 2013. (Cité en page 129.)

[Lee 2014] Hak J Lee, Michael A Liss et Ithaar H Derweesh. *Outcomes of partial nephrectomy for clinical T1b and T2 renal tumors*. Current opinion in urology, vol. 24, no. 5, pages 448–452, 2014. (Cité en page 75.)

[Lee 2017] GI Lee, MR Lee, I Green, M Allaf et MR Marohn. *Surgeons' physical discomfort and symptoms during robotic surgery : a comprehensive ergonomic survey study*. Surgical endoscopy, vol. 31, pages 1697–1706, 2017. (Cité en page 130.)

[Lei 2006] Cheng Lei et Yee-Hong Yang. *Tri-focal tensor-based multiple video synchronization with subframe optimization*. IEEE Transactions on Image Processing, vol. 15, no. 9, pages 2473–2480, 2006. (Cité en page 128.)

[Lenc 2016] Karel Lenc et Andrea Vedaldi. *Learning covariant feature detectors*. In Computer Vision–ECCV 2016 Workshops : Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14, pages 100–117. Springer, 2016. (Cité en page 217.)

[Leordeanu 2005] Marius Leordeanu et Martial Hebert. *A spectral technique for correspondence problems using pairwise constraints*. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, volume 2, pages 1482–1489. IEEE, 2005. (Cité en page 218.)

[Leow 2023] Jeffrey J Leow, Shagnik Ray, Shawn Dason, Eric A Singer et Steven L Chang. *The promise of neoadjuvant and adjuvant therapies for renal cancer*. Urologic Clinics, vol. 50, no. 2, pages 285–303, 2023. (Cité en page 77.)

[Leroy 2024] Vincent Leroy, Yohann Cabon et Jérôme Revaud. *Grounding image matching in 3d with mast3r*. In European Conference on Computer Vision, pages 71–91. Springer, 2024. (Cité en page 190.)

[Leutenegger 2011] Stefan Leutenegger, Margarita Chli et Roland Y Siegwart. *BRISK : Binary robust invariant scalable keypoints*. In 2011 International conference on computer vision, pages 2548–2555. Ieee, 2011. (Cité en page 146.)

[Levesque 2019] Marc Levesque, Jonathan D. Halderman, Brian D. Hoffman, Derek C. Liou et Nariman Majdi Nasab. *Closed-loop control of illumination in an endoscopic camera system*, 2019. (Cité en page 37.)

[Levey 2015] Andrew S Levey, Cassandra Becker et Lesley A Inker. *Glomerular filtration rate and albuminuria for detection and staging of acute and chronic kidney disease in adults : a systematic review*. Jama, vol. 313, no. 8, pages 837–846, 2015. (Cité en page 72.)

[Levi 2016] Gil Levi et Tal Hassner. *LATCH : learned arrangements of three patch codes*. In 2016 IEEE winter conference on applications of computer vision (WACV), pages 1–9. IEEE, 2016. (Cité en page 145.)

[Lévy 2023] Bruno Lévy, Sylvain Petitjean, Nicolas Ray et Jérome Maillot. *Least squares conformal maps for automatic texture atlas generation*. In Seminal Graphics Papers : Pushing the Boundaries, Volume 2, pages 193–202. 2023. (Cité en pages 231 et 253.)

[Li 2010] Xiangru Li et Zhanyi Hu. *Rejecting mismatches by correspondence function*. International Journal of Computer Vision, vol. 89, pages 1–17, 2010. (Cité en page 218.)

[Li 2015] Xinchao Li, Martha Larson et Alan Hanjalic. *Pairwise geometric matching for large-scale object retrieval*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 5153–5161, 2015. (Cité en page 218.)

[Li 2019] Ranyang Li, Junjun Pan, Yaqing Si, Bin Yan, Yong Hu et Hong Qin. *Specular reflections removal for endoscopic image sequences with adaptive-RPCA decomposition*. IEEE transactions on medical imaging, vol. 39, no. 2, pages 328–340, 2019. (Cité en page 132.)

[Li 2020a] Jinze Li, Lei Peng, Dehong Cao, Bo Cheng, Haocheng Gou, Yunxiang Li et Qiang Wei. *Comparison of perioperative outcomes of robot-assisted vs. laparoscopic radical nephrectomy : a systematic review and meta-analysis*. Frontiers in Oncology, vol. 10, page 551052, 2020. (Cité en page 76.)

[Li 2020b] Yang Li, Florian Richter, Jingpei Lu, Emily K Funk, Ryan K Orosco, Jianke Zhu et Michael C Yip. *Super : A surgical perception framework for endoscopic tissue manipulation with surgical robotics*. IEEE Robotics and Automation Letters, vol. 5, no. 2, pages 2294–2301, 2020. (Cité en page 226.)

[Li 2021] Yilin Li, Qingduo Kong, Hongyi Wei et Yongjun Wang. *Comparison of the complications between minimally invasive surgery and open surgical treatments for early-stage cervical cancer : A systematic review and meta-analysis*. PLoS One, vol. 16, no. 7, page e0253143, 2021. (Cité en page 25.)

[Li 2022a] Fengze Li, Jieming Ma, Zhongbei Tian, Ji Ge, Hai-Ning Liang, Yungang Zhang et Tianxi Wen. *Mirror-Yolo : a novel attention focus, instance segmentation and mirror detection model*. In 2022 7th International Conference on Frontiers of Signal Processing (ICFSP), pages 76–80. IEEE, 2022. (Cité en page 132.)

[Li 2022b] Le Li, Xing Zeng, Chunguang Yang, Waikeong Un et Zhiquan Hu. *Three-dimensional (3D) reconstruction and navigation in robotic-assisted partial nephrectomy (RAPN) for renal masses in the solitary kidney : A comparative study*. The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 18, no. 1, page e2337, 2022. (Cité en page 94.)

[Li 2022c] Zizhuo Li, Yong Ma, Xiaoguang Mei, Jun Huang et Jiayi Ma. *Guided neighborhood affine subspace embedding for feature matching*. Pattern Recognition, vol. 124, page 108489, 2022. (Cité en page 218.)

[Li 2023] Zhaoshuo Li, Wei Ye, Dilin Wang, Francis X Creighton, Russell H Taylor, Ganesh Venkatesh et Mathias Unberath. *Temporally consistent online*

depth estimation in dynamic scenes. In Proceedings of the IEEE/CVF winter conference on applications of computer vision, pages 3018–3027, 2023. (Cité en page 221.)

[Li 2024a] Fengze Li, Jieming Ma, Hai-Ning Liang, Zhongbei Tian, Zhijing Wu, Tianxi Wen et Dawei Liu. *A Comprehensive Survey of Specularity Detection : State-of-the-Art Techniques and Breakthroughs*. 2024. (Cité en page 131.)

[Li 2024b] Guiji Li et Ke Nai. *Robust tracking via coarse-to-fine redetection and spatial–temporal reliability evaluation*. Expert Systems with Applications, vol. 256, page 124927, 2024. (Cité en page 230.)

[Liang 2022] Xiaohong Liang, Caiqiang Xue, Xiaoyu Huang, Jinyan Wei et Junlin Zhou. *Value of energy spectrum CT parameters in the differential diagnosis of high-grade clear cell renal cell carcinoma and type II papillary renal cell carcinoma*. Acta Radiologica, vol. 63, no. 4, pages 545–552, 2022. (Cité en page 71.)

[Liao 2022] Zhanchen Liao, Zhitao Fang, Siqi Gou, Yong Luo, Yiqi Liu, Zhun He, Xin Li, Yansong Peng, Zheng Fu, Dongjin Li*et al.* *The role of diet in renal cell carcinoma incidence : an umbrella review of meta-analyses of observational studies*. BMC medicine, vol. 20, no. 1, page 39, 2022. (Cité en page 69.)

[Lin 2014] Wen-Yan Daniel Lin, Ming-Ming Cheng, Jiangbo Lu, Hongsheng Yang, Minh N Do et Philip Torr. *Bilateral functions for global motion modeling*. In Computer Vision–ECCV 2014 : 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13, pages 341–356. Springer, 2014. (Cité en page 218.)

[Lin 2017a] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He et Piotr Dollár. *Focal loss for dense object detection*. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988, 2017. (Cité en page 242.)

[Lin 2017b] Wen-Yan Lin, Fan Wang, Ming-Ming Cheng, Sai-Kit Yeung, Philip HS Torr, Minh N Do et Jiangbo Lu. *CODE : Coherence based decision boundaries for feature correspondence*. IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 1, pages 34–47, 2017. (Cité en page 218.)

[Lin 2020] Jiaying Lin, Guodong Wang et Rynson WH Lau. *Progressive mirror detection*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 3697–3705, 2020. (Cité en page 132.)

[Lin 2023a] Jiaying Lin et Rynson WH Lau. *Self-supervised pre-training for mirror detection*. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 12227–12236, 2023. (Cité en page 132.)

[Lin 2023b] Jiaying Lin, Xin Tan et Rynson WH Lau. *Learning to detect mirrors from videos via dual correspondences*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 9109–9118, 2023. (Cité en page 133.)

[Lin 2023c] Shan Lin, Albert J Miao, Jingpei Lu, Shunkai Yu, Zih-Yun Chiu, Florian Richter et Michael C Yip. *Semantic-super : a semantic-aware surgical*

perception framework for endoscopic tissue identification, reconstruction, and tracking. In 2023 IEEE International Conference on Robotics and Automation (ICRA), pages 4739–4746. IEEE, 2023. (Cité en page 226.)

[Lin 2023d] Zhefan Lin, Chen Lei et Liangjing Yang. *Modern image-guided surgery : A narrative review of medical image processing and visualization*. Sensors, vol. 23, no. 24, page 9872, 2023. (Cité en pages 38 et 39.)

[Lindenberger 2023] Philipp Lindenberger, Paul-Edouard Sarlin et Marc Pollefeys. *Lightglue : Local feature matching at light speed*. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 17627–17638, 2023. (Cité en pages 146 et 219.)

[Litynski 1999] Grzegorz S Litynski. *Mouret, Dubois, and Perissat : the laparoscopic breakthrough in Europe (1987-1988)*. JSLS : Journal of the Society of Laparoendoscopic Surgeons, vol. 3, no. 2, page 163, 1999. (Cité en page 23.)

[Liu 2010] Hairong Liu et Shuicheng Yan. *Common visual pattern discovery via spatially coherent correspondences*. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1609–1616. IEEE, 2010. (Cité en page 218.)

[Liu 2012] Yiguang Liu, Menglong Yang et Zhisheng You. *Video synchronization based on events alignment*. Pattern Recognition Letters, vol. 33, no. 10, pages 1338–1348, 2012. (Cité en page 129.)

[Liu 2017] Xinyang Liu, Christina E Rice et Raj Shekhar. *Fast calibration of electromagnetically tracked oblique-viewing rigid endoscopes*. International journal of computer assisted radiology and surgery, vol. 12, pages 1685–1695, 2017. (Cité en page 110.)

[Liu 2019a] Chen Liu, Kihwan Kim, Jinwei Gu, Yasutaka Furukawa et Jan Kautz. *Planercnn : 3d plane detection and reconstruction from a single image*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 4450–4459, 2019. (Cité en page 133.)

[Liu 2019b] Xiaoqiu Liu, Giulia Peveri, Cristina Bosetti, Vincenzo Bagnardi, Claudia Specchia, Silvano Gallus et Alessandra Lugo. *Dose-response relationships between cigarette smoking and kidney cancer : a systematic review and meta-analysis*. Critical Reviews in Oncology/Hematology, vol. 142, pages 86–93, 2019. (Cité en page 69.)

[Liu 2020] Liang Liu, Jiangning Zhang, Ruifei He, Yong Liu, Yabiao Wang, Ying Tai, Donghao Luo, Chengjie Wang, Jilin Li et Feiyue Huang. *Learning by analogy : Reliable supervision from transformations for unsupervised optical flow estimation*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 6489–6498, 2020. (Cité en page 222.)

[Liu 2021a] Hui Liu, Qing-Fang Kong, Jian Li, Yu-Qing Wu, Ke-Hao Pan, Bin Xu, Ya-Li Wang et Ming Chen. *A meta-analysis for comparison of partial nephrectomy vs. radical nephrectomy in patients with pT3a renal cell carcinoma*.

Translational andrology and urology, vol. 10, no. 3, page 1170, 2021. (Cité en page 76.)

[Liu 2021b] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin et Baining Guo. *Swin transformer : Hierarchical vision transformer using shifted windows*. In Proceedings of the IEEE/CVF international conference on computer vision, pages 10012–10022, 2021. (Cité en page 132.)

[Liu 2022a] Chang Liu, Xiao-Fan Chen, Chun-Juan Bo et Dong Wang. *Long-term visual tracking : review and experimental comparison*. Machine Intelligence Research, vol. 19, pages 512–530, 2022. (Cité en page 229.)

[Liu 2022b] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell et Saining Xie. *A convnet for the 2020s*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 11976–11986, 2022. (Cité en page 224.)

[Liu 2023] Ziteng Liu, Wenpeng Gao, Jiahua Zhu, Zhi Yu et Yili Fu. *Surface deformation tracking in monocular laparoscopic video*. Medical image analysis, vol. 86, page 102775, 2023. (Cité en page 220.)

[Liu 2024a] Hengyu Liu, Yifan Liu, Chenxin Li, Wuyang Li et Yixuan Yuan. *Lgs : A light-weight 4d gaussian splatting for efficient surgical scene reconstruction*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 660–670. Springer, 2024. (Cité en page 190.)

[Liu 2024b] Wen Pei Liu et Wendy Chan. *Touchscreen user interface for interacting with a virtual model*, 2024. (Cité en page 31.)

[Lobo 2020] Jennifer M Lobo, Matthew B Clements, Daniel P Bitner, Matthew D Mikula, Sean W Noona, Mark I Sultan, Helen P Cathro, Drew L Lambert, Noah S Schenkman et Tracey L Krupski. *Does renal mass biopsy influence multidisciplinary treatment recommendations ?* Scandinavian Journal of Urology, vol. 54, no. 1, pages 27–32, 2020. (Cité en page 72.)

[Long 2012] Jean-Alexandre Long, Rachid Yakoubi, Byron Lee, Julien Guillotreau, Riccardo Autorino, Humberto Laydner, Remi Eyraud, Robert J Stein, Jihad H Kaouk et Georges-Pascal Haber. *Robotic versus laparoscopic partial nephrectomy for complex tumors : comparison of perioperative outcomes*. European urology, vol. 61, no. 6, pages 1257–1262, 2012. (Cité en page 81.)

[Longo 2021] Umile Giuseppe Longo, Sergio De Salvatore, Vincenzo Candela, Giuliano Zollo, Giovanni Calabrese, Sara Fioravanti, Lucia Giannone, Anna Marchetti, Maria Grazia De Marinis et Vincenzo Denaro. *Augmented reality, virtual reality and artificial intelligence in orthopedic surgery : a systematic review*. Applied Sciences, vol. 11, no. 7, page 3253, 2021. (Cité en page 51.)

[Loper 2023] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll et Michael J Black. *SMPL : A skinned multi-person linear model*. Seminal Graphics Papers : Pushing the Boundaries, Volume 2, pages 851–866, 2023. (Cité en page 178.)

[Lopes Vendrami 2017] Camila Lopes Vendrami, Carolina Parada Villavicencio, Todd J DeJulio, Argha Chatterjee, David D Casalino, Jeanne M Horowitz, Daniel T Oberlin, Guang-Yu Yang, Paul Nikolaidis et Frank H Miller. *Differentiation of solid renal tumors with multiparametric MR imaging.* Radiographics, vol. 37, no. 7, pages 2026–2042, 2017. (Cité en page 71.)

[Loshchilov 2017] Ilya Loshchilov et Frank Hutter. *Decoupled weight decay regularization.* arXiv preprint arXiv :1711.05101, 2017. (Cité en pages 202, 242 et 262.)

[Lowe 2004] David G Lowe. *Distinctive image features from scale-invariant keypoints.* International journal of computer vision, vol. 60, pages 91–110, 2004. (Cité en pages 129, 138, 139, 146, 217, 244, 250, 251, 270 et 271.)

[Lowrance 2014] William T Lowrance, Juan Ordoñez, Natalia Udaltsova, Paul Russo et Alan S Go. *CKD and the risk of incident cancer.* Journal of the American Society of Nephrology, vol. 25, no. 10, pages 2327–2334, 2014. (Cité en page 69.)

[Lu 2012] Cheng Lu et Mrinal Mandal. *A robust technique for motion-based video sequences temporal alignment.* IEEE Transactions on Multimedia, vol. 15, no. 1, pages 70–82, 2012. (Cité en page 129.)

[Lu 2025] Xiaojun Lu, Zhe Yan et Ziyun Fan. *GaLAM : Two-stage outlier detection algorithm.* IEEE Access, 2025. (Cité en page 219.)

[Lucas 1981] Bruce D Lucas et Takeo Kanade. *An iterative image registration technique with an application to stereo vision.* In IJCAI'81 : 7th international joint conference on Artificial intelligence, volume 2, pages 674–679, 1981. (Cité en page 221.)

[Ludwig 2017] Wesley W Ludwig, Michael A Gorin, Phillip M Pierorazio et Mohamad E Allaf. *Frontiers in robot-assisted retroperitoneal oncological surgery.* Nature Reviews Urology, vol. 14, no. 12, pages 731–741, 2017. (Cité en page 78.)

[Ludwig 2020] Daniel R Ludwig, David H Ballard, Anup S Shetty, Cary L Siegel et Motoyo Yano. *Apparent diffusion coefficient distinguishes malignancy in T1-hyperintense small renal masses.* American Journal of Roentgenology, vol. 214, no. 1, pages 114–121, 2020. (Cité en page 71.)

[Lukezic 2017] Alan Lukezic, Tomas Vojir, Luka ˇCehovin Zajc, Jiri Matas et Matej Kristan. *Discriminative correlation filter with channel and spatial reliability.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 6309–6318, 2017. (Cité en page 222.)

[Luo 2020a] Huoling Luo, Dalong Yin, Shugeng Zhang, Deqiang Xiao, Baochun He, Fanzheng Meng, Yanfang Zhang, Wei Cai, Shenghao He, Wenyu Zhang*et al.* *Augmented reality navigation for liver resection with a stereoscopic laparoscope.* Computer methods and programs in biomedicine, vol. 187, page 105099, 2020. (Cité en page 172.)

[Luo 2020b] Zixin Luo, Lei Zhou, Xuyang Bai, Hongkai Chen, Jiahui Zhang, Yao Yao, Shiwei Li, Tian Fang et Long Quan. *Aslfeat : Learning local features of accurate shape and localization*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 6589–6598, 2020. (Cité en page 217.)

[Lüthi 2017] Marcel Lüthi, Thomas Gerig, Christoph Jud et Thomas Vetter. *Gaussian process morphable models*. IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 8, pages 1860–1873, 2017. (Cité en page 175.)

[Ma 2014] Jiayi Ma, Ji Zhao, Jinwen Tian, Alan L Yuille et Zhuowen Tu. *Robust point matching via vector field consensus*. IEEE Transactions on Image Processing, vol. 23, no. 4, pages 1706–1721, 2014. (Cité en page 218.)

[Ma 2015] Chao Ma, Xiaokang Yang, Chongyang Zhang et Ming-Hsuan Yang. *Long-term correlation tracking*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 5388–5396, 2015. (Cité en page 228.)

[Ma 2018] Qing Ma, Xu Du, Jiahao Wang, Yong Ma et Jiayi Ma. *Robust feature matching via gaussian field criterion for remote sensing image registration*. Journal of Real-Time Image Processing, vol. 15, pages 523–536, 2018. (Cité en page 218.)

[Ma 2019a] Jiayi Ma, Xingyu Jiang, Junjun Jiang, Ji Zhao et Xiaojie Guo. *LMR : Learning a two-class classifier for mismatch removal*. IEEE Transactions on Image Processing, vol. 28, no. 8, pages 4045–4059, 2019. (Cité en page 219.)

[Ma 2019b] Jiayi Ma, Ji Zhao, Junjun Jiang, Huabing Zhou et Xiaojie Guo. *Locality preserving matching*. International Journal of Computer Vision, vol. 127, pages 512–531, 2019. (Cité en page 218.)

[Ma 2019c] Ruibin Ma, Rui Wang, Stephen Pizer, Julian Rosenman, Sarah K McGill et Jan-Michael Frahm. *Real-time 3D reconstruction of colonoscopic surfaces for determining missing regions*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 573–582. Springer, 2019. (Cité en page 223.)

[Ma 2022] Jiayi Ma, Zizhuo Li, Kaining Zhang, Zhenfeng Shao et Guobao Xiao. *Robust feature matching via neighborhood manifold representation consensus*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 183, pages 196–209, 2022. (Cité en page 218.)

[Maas 2013] Andrew L Maas, Awni Y Hannun, Andrew Y Ng*et al.* *Rectifier nonlinearities improve neural network acoustic models*. In Proc. icml, volume 30, page 3. Atlanta, GA, 2013. (Cité en page 202.)

[MacLennan 2012] Steven MacLennan, Mari Imamura, Marie C Lapitan, Muhammad Imran Omar, Thomas BL Lam, Ana M Hilvano-Cabungcal, Pam Royle, Fiona Stewart, Graeme MacLennan, Sara J MacLennan*et al.* *Systematic review of perioperative and quality-of-life outcomes following surgical mana-*

gement of localised renal cancer. European urology, vol. 62, no. 6, pages 1097–1117, 2012. (Cité en pages 75 et 76.)

[Macleod 2013] Liam C Macleod, James M Hotaling, Jonathan L Wright, Michael T Davenport, John L Gore, Jonathan Harper et Emily White. *Risk factors for renal cell carcinoma in the VITAL study.* The Journal of urology, vol. 190, no. 5, pages 1657–1661, 2013. (Cité en page 69.)

[Maddox 2018] Michael M Maddox, Allison Feibus, James Liu, Julie Wang, Raju Thomas et Jonathan L Silberstein. *3D-printed soft-tissue physical models of renal malignancies for individualized surgical simulation : a feasibility study.* Journal of robotic surgery, vol. 12, pages 27–33, 2018. (Cité en page 45.)

[Magaribuchi 2024] Toshihiro Magaribuchi, Kimihiko Masui, Takayuki Goto, Ryoichi Saito et Takashi Kobayashi. *Current status and challenges of 3D navigation in partial nephrectomy.* Mini-invasive Surgery, vol. 8, pages N–A, 2024. (Cité en pages 94 et 96.)

[Mahmoud 2016] Nader Mahmoud, Iñigo Cirauqui, Alexandre Hostettler, Christophe Doignon, Luc Soler, Jacques Marescaux et Jose Maria Martinez Montiel. *ORBSLAM-based endoscope tracking and 3D reconstruction.* In International workshop on computer-assisted and robotic endoscopy, pages 72–83. Springer, 2016. (Cité en page 223.)

[Mahmoud 2018] Nader Mahmoud, Toby Collins, Alexandre Hostettler, Luc Soler, Christophe Doignon et Jose Maria Martinez Montiel. *Live tracking and dense reconstruction for handheld monocular endoscopy.* IEEE transactions on medical imaging, vol. 38, no. 1, pages 79–89, 2018. (Cité en page 223.)

[Mair 2010] Elmar Mair, Gregory D Hager, Darius Burschka, Michael Suppa et Gerhard Hirzinger. *Adaptive and generic corner detection based on the accelerated segment test.* In Computer Vision–ECCV 2010 : 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part II 11, pages 183–196. Springer, 2010. (Cité en page 145.)

[Makiyama 2023] Kazuhide Makiyama, Mitsuru Komeya, Tomoyuki Tatenuma, Go Noguchi et Shinji Ohtake. *Patient-specific simulations and navigation systems for partial nephrectomy.* International Journal of Urology, vol. 30, no. 12, pages 1087–1095, 2023. (Cité en pages 94 et 96.)

[Makki 2023] Karim Makki, Kilian Chandelon et Adrien Bartoli. *Elliptical specularity detection in endoscopy with application to normal reconstruction.* International Journal of Computer Assisted Radiology and Surgery, vol. 18, no. 7, pages 1323–1328, 2023. (Cité en pages 131 et 219.)

[Makuuchi 1998] Masatoshi Makuuchi, Guido Torzilli et Junji Machi. *History of intraoperative ultrasound.* Ultrasound in medicine & biology, vol. 24, no. 9, pages 1229–1242, 1998. (Cité en page 47.)

[Malti 2012] Abed Malti, Adrien Bartoli et Toby Collins. *Template-based conformal shape-from-motion-and-shading for laparoscopy.* In International Conference

on Information Processing in Computer-Assisted Interventions, pages 1–10. Springer, 2012. (Cité en page 222.)

[Malti 2014] Abed Malti et Adrien Bartoli. *Combining conformal deformation and cook–torrance shading for 3-d reconstruction in laparoscopy.* IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, pages 1684–1692, 2014. (Cité en page 222.)

[Manfredi 2022] Matteo Manfredi, Federico Piramide, Daniele Amparore, Mariano Burgio, Giovanni Busacca, Marco Colombo, Juliette Meziere, Stefano Granato, Alberto Piana, Sabrina De Cillis*et al. Augmented reality : The smart way to guide robotic urologic surgery.* Mini-invasive Surgery, vol. 6, pages N–A, 2022. (Cité en page 94.)

[Manni 2024] G Manni, C Lauretti, F Prata, R Papalia, L Zollo et P Soda. *BodySLAM : A Generalized Monocular Visual SLAM Framework for Surgical Applications.* arXiv preprint arXiv :2408.03078, 2024. (Cité en page 223.)

[Marcal 2013] Leonardo P Marcal, Madhavi Patnana, Priya Bhosale et Deepak G Bedi. *Intraoperative abdominal ultrasound in oncologic imaging.* World Journal of Radiology, vol. 5, no. 3, page 51, 2013. (Cité en page 47.)

[Marconi 2016] Lorenzo Marconi, Saeed Dabestani, Thomas B Lam, Fabian Hofmann, Fiona Stewart, John Norrie, Axel Bex, Karim Bensalah, Steven E Canfield, Milan Hora*et al. Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy.* European urology, vol. 69, no. 4, pages 660–673, 2016. (Cité en page 72.)

[Margue 2023] Gaëlle Margue, Alexandre Ingels, Karim Bensalah, Nicolas Doumerc, Christophe Vaessen, Morgan Roupret, François Audenet, Arnaud Mejean, Franck Bruyere, Jonathan Olivier*et al. Late complications and 5 years outcomes of robotic partial nephrectomy in France : prospective assessment in the French Kidney Cancer Research Network (UroCCR 10).* World Journal of Urology, vol. 41, no. 8, pages 2281–2288, 2023. (Cité en page 81.)

[Margulis 2007] Vitaly Margulis, Ricardo F Sánchez-Ortiz, Pheroze Tamboli, Daniel D Cohen, David A Swanson et Christopher G Wood. *Renal cell carcinoma clinically involving adjacent organs : experience with aggressive surgical management.* Cancer : Interdisciplinary International Journal of the American Cancer Society, vol. 109, no. 10, pages 2025–2030, 2007. (Cité en page 78.)

[Marmol Velez 2017] Andres Marmol Velez, Thierry Peynot, Anders Eriksson, Anjali Jaiprakash et Jonathan Roberts. *Evaluation of keypoint detectors and descriptors in arthroscopic images for feature-based matching applications.* IEEE Robotics and Automation Letters, no. 99, pages 1–8, 2017. (Cité en page 220.)

[Marshall 2019] Keith Marshall, Rebecca Anne Cuthbertson et Gordon Thomas Deanel. *Camera control*, 2019. (Cité en pages 31 et 34.)

[Martin 2020] Xavier Martin, Jocelyne Troccaz, Jean-Yves Chapelon, Cyril Lafon*et al.* *La Robotisation en Chirurgie : Etat des Lieux*, 2020. (Cité en page 29.)

[Masson-Lecomte 2013] Alexandra Masson-Lecomte, Karim Bensalah, Elise Seringe, Christophe Vaessen, Alexandre de la Taille, Nicolas Doumerc, Pascal Rischmann, Franck Bruyère, Laurent Soustelle, Stéphane Droupy*et al.* *A prospective comparison of surgical and pathological outcomes obtained after robot-assisted or pure laparoscopic partial nephrectomy in moderate to complex renal tumours : results from a French multicentre collaborative study.* BJU international, vol. 111, no. 2, pages 256–263, 2013. (Cité en page 81.)

[Matas 2004] Jiri Matas, Ondrej Chum, Martin Urban et Tomás Pajdla. *Robust wide-baseline stereo from maximally stable extremal regions.* Image and vision computing, vol. 22, no. 10, pages 761–767, 2004. (Cité en page 145.)

[Matthews 2004] Lain Matthews, Takahiro Ishikawa et Simon Baker. *The template update problem.* IEEE transactions on pattern analysis and machine intelligence, vol. 26, pages 810–815, 2004. (Cité en page 228.)

[Maurer 2017] Martin H Maurer, Kirsi Hannele Härmä et Harriet Thoeny. *Diffusion-weighted genitourinary imaging.* Radiologic Clinics, vol. 55, no. 2, pages 393–411, 2017. (Cité en page 71.)

[Maxwell 2007] NJ Maxwell, N Saleem Amer, E Rogers, D Kiely, P Sweeney et AP Brady. *Renal artery embolisation in the palliative treatment of renal carcinoma.* The British journal of radiology, vol. 80, no. 950, pages 96–102, 2007. (Cité en page 76.)

[May 2009] Matthias May, S Brookman-Amissah, S Pflanz, J Roigas, B Hoschke et F Kendel. *Pre-operative renal arterial embolisation does not provide survival benefit in patients with radical nephrectomy for renal cell carcinoma.* The British journal of radiology, vol. 82, no. 981, pages 724–731, 2009. (Cité en page 76.)

[McDowall 2017] Ian E. McDowall, Jeffrey M. Dicarlo, Brian D. Hoffman et William Jason Culman. *Simultaneous white light and hyperspectral light imaging systems*, 2017. (Cité en page 37.)

[Mehrabi 2015] Maziar Mehrabi. *Video synchronization in the cloud using visible light communication.* 2015. (Cité en page 127.)

[Mei 2011] Xue Mei et Haibin Ling. *Robust visual tracking and vehicle classification via sparse representation.* IEEE transactions on pattern analysis and machine intelligence, vol. 33, pages 2259–2272, 2011. (Cité en page 229.)

[Mei 2021] Haiyang Mei, Bo Dong, Wen Dong, Pieter Peers, Xin Yang, Qiang Zhang et Xiaopeng Wei. *Depth-aware mirror segmentation.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 3044–3053, 2021. (Cité en page 133.)

[Melloni 2015] Andrea Melloni, Silvia Lameri, Paolo Bestagini, Marco Tagliasacchi et Stefano Tubaro. *Near-duplicate detection and alignment for multi-view*

videos. In 2015 IEEE International Conference on Image Processing (ICIP), pages 2444–2448. IEEE, 2015. (Cité en pages 128 et 137.)

[Mellouki 2021] Adil Mellouki, Imad Bentellis, Arnoult Morrone, Nicolas Doumerc, Jean-Baptiste Beauval, Morgane Roupret, François-Xavier Nouhaud, Cedric Lebacle, Jean-Alexandre Long, Daniel Chevallier*et al.* *Evaluation of oncological outcomes of robotic partial nephrectomy according to the type of hilar control approach (On-clamp vs Off-clamp), a multicentric study of the French network of research on kidney cancer—UROCCR 58-NCT03293563*. World Journal of Urology, pages 1–8, 2021. (Cité en page 81.)

[Melo 2011] Rui Melo, Joao P Barreto et Gabriel Falcao. *A new solution for camera calibration and real-time image distortion correction in medical endoscopy–initial technical evaluation*. IEEE Transactions on Biomedical Engineering, vol. 59, no. 3, pages 634–644, 2011. (Cité en page 110.)

[Mendoza-Ramírez 2023] Carlos E Mendoza-Ramírez, Juan C Tudon-Martinez, Luis C Félix-Herrán, Jorge de J Lozoya-Santos et Adriana Vargas-Martínez. *Augmented reality : survey*. Applied Sciences, vol. 13, no. 18, page 10491, 2023. (Cité en page 50.)

[Meyer-Szary 2022] Jarosław Meyer-Szary, Marlon Souza Luis, Szymon Mikulski, Agastya Patel, Finn Schulz, Dmitry Tretiakow, Justyna Fercho, Kinga Jaguszewska, Mikołaj Frankiewicz, Ewa Pawłowska*et al.* *The role of 3D printing in planning complex medical procedures and training of medical professionals—cross-sectional multispecialty review*. International journal of environmental research and public health, vol. 19, no. 6, page 3331, 2022. (Cité en page 45.)

[Meyer 2008] Benjamin Meyer, Timo Stich, Marcus A Magnor et Marc Pollefeys. *Subframe Temporal Alignment of Non-Stationary Cameras*. In BMVC, pages 1–10, 2008. (Cité en page 129.)

[Mezger 2013] Uli Mezger, Claudia Jendrewski et Michael Bartels. *Navigation in surgery*. Langenbeck's archives of surgery, vol. 398, pages 501–514, 2013. (Cité en page 39.)

[Michiels 2019] C Michiels, E Jambon et JC Bernhard. *Measurement of the accuracy of 3D-printed medical models to be used for robot-assisted partial nephrectomy*. American Journal of Roentgenology, vol. 213, no. 3, pages 626–631, 2019. (Cité en page 45.)

[Michiels 2020] C Michiels, A Boulenger de Hauteclocque, J Sarrazin, F Cornelis, S Ricard, M Faessel, N Grenier, F Bos et JC Bernhard. *Utilisation de l'impression 3D avant néphrectomie partielle pour tumeur du rein*. Progrès en Urologie, vol. 30, no. 15, pages 970–975, 2020. (Cité en page 46.)

[Michiels 2021a] Clément Michiels, Eva Jambon, Julien Sarrazin, Astrid Boulenger de Hauteclocque, Solène Ricard, Nicholas Grenier, Matthieu Faessel, Frédéric Bos et Jean-Christophe Bernhard. *Comprehensive review of 3D printing use in medicine : comparison with practical applications in urology*.

Progres en Urologie : Journal de L'association Francaise D'urologie et de la Societe Francaise D'urologie, vol. 31, no. 12, pages 762–771, 2021. (Cité en page 46.)

[Michiels 2021b] Clément Michiels, Zine-Eddine Khene, Thomas Prudhomme, Astrid Boulenger de Hauteclocque, François H Cornelis, Mélanie Percot, Hélène Simeon, Laure Dupitout, Henri Bensadoun, Grégoire Capon*et al. 3D-Image guided robotic-assisted partial nephrectomy : a multi-institutional propensity score-matched analysis (UroCCR study 51)*. World Journal of Urology, pages 1–11, 2021. (Cité en page 50.)

[Mikolajczyk 2004] Krystian Mikolajczyk et Cordelia Schmid. *Scale & affine invariant interest point detectors*. International journal of computer vision, vol. 60, pages 63–86, 2004. (Cité en pages 145 et 217.)

[Mileto 2015] Achille Mileto, Rendon C Nelson, Erik K Paulson et Daniele Marin. *Dual-energy MDCT for imaging the renal mass*. American Journal of Roentgenology, vol. 204, no. 6, pages W640–W647, 2015. (Cité en page 71.)

[Milgram 1995] Paul Milgram, Haruo Takemura, Akira Utsumi et Fumio Kishino. *Augmented reality : A class of displays on the reality-virtuality continuum*. In Telemanipulator and telepresence technologies, volume 2351, pages 282–292. Spie, 1995. (Cité en page 50.)

[Miller 2008] David C Miller, Matthias Schonlau, Mark S Litwin, Julie Lai et Christopher S Saigal. *Renal and cardiovascular morbidity after partial or radical nephrectomy*. Cancer : Interdisciplinary International Journal of the American Cancer Society, vol. 112, no. 3, pages 511–520, 2008. (Cité en page 79.)

[Milletari 2016] Fausto Milletari, Nassir Navab et Seyed-Ahmad Ahmadi. *V-net : Fully convolutional neural networks for volumetric medical image segmentation*. In 2016 fourth international conference on 3D vision (3DV), pages 565–571. Ieee, 2016. (Cité en page 242.)

[Min 2014] Dongbo Min, Sunghwan Choi, Jiangbo Lu, Bumsub Ham, Kwanghoon Sohn et Minh N Do. *Fast global image smoothing based on weighted least squares*. IEEE Transactions on Image Processing, vol. 23, no. 12, pages 5638–5653, 2014. (Cité en page 253.)

[Minervini 2011] Andrea Minervini, Vincenzo Ficarra, Francesco Rocco, Alessandro Antonelli, Roberto Bertini, Giorgio Carmignani, Sergio Cosciani Cunico, Dario Fontana, Nicola Longo, Giuseppe Martorana*et al. Simple enucleation is equivalent to traditional partial nephrectomy for renal cell carcinoma : results of a nonrandomized, retrospective, comparative study*. The Journal of urology, vol. 185, no. 5, pages 1604–1610, 2011. (Cité en page 79.)

[Ministère de l'éducation nationale 2017] Ministère de l'éducation nationale. *Arrêté du 21 avril 2017 relatif aux connaissances, aux compétences et aux maquettes de formation des diplômes d'études spécialisées et fixant la liste de ces diplômes et des options et formations*

spécialisées transversales du troisième cycle des études de médecine. https ://www.legifrance.gouv.fr/jorf/id/JORFARTI000034502903, 2017. Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche. (Cité en page 4.)

[Ministère des solidarités et de la santé 2021] Ministère des solidarités et de la santé. *Arrêté du 25 septembre 2021 pris en application des articles R. 6152-301 et suivants du code de la santé publique et relatif à l'organisation des épreuves du concours de praticien hospitalier des établissements publics de santé*. https ://www.legifrance.gouv.fr/download/pdf ?id=r4UtSwLuyuyEpihQw3VNGt2zup93I6QTVI 2021. (Cité en page 4.)

[Mir 2015] Maria C Mir, Cesar Ercole, Toshio Takagi, Zhiling Zhang, Lily Velet, Erick M Remer, Sevag Demirjian et Steven C Campbell. *Decline in renal function after partial nephrectomy : etiology and prevention*. The Journal of urology, vol. 193, no. 6, pages 1889–1898, 2015. (Cité en page 81.)

[Mir 2017] Maria Carmen Mir, Ithaar Derweesh, Francesco Porpiglia, Homayoun Zargar, Alexandre Mottrie et Riccardo Autorino. *Partial nephrectomy versus radical nephrectomy for clinical T1b and T2 renal tumors : a systematic review and meta-analysis of comparative studies*. European urology, vol. 71, no. 4, pages 606–617, 2017. (Cité en page 75.)

[Mir 2018] Maria Carmen Mir, Umberto Capitanio, Riccardo Bertolo, Idir Ouzaid, Maciej Salagierski, Maximilian Kriegmair, Alessandro Volpe, Michael AS Jewett, Alexander Kutikov, Phillip M Pierorazio*et al. Role of active surveillance for localized small renal masses*. European urology oncology, vol. 1, no. 3, pages 177–187, 2018. (Cité en page 75.)

[Miranda-Luna 2008] Rosebet Miranda-Luna, Christian Daul, Walter CPM Blondel, Yahir Hernandez-Mier, Didier Wolf et François Guillemin. *Mosaicing of bladder endoscopic image sequences : Distortion calibration and registration algorithm*. IEEE Transactions on Biomedical Engineering, vol. 55, no. 2, pages 541–553, 2008. (Cité en page 221.)

[Mitsui 2012] Yozo Mitsui, Hiroaki Shiina, Naoko Arichi, Takeo Hiraoka, Shogo Inoue, Masahiro Sumura, Satoshi Honda, Hiroaki Yasumoto et Mikio Igawa. *Indocyanine green (ICG)-based fluorescence navigation system for discrimination of kidney cancer from normal parenchyma : application during partial nephrectomy*. International urology and nephrology, vol. 44, pages 753–759, 2012. (Cité en page 48.)

[Moch 2022] Holger Moch, Mahul B Amin, Daniel M Berney, Eva M Compérat, Anthony J Gill, Arndt Hartmann, Santosh Menon, Maria R Raspollini, Mark A Rubin, John R Srigley*et al. The 2022 World Health Organization classification of tumours of the urinary system and male genital organs—part A : renal, penile, and testicular tumours*. European urology, vol. 82, no. 5, pages 458–468, 2022. (Cité en page 73.)

[Modrzejewski 2019] Richard Modrzejewski, Toby Collins, Barbara Seeliger, Adrien Bartoli, Alexandre Hostettler et Jacques Marescaux. *An in vivo porcine dataset and evaluation methodology to measure soft-body laparoscopic liver registration accuracy with an extended algorithm that handles collisions.* International journal of computer assisted radiology and surgery, vol. 14, pages 1237–1245, 2019. (Cité en page 172.)

[Moldovanu 2024] Claudia-Gabriela Moldovanu. *Virtual and augmented reality systems and three-dimensional printing of the renal model—novel trends to guide preoperative planning for renal cancer.* Asian Journal of Urology, 2024. (Cité en pages 45, 46 et 50.)

[Molle 2023] Fernando Molle, Maria Cristina Savastano, Federico Giannuzzi, Claudia Fossataro, Davide Brando, Andrea Molle, Maria Teresa Rebecchi, Benedetto Falsini, Roberta Mattei, Giorgia Mirisola*et al. 3D Da Vinci robotic surgery : is it a risk to the surgeon's eye health ?* Journal of Robotic Surgery, vol. 17, no. 5, pages 1995–2000, 2023. (Cité en page 131.)

[Moore 2016] Steven C Moore, I-Min Lee, Elisabete Weiderpass, Peter T Campbell, Joshua N Sampson, Cari M Kitahara, Sarah K Keadle, Hannah Arem, Amy Berrington De Gonzalez, Patricia Hartge*et al. Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults.* JAMA internal medicine, vol. 176, no. 6, pages 816–825, 2016. (Cité en page 69.)

[Morgand 2014] Alexandre Morgand et Mohamed Tamaazousti. *Generic and real-time detection of specular reflections in images.* In 2014 International conference on computer vision theory and applications (VISAPP), volume 1, pages 274–282. IEEE, 2014. (Cité en page 132.)

[Mori 2020] Junichi Mori. *Stereo camera calibration method and image processing device for stereo camera*, 2020. (Cité en page 37.)

[Morrissette 2015] Tyler J. Morrissette, Tao Zhao, III Joseph P. Orban et Brian David Hoffman. *Three-dimensional target devices, assemblies and methods for calibrating an endoscopic camera*, 2015. (Cité en page 37.)

[Morteza Safdarnejad 2017] Seyed Morteza Safdarnejad et Xiaoming Liu. *Spatio-Temporal Alignment of Non-Overlapping Sequences From Independently Panning Cameras.* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3808–3816, 2017. (Cité en page 129.)

[Mountney 2010] Peter Mountney, Danail Stoyanov et Guang-Zhong Yang. *Three-dimensional tissue deformation recovery and tracking.* IEEE Signal Processing Magazine, vol. 27, no. 4, pages 14–24, 2010. (Cité en page 226.)

[Mueller-Lisse 2010] Ullrich G Mueller-Lisse et Ulrike L Mueller-Lisse. *Imaging of advanced renal cell carcinoma.* World journal of urology, vol. 28, pages 253–261, 2010. (Cité en page 71.)

[Muglia 2015] Valdair F Muglia et Adilson Prando. *Renal cell carcinoma : histological classification and correlation with imaging findings.* Radiologia brasileira, vol. 48, pages 166–174, 2015. (Cité en page 69.)

[Mühlbauer 2021] Julia Mühlbauer, Luisa Egen, Karl-Friedrich Kowalewski, Maurizio Grilli, Margarete T Walach, Niklas Westhoff, Philipp Nuhn, Fabian C Laqua, Bettina Baessler et Maximilian C Kriegmair. *Radiomics in renal cell carcinoma—a systematic review and meta-analysis.* Cancers, vol. 13, no. 6, page 1348, 2021. (Cité en page 72.)

[Mullen 2008] Patrick Mullen, Yiying Tong, Pierre Alliez et Mathieu Desbrun. *Spectral conformal parameterization.* In Computer Graphics Forum, volume 27, pages 1487–1494. Wiley Online Library, 2008. (Cité en pages 231 et 253.)

[Müller 2007] Matthias Müller, Bruno Heidelberger, Marcus Hennix et John Ratcliff. *Position based dynamics.* Journal of Visual Communication and Image Representation, vol. 18, no. 2, pages 109–118, 2007. (Cité en page 175.)

[Müller 2021] Philip C Müller, Caroline Haslebacher, Daniel C Steinemann, Beat P Müller-Stich, Thilo Hackert, Matthias Peterhans et Benjamin Eigl. *Image-guided minimally invasive endopancreatic surgery using a computer-assisted navigation system.* Surgical endoscopy, vol. 35, pages 1610–1617, 2021. (Cité en page 176.)

[Mur-Artal 2015] Raul Mur-Artal, Jose Maria Martinez Montiel et Juan D Tardos. *ORB-SLAM : A versatile and accurate monocular SLAM system.* IEEE transactions on robotics, vol. 31, no. 5, pages 1147–1163, 2015. (Cité en pages 99 et 222.)

[Murai 2025] Riku Murai, Eric Dexheimer et Andrew J Davison. *MASt3R-SLAM : Real-time dense SLAM with 3D reconstruction priors.* In Proceedings of the Computer Vision and Pattern Recognition Conference, pages 16695–16705, 2025. (Cité en page 190.)

[Myronenko 2010] Andriy Myronenko et Xubo Song. *Point set registration : Coherent point drift.* IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 12, pages 2262–2275, 2010. (Cité en page 97.)

[Nagaya 2017] Tadanobu Nagaya, Yu A Nakamura, Peter L Choyke et Hisataka Kobayashi. *Fluorescence-guided surgery.* Frontiers in oncology, vol. 7, page 314, 2017. (Cité en page 48.)

[Nagel 2021] Markus Nagel, Marios Fournarakis, Rana Ali Amjad, Yelysei Bondarenko, Mart Van Baalen et Tijmen Blankevoort. *A white paper on neural network quantization.* arXiv preprint arXiv :2106.08295, 2021. (Cité en page 243.)

[Nakamura 2010] Kazuhiro Nakamura, Yukio Naya, Satoki Zenbutsu, Kazuhiro Araki, Shuko Cho, Sho Ohta, Naoki Nihei, Hiroyoshi Suzuki, Tomohiko Ichikawa et Tatsuo Igarashi. *Surgical navigation using three-dimensional computed tomography images fused intraoperatively with live video.* Journal of endourology, vol. 24, no. 4, pages 521–524, 2010. (Cité en pages 94 et 176.)

[Nasab 2019] Mohammad Hossein Abedin Nasab. Handbook of robotic and image-guided surgery. Elsevier, 2019. (Cité en page 38.)

[Netter 2019] Frank H Netter. Atlas d'anatomie humaine. Elsevier Health Sciences, 2019. (Cité en pages 57, 60, 61, 62, 63 et 64.)

[Ngo 2024] Tuan Duc Ngo, Peiye Zhuang, Chuang Gan, Evangelos Kalogerakis, Sergey Tulyakov, Hsin-Ying Lee et Chaoyang Wang. *DELTA : Dense Efficient Long-range 3D Tracking for any video.* arXiv preprint arXiv :2410.24211, 2024. (Cité en page 224.)

[Ngu 2017] James Chi-Yong Ngu, Charles Bih-Shiou Tsang et Dean Chi-Siong Koh. *The da Vinci Xi : a review of its capabilities, versatility, and potential role in robotic colorectal surgery.* Robotic Surgery : Research and Reviews, pages 77–85, 2017. (Cité en page 29.)

[Ni 2009] Kai Ni, Hailin Jin et Frank Dellaert. *GroupSAC : Efficient consensus in the presence of groupings.* In 2009 IEEE 12th International Conference on Computer Vision, pages 2193–2200. IEEE, 2009. (Cité en page 218.)

[Ni 2022] Yixiu Ni et Xiaohua Yang. *A systematic review and meta-analysis of comparison of outcomes of robot-assisted versus open partial nephrectomy in clinical T1 renal cell carcinoma patients.* Urologia Internationalis, vol. 106, no. 8, pages 757–767, 2022. (Cité en page 81.)

[Ni 2023] Junjie Ni, Yijin Li, Zhaoyang Huang, Hongsheng Li, Hujun Bao, Zhaopeng Cui et Guofeng Zhang. *Pats : Patch area transportation with subdivision for local feature matching.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 17776–17786, 2023. (Cité en page 146.)

[Nicodemus 1977] Fred Edwin Nicodemus, Joseph C Richmond, Jack J Hsia, Irving W Ginsberg, Thomas Limperis*et al.* Geometrical considerations and nomenclature for reflectance, volume 160. US Department of Commerce, National Bureau of Standards Washington, DC, USA, 1977. (Cité en page 123.)

[Niemeyer 1999] Gunter D. Niemeyer, William C. Nowlin et Gary S. Guthart. *Alignment of master and slave in a minimally invasive surgical apparatus*, 1999. (Cité en pages 31 et 32.)

[Niemeyer 2005] Gunter D. Niemeyer, Gary S. Guthart, William C. Nowlin, Nitish Swarup, Gregory K Toth et Robert G. Younge. *Camera referenced control in a minimally invasive surgical apparatus*, 2005. (Cité en page 35.)

[Nimmagadda 2022] Naren Nimmagadda, James M Ferguson, Nicholas L Kavoussi, Bryn Pitt, Eric J Barth, Josephine Granna, Robert J Webster III et S Duke Herrell III. *Patient-specific, touch-based registration during robotic, image-guided partial nephrectomy.* World journal of urology, vol. 40, no. 3, pages 671–677, 2022. (Cité en page 100.)

[Ntourakis 2016] Dimitrios Ntourakis, Ricardo Memeo, Luc Soler, Jacques Marescaux, Didier Mutter et Patrick Pessaux. *Augmented reality guidance for the*

resection of missing colorectal liver metastases : an initial experience. World journal of surgery, vol. 40, pages 419–426, 2016. (Cité en page 171.)

[Oberweger 2015] Markus Oberweger, Paul Wohlhart et Vincent Lepetit. *Training a feedback loop for hand pose estimation*. In Proceedings of the IEEE international conference on computer vision, pages 3316–3324, 2015. (Cité en page 202.)

[Okamoto 2015] Tomoyoshi Okamoto, Shinji Onda, Jungo Yasuda, Katsuhiko Yanaga, Naoki Suzuki et Asaki Hattori. *Navigation surgery using an augmented reality for pancreatectomy*. Digestive surgery, vol. 32, no. 2, pages 117–123, 2015. (Cité en page 171.)

[Okhunov 2021] Zhamshid Okhunov, Michael A Gorin, Francis A Jefferson, Andrew Shea Afyouni, Mohamad E Allaf, Phillip M Pierorazio, Roshan M Patel, Linda M Huynh, Shlomi Tapiero, Kathryn Osann*et al. Can preoperative renal mass biopsy change clinical practice and reduce surgical intervention for small renal masses ?* In Urologic Oncology : Seminars and Original Investigations, volume 39, pages 735–e17. Elsevier, 2021. (Cité en page 72.)

[Okunrintemi 2016] Victor Okunrintemi, Faiz Gani et Timothy M Pawlik. *National trends in postoperative outcomes and cost comparing minimally invasive versus open liver and pancreatic surgery*. Journal of Gastrointestinal Surgery, vol. 20, no. 11, pages 1836–1843, 2016. (Cité en page 24.)

[Oliva 2001] Aude Oliva et Antonio Torralba. *Modeling the shape of the scene : A holistic representation of the spatial envelope*. International journal of computer vision, vol. 42, pages 145–175, 2001. (Cité en page 146.)

[Onda 2013] Shinji Onda, Tomoyoshi Okamoto, Masaru Kanehira, Shuichi Fujioka, Naoki Suzuki, Asaki Hattori et Katsuhiko Yanaga. *Short rigid scope and stereo-scope designed specifically for open abdominal navigation surgery : clinical application for hepatobiliary and pancreatic surgery*. Journal of hepatobiliary-pancreatic sciences, vol. 20, pages 448–453, 2013. (Cité en page 172.)

[ONNX developers 2021] ONNX developers. *ONNX Runtime*. https ://onnxruntime.ai/, 2021. Version : 1.22.2. (Cité en page 243.)

[Oquab 2023] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby*et al. Dinov2 : Learning robust visual features without supervision*. arXiv preprint arXiv :2304.07193, 2023. (Cité en pages 202 et 225.)

[Orr 1996] Mark JL Orr*et al. Introduction to radial basis function networks*, 1996. (Cité en page 97.)

[Ortiz 2005] Francisco Ortiz et Fernando Torres. *A new inpainting method for highlights elimination by colour morphology*. In International Conference on Pattern Recognition and Image Analysis, pages 368–376. Springer, 2005. (Cité en page 132.)

[Ortiz 2006] Francisco Ortiz et Fernando Torres. *Automatic detection and elimination of specular reflectance in color images by means of ms diagram and vector connected filters.* IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 36, no. 5, pages 681–687, 2006. (Cité en page 132.)

[Ortiz 2023] Lizmarie Comenencia Ortiz, David I. Moreira Ridsdale, Ashwinram Suresh, Kyle W. Clements, Alan W. Petersen et Cheng-Ling Chang. *Force sensing medical instrument*, 2023. (Cité en page 37.)

[Özgür 2018] Erol Özgür, Bongjin Koo, Bertrand Le Roy, Emmanuel Buc et Adrien Bartoli. *Preoperative liver registration for augmented monocular laparoscopy using backward–forward biomechanical simulation.* International journal of computer assisted radiology and surgery, vol. 13, pages 1629–1640, 2018. (Cité en page 174.)

[Özyeşil 2017] Onur Özyeşil, Vladislav Voroninski, Ronen Basri et Amit Singer. *A survey of structure from motion*.* Acta Numerica, vol. 26, pages 305–364, 2017. (Cité en page 171.)

[Ozyoruk 2021] Kutsev Bengisu Ozyoruk, Guliz Irem Gokceler, Taylor L Bobrow, Gulfize Coskun, Kagan Incetan, Yasin Almalioglu, Faisal Mahmood, Eva Curto, Luis Perdigoto, Marina Oliveira*et al.* *EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos.* Medical image analysis, vol. 71, page 102058, 2021. (Cité en page 226.)

[Padovan 2022] Erica Padovan, Giorgia Marullo, Leonardo Tanzi, Pietro Piazzolla, Sandro Moos, Francesco Porpiglia et Enrico Vezzetti. *A deep learning framework for real-time 3D model registration in robot-assisted laparoscopic surgery.* The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 18, no. 3, page e2387, 2022. (Cité en pages 98, 99 et 178.)

[Palmer 1962] Raoul Palmer. *Essais de sterilization tubaire coelioscopique par electrocoagulation isthmique.* Bull Fed Soc Gynecol Obstet Lang Fr, vol. 14, page 298, 1962. (Cité en page 23.)

[Pan 2019] Heng-Chih Pan, Chiao-Yin Sun, I-Wen Wu, Tien-Ling Tsai, Chi-Chin Sun et Chin-Chan Lee. *Higher risk of malignant neoplasms in young adults with end-stage renal disease receiving haemodialysis : A nationwide population-based study.* Nephrology, vol. 24, no. 11, pages 1165–1171, 2019. (Cité en page 69.)

[Pantelidou 2016] Maria Pantelidou, Ben Challacombe, Andrew McGrath, Matthew Brown, Shahzad Ilyas, Konstantinos Katsanos et Andreas Adam. *Percutaneous radiofrequency ablation versus robotic-assisted partial nephrectomy for the treatment of small renal cell carcinoma.* Cardiovascular and Interventional Radiology, vol. 39, pages 1595–1603, 2016. (Cité en page 74.)

[Parker 2013] Patricia A Parker, Frances Alba, Bryan Fellman, Diana L Urbauer, Yisheng Li, Jose A Karam, Nizar Tannir, Eric Jonasch, Christopher G Wood

et Surena F Matin. *Illness uncertainty and quality of life of patients with small renal tumors undergoing watchful waiting : a 2-year prospective study.* European urology, vol. 63, no. 6, pages 1122–1127, 2013. (Cité en page 75.)

[Parlement Européen 2005] Parlement Européen. *Directive 2005/36/EC of the European Parliament and of the Council on the Recognition of Professional Qualifications*, 2005. (Cité en page 4.)

[Patat 2025] F. Patat et M. Ronot. *Échographie.* In Les fondamentaux de l'imagerie médicale. Collège National des Enseignants de Biophysique et de Médecine Nucléaire, 2025. (Cité en pages 47 et 48.)

[Patel 2017] Hiten D Patel, Phillip M Pierorazio, Michael H Johnson, Ritu Sharma, Emmanuel Iyoha, Mohamad E Allaf, Eric B Bass et Stephen M Sozio. *Renal functional outcomes after surgery, ablation, and active surveillance of localized renal tumors : a systematic review and meta-analysis.* Clinical Journal of the American Society of Nephrology, vol. 12, no. 7, pages 1057–1069, 2017. (Cité en page 74.)

[Patel 2020a] Hiten D Patel, Paige E Nichols, Zhuo Tony Su, Mohit Gupta, Joseph G Cheaib, Mohamad E Allaf et Phillip M Pierorazio. *Renal mass biopsy is associated with reduction in surgery for early-stage kidney cancer.* Urology, vol. 135, pages 76–81, 2020. (Cité en page 72.)

[Patel 2020b] Sunil H Patel, Robert G Uzzo, Alessandro Larcher, Benoit Peyronnet, Brian R Lane, Deepak Pruthi, Madhumitha Reddy, Umberto Capitanio, Shreyas Joshi, Sabrina Noyes*et al. Oncologic and functional outcomes of radical and partial nephrectomy in pT3a pathologically upstaged renal cell carcinoma : A multi-institutional analysis.* Clinical Genitourinary Cancer, vol. 18, no. 6, pages e723–e729, 2020. (Cité en page 76.)

[Patel 2021] Amit K Patel, Brian R Lane, Prateek Chintalapati, Lina Fouad, Mohit Butaney, Jeffrey Budzyn, Anna Johnson, Ji Qi, Edward Schervish et Craig G Rogers. *Utilization of renal mass biopsy for T1 renal lesions across Michigan : results from MUSIC-KIDNEY, a statewide quality improvement collaborative.* European Urology Open Science, vol. 30, pages 37–43, 2021. (Cité en page 72.)

[Pătrăucean 2024] Viorica Pătrăucean, Xu Owen He, Joseph Heyward, Chuhan Zhang, Mehdi SM Sajjadi, George-Cristian Muraru, Artem Zholus, Mahdi Karami, Ross Goroshin, Yutian Chen*et al. TRecViT : A Recurrent Video Transformer.* arXiv preprint arXiv :2412.14294, 2024. (Cité en page 225.)

[Patwardhan 2017] Viraj A. Patwardhan, John A Barton, Mathew Clopp et John E. Sell. *Compact binocular image capture device*, 2017. (Cité en page 36.)

[Pautrat 2023] Rémi Pautrat, Iago Suárez, Yifan Yu, Marc Pollefeys et Viktor Larsson. *Gluestick : Robust image matching by sticking points and lines together.* In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9706–9716, 2023. (Cité en page 146.)

[Pecoraro 2019] Angela Pecoraro, Carlotta Palumbo, Sophie Knipper, Francesco A Mistretta, Zhe Tian, Shahrokh F Shariat, Fred Saad, Alberto Briganti, Cristian Fiori, Francesco Porpiglia*et al.* *Cryoablation predisposes to higher cancer specific mortality relative to partial nephrectomy in patients with nonmetastatic pT1b kidney cancer.* The Journal of Urology, vol. 202, no. 6, pages 1120–1126, 2019. (Cité en page 77.)

[Pecoraro 2024] Angela Pecoraro, Federico Piramide et Daniele Amparore. *3D augmented reality-guided robotic partial nephrectomy.* Mini-invasive Surgery, vol. 8, pages N–A, 2024. (Cité en pages 98 et 99.)

[Peng 2010] En Peng et Ling Li. *Camera calibration using one-dimensional information and its applications in both controlled and uncontrolled environments.* Pattern Recognition, vol. 43, no. 3, pages 1188–1198, 2010. (Cité en page 107.)

[Pessaux 2015] Patrick Pessaux, Michele Diana, Luc Soler, Tullio Piardi, Didier Mutter et Jacques Marescaux. *Towards cybernetic surgery : robotic and augmented reality-assisted liver segmentectomy.* Langenbeck's archives of surgery, vol. 400, pages 381–385, 2015. (Cité en page 171.)

[Petros 2019] Firas G Petros, Aradhana M Venkatesan, Diana Kaya, Chaan S Ng, Bryan M Fellman, Jose A Karam, Christopher G Wood et Surena F Matin. *Conditional survival of patients with small renal masses undergoing active surveillance.* BJU international, vol. 123, no. 3, pages 447–455, 2019. (Cité en page 75.)

[Peyronnet 2016] Benoit Peyronnet, Thomas Seisen, Emmanuel Oger, Christophe Vaessen, Yohann Grassano, Thibaut Benoit, Julie Carrouget, Benjamin Pradère, Zineddine Khene, Anthony Giwerc*et al.* *Comparison of 1800 robotic and open partial nephrectomies for renal tumors.* Annals of surgical oncology, vol. 23, pages 4277–4283, 2016. (Cité en page 81.)

[Peyronnet 2018] Benoît Peyronnet, Lauranne Tondut, Jean-Christophe Bernhard, Christophe Vaessen, Nicolas Doumerc, Philippe Sebe, Benjamin Pradere, Bertrand Guillonneau, Zine-Eddine Khene, Francois-Xavier Nouhaud*et al.* *Impact of hospital volume and surgeon volume on robot-assisted partial nephrectomy outcomes : a multicentre study.* BJU international, vol. 121, no. 6, pages 916–922, 2018. (Cité en page 82.)

[Piana 2024] Alberto Piana, Daniele Amparore, Michele Sica, Gabriele Volpi, Enrico Checcucci, Federico Piramide, Sabrina De Cillis, Giovanni Busacca, Gianluca Scarpelli, Flavio Sidoti, Stefano Alba, Pietro Piazzolla, Cristian Fiori, Francesco Porpiglia et Michele Di Dio. *Automatic 3D augmented-reality robot-assisted partial nephrectomy using machine learning : our pioneer experience.* Cancers, vol. 16, no. 5, page 1047, 2024. (Cité en page 177.)

[Piazzolla 2025] P Piazzolla, G Volpi, A Piana, E Checcucci, D Amparore, F Piramide, S De Cillis, P Verri, G Mesterca, A Bellin*et al.* *Artificial intelligence guidance for 3D augmented reality robotic surgery : When the machine falls*

the human assistance is still alive. Urology Video Journal, vol. 25, page 100307, 2025. (Cité en page 98.)

[Piccinelli 2025] Luigi Piccinelli, Christos Sakaridis, Yung-Hsu Yang, Mattia Segu, Siyuan Li, Wim Abbeloos et Luc Van Gool. *UniDepthV2 : Universal monocular metric depth estimation made simpler.* arXiv preprint arXiv :2502.20110, 2025. (Cité en page 109.)

[Pierorazio 2013] Phillip Pierorazio, James McKiernan et Mohamad Allaf. *633 quality of life on active surveillance for a small renal masses versus immediate intervention : interim analysis of the DISSRM (DELAYED INTERVENTION AND SURVEILLANCE FOR SMALL RENAL MASSES) registry.* The Journal of Urology, vol. 189, no. 4S, pages e259–e259, 2013. (Cité en page 75.)

[Pierorazio 2016] Phillip M Pierorazio, Michael H Johnson, Hiten D Patel, Stephen M Sozio, Ritu Sharma, Emmanuel Iyoha, Eric B Bass et Mohamad E Allaf. *Management of renal masses and localized renal cancer [Internet].* 2016. (Cité en page 72.)

[Pinto 2024] Paulo Victor Alves Pinto, Fernando Morbeck Almeida Coelho, Alice Schuch, Mauricio Zapparoli et Ronaldo Hueb Baroni. *Pre-operative imaging evaluation of renal cell carcinoma.* Revista da Associação Médica Brasileira, vol. 70, no. suppl 1, page e2024S107, 2024. (Cité en page 72.)

[Piramide 2022] Federico Piramide, Daniele Amparore, Angela Pecoraro, Paolo Verri, Enrico Checcucci, Sabrina De Cillis, Alberto Piana, Matteo Manfredi, Cristian Fiori et Francesco Porpiglia. *Augmented reality 3D robot-assisted partial nephrectomy : Tips and tricks to improve surgical strategies and outcomes.* Urology Video Journal, vol. 13, page 100137, 2022. (Cité en page 94.)

[Pizarro 2012] Daniel Pizarro et Adrien Bartoli. *Feature-based deformable surface detection with self-occlusion reasoning.* International journal of computer vision, vol. 97, pages 54–70, 2012. (Cité en page 218.)

[Plewes 2012] Donald B Plewes et Walter Kucharczyk. *Physics of MRI : a primer.* Journal of magnetic resonance imaging, vol. 35, no. 5, pages 1038–1054, 2012. (Cité en page 43.)

[Plotz 2012] Thomas Plotz, Chen Chen, Nils Y Hammerla et Gregory D Abowd. *Automatic synchronization of wearable sensors and video-cameras for ground truth annotation–A practical approach.* In 2012 16th international symposium on wearable computers, pages 100–103. IEEE, 2012. (Cité en page 129.)

[Plötz 2018] Tobias Plötz et Stefan Roth. *Neural nearest neighbors networks.* Advances in Neural information processing systems, vol. 31, 2018. (Cité en page 219.)

[Polok 2015] Lukáš Polok, Lukáš Klicnar, Vítězslav Beran, Pavel Smrž et Pavel Zemčík. *Quality assurance in large collections of video sequences.* In 2015 IEEE International Conference on Image Processing (ICIP), pages 3580–3584. IEEE, 2015. (Cité en page 128.)

[Pontana 2025] F. Pontana, B. Longère et A. Lalande. *Tomodensitométrie (TDM).* In Les fondamentaux de l'imagerie médicale. Collège National des Enseignants de Biophysique et de Médecine Nucléaire, 2025. (Cité en pages 40, 41 et 42.)

[Pooley 2003] Daniel W Pooley, Michael J Brooks, AJ Van Den Hengel et Wojciech Chojnacki. *A voting scheme for estimating the synchrony of moving-camera videos.* In Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429), volume 1, pages I–413. IEEE, 2003. (Cité en page 128.)

[Porpiglia 2018a] Francesco Porpiglia, Daniele Amparore, Enrico Checcucci, Riccardo Autorino, Matteo Manfredi, Giada Iannizzi, Cristian Fiori, ESUT Research Group*et al.* *Current use of three-dimensional model technology in urology : a road map for personalised surgical planning.* European urology focus, vol. 4, no. 5, pages 652–656, 2018. (Cité en page 94.)

[Porpiglia 2018b] Francesco Porpiglia, Cristian Fiori, Enrico Checcucci, Daniele Amparore et Riccardo Bertolo. *Hyperaccuracy three-dimensional reconstruction is able to maximize the efficacy of selective clamping during robot-assisted partial nephrectomy for complex renal masses.* European urology, vol. 74, no. 5, pages 651–660, 2018. (Cité en pages 94 et 176.)

[Porpiglia 2019] Francesco Porpiglia, Enrico Checcucci, Daniele Amparore, Matteo Manfredi, Federica Massa, Pietro Piazzolla, Diego Manfrin, Alberto Piana, Daniele Tota, Enrico Bollito*et al.* *Three-dimensional elastic augmented-reality robot-assisted radical prostatectomy using hyperaccuracy three-dimensional reconstruction technology : a step further in the identification of capsular involvement.* European urology, vol. 76, no. 4, pages 505–514, 2019. (Cité en page 171.)

[Porpiglia 2020] Francesco Porpiglia, Enrico Checcucci, Daniele Amparore, Federico Piramide, Gabriele Volpi, Stefano Granato, Paolo Verri, Matteo Manfredi, Andrea Bellin, Pietro Piazzolla*et al.* *Three-dimensional augmented reality robot-assisted partial nephrectomy in case of complex tumours (PADUA superior to 10) : a new intraoperative tool overcoming the ultrasound guidance.* European urology, vol. 78, no. 2, pages 229–238, 2020. (Cité en pages 94, 95, 100 et 176.)

[Porpiglia 2021] Francesco Porpiglia, Andrea Mari, Daniele Amparore, Cristian Fiori, Alessandro Antonelli, Walter Artibani, Pierluigi Bove, Eugenio Brunocilla, Umberto Capitanio, Luigi Da Pozzo*et al.* *Transperitoneal vs retroperitoneal minimally invasive partial nephrectomy : comparison of perioperative outcomes and functional follow-up in a large multi-institutional cohort (The RECORD 2 Project).* Surgical Endoscopy, vol. 35, pages 4295–4304, 2021. (Cité en pages 76 et 81.)

[Posielski 2019] Natasza M Posielski, Anthony Bui, Shane A Wells, Sara L Best, Lori Mankowski Gettle, Timothy J Ziemlewicz, Meghan G Lubner, J Louis

Hinshaw, Fred T Lee Jr, Glenn O Allen*et al.* *Risk factors for complications and nondiagnostic results following 1,155 consecutive percutaneous core renal mass biopsies.* The Journal of Urology, vol. 201, no. 6, pages 1080–1087, 2019. (Cité en page 72.)

[Potje 2023] Guilherme Potje, Felipe Cadar, André Araujo, Renato Martins et Erickson R Nascimento. *Enhancing deformable local features by jointly learning to detect and describe keypoints.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1306–1315, 2023. (Cité en page 217.)

[Potje 2024] Guilherme Potje, Felipe Cadar, André Araujo, Renato Martins et Erickson R Nascimento. *Xfeat : Accelerated features for lightweight image matching.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 2682–2691, 2024. (Cité en page 146.)

[Potretzke 2016] Aaron M Potretzke, Brent Alexander Knight, Homayoun Zargar, Jihad H Kaouk, Ravi Barod, Craig G Rogers, Alon Mass, Michael D Stifelman, Michael H Johnson, Mohamad E Allaf*et al.* *Urinary fistula after robot-assisted partial nephrectomy : a multicentre analysis of 1 791 patients.* BJU international, vol. 117, no. 1, pages 131–137, 2016. (Cité en page 81.)

[Pratt 2010] Philip Pratt, Danail Stoyanov, Marco Visentini-Scarzanella et Guang-Zhong Yang. *Dynamic guidance for robotic surgery using image-constrained biomechanical models.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 77–85. Springer, 2010. (Cité en page 226.)

[Pratt 2012] Philip Pratt, Erik Mayer, Justin Vale, Daniel Cohen, Eddie Edwards, Ara Darzi et Guang-Zhong Yang. *An effective visualisation and registration system for image-guided robotic partial nephrectomy.* Journal of Robotic Surgery, vol. 6, pages 23–31, 2012. (Cité en pages 95 et 176.)

[Prevost 2020] Gian Andrea Prevost, Benjamin Eigl, Iwan Paolucci, Tobias Rudolph, Matthias Peterhans, Stefan Weber, Guido Beldi, Daniel Candinas et Anja Lachenmayer. *Efficiency, accuracy and clinical applicability of a new image-guided surgery system in 3D laparoscopic liver surgery.* Journal of gastrointestinal surgery, vol. 24, no. 10, pages 2251–2258, 2020. (Cité en page 171.)

[Pribanic 2015] Tomislav Pribanic, Marko Lelas et Igor Krois. *Sequence-to-sequence alignment using a pendulum.* IET computer vision, vol. 9, no. 4, pages 570–575, 2015. (Cité en page 127.)

[Prisco 2010] Giuseppe Prisco et David J Rosa. *Robotic surgical system with joint motion controller adapted to reduce instrument tip vibrations*, 2010. (Cité en page 35.)

[Puliatti 2022] Stefano Puliatti, Ahmed Eissa, Enrico Checcucci, Pietro Piazza, Marco Amato, Stefania Ferretti, Simone Scarcella, Juan Gomez Rivas, Mark Taratkin, Josè Marenco*et al.* *New imaging technologies for robotic kidney*

cancer surgery. Asian Journal of Urology, vol. 9, no. 3, pages 253–262, 2022. (Cité en pages 87 et 94.)

[Purushwalkam 2020] Senthil Purushwalkam, Tian Ye, Saurabh Gupta et Abhinav Gupta. *Aligning videos in space and time*. In Computer Vision–ECCV 2020 : 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVI 16, pages 262–278. Springer, 2020. (Cité en page 129.)

[Qi 2017] Charles Ruizhongtai Qi, Li Yi, Hao Su et Leonidas J Guibas. *Pointnet++ : Deep hierarchical feature learning on point sets in a metric space*. Advances in neural information processing systems, vol. 30, 2017. (Cité en page 98.)

[Rabindran 2024] Dinesh Rabindran, Ryan C. Abbott, Daniel H. Gomez et John Ryan Steger. *Systems and methods for control of end effectors*, 2024. (Cité en page 37.)

[Radford 2021] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark*et al.* *Learning transferable visual models from natural language supervision*. In International conference on machine learning, pages 8748–8763. PmLR, 2021. (Cité en page 225.)

[Raguram 2012] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas et Jan-Michael Frahm. *USAC : A universal framework for random sample consensus*. IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 8, pages 2022–2038, 2012. (Cité en page 218.)

[Raguse 2009] Karsten Raguse et Christian Heipke. *Synchronization of image sequences-a photogrammetric method*. Photogrammetric Engineering & Remote Sensing, vol. 75, no. 5, pages 535–546, 2009. (Cité en page 129.)

[Rais-Bahrami 2012] Soroush Rais-Bahrami, Arvin K George, Amin S Herati, Arun K Srinivasan, Lee Richstone et Louis R Kavoussi. *Off-clamp versus complete hilar control laparoscopic partial nephrectomy : comparison by clinical stage*. BJU international, vol. 109, no. 9, 2012. (Cité en page 81.)

[Rallis 2021] Kathrine S Rallis, Sam O Kleeman, Michael Grant, Katherine L Ordidge, Anju Sahdev et Thomas Powles. *Radiomics for renal cell carcinoma : Predicting outcomes from immunotherapy and targeted therapies—A narrative review*. European Urology Focus, vol. 7, no. 4, pages 717–721, 2021. (Cité en page 72.)

[Ramírez-Hernández 2020] Luis R Ramírez-Hernández, Julio C Rodríguez-Quinoñez, Moises J Castro-Toscano, Daniel Hernández-Balbuena, Wendy Flores-Fuentes, Raúl Rascón-Carmona, Lars Lindner et Oleg Sergiyenko. *Improve three-dimensional point localization accuracy in stereo vision systems using a novel camera calibration method*. International Journal of Advanced Robotic Systems, vol. 17, no. 1, page 1729881419896717, 2020. (Cité en page 106.)

[Rassweiler 2014] Jens Rassweiler, Marie-Claire Rassweiler, Michael Müller, Hannes Kenngott, Hans-Peter Meinzer, Dogu Teber, ESUT Expert Group*et al.* *Sur-*

gical navigation in urology : European perspective. Current opinion in urology, vol. 24, no. 1, pages 81–97, 2014. (Cité en page 94.)

[Rau 2019] Anita Rau, PJ Eddie Edwards, Omer F Ahmad, Paul Riordan, Mirek Janatka, Laurence B Lovat et Danail Stoyanov. *Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy.* International journal of computer assisted radiology and surgery, vol. 14, no. 7, pages 1167–1176, 2019. (Cité en page 226.)

[Rau 2023] Anita Rau, Binod Bhattarai, Lourdes Agapito et Danail Stoyanov. *Bimodal camera pose prediction for endoscopy.* IEEE Transactions on Medical Robotics and Bionics, vol. 5, no. 4, pages 978–989, 2023. (Cité en page 228.)

[Ravi 2024] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson*et al. Sam 2 : Segment anything in images and videos.* arXiv preprint arXiv :2408.00714, 2024. (Cité en page 238.)

[Ravikumar 2017] Nishant Ravikumar, Ali Gooya, Alejandro F Frangi et Zeike A Taylor. *Generalised coherent point drift for group-wise registration of multi-dimensional point sets.* In Medical Image Computing and Computer Assisted Intervention- MICCAI 2017 : 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20, pages 309–316. Springer, 2017. (Cité en page 172.)

[Reese 2013] Adam C Reese, Jared M Whitson et Maxwell V Meng. *Natural history of untreated renal cell carcinoma with venous tumor thrombus.* In Urologic Oncology : Seminars and Original Investigations, volume 31, pages 1305–1309. Elsevier, 2013. (Cité en page 78.)

[Reizine 2024] Edouard Reizine, Maxime Blain, Lorenzo Pescatori, Benjamin Longère, Alexandre Ingels, Wafa Boughamni, Mohamed Bouanane, Sébastien Mulé et Alain Luciani. *Applicability of Bosniak 2019 for renal mass classification on portal venous phase at the era of spectral CT imaging using rapid kV-switching dual-energy CT.* European Radiology, vol. 34, no. 3, pages 1816–1824, 2024. (Cité en page 71.)

[Revaud 2019] Jerome Revaud, Cesar De Souza, Martin Humenberger et Philippe Weinzaepfel. *R2d2 : Reliable and repeatable detector and descriptor.* Advances in neural information processing systems, vol. 32, 2019. (Cité en page 217.)

[Reynolds 2009] Douglas A Reynolds*et al. Gaussian mixture models.* Encyclopedia of biometrics, vol. 741, no. 659-663, page 3, 2009. (Cité en page 97.)

[Richa 2008] Rogério Richa, Philippe Poignet et Chao Liu. *Efficient 3D tracking for motion compensation in beating heart surgery.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 684–691. Springer, 2008. (Cité en page 221.)

[Richmond 2019] Geoffrey A. Richmond et Brian D. Hoffman. *Systems and methods for emulating far-range lighting for an operational scene illuminated by close-range light*, 2019. (Cité en page 37.)

[Richter 2021] Florian Richter, Shihao Shen, Fei Liu, Jingbin Huang, Emily K Funk, Ryan K Orosco et Michael C Yip. *Autonomous robotic suction to clear the surgical field for hemostasis using image-based blood flow detection.* IEEE Robotics and Automation Letters, vol. 6, no. 2, pages 1383–1390, 2021. (Cité en page 219.)

[Richter 2024] Alexander Richter, Till Steinmann, Jean-Claude Rosenthal et Stefan J Rupitsch. *Advances in real-time 3d reconstruction for medical endoscopy.* Journal of imaging, vol. 10, no. 5, page 120, 2024. (Cité en pages 171 et 191.)

[Ricolfe-Viala 2021] Carlos Ricolfe-Viala et Alicia Esparza. *The influence of autofocus lenses in the camera calibration process.* IEEE Transactions on Instrumentation and Measurement, vol. 70, pages 1–15, 2021. (Cité en page 107.)

[Robert 2023] Sébastien C Robert, Tyler Cossetto, Timothy L Miao, Katherine Li, Eric Habib, Valentin Mocanu, Greg Garvin, Roya Etemad-Rezai et Derek W Cool. *Complications after renal mass biopsy : frequency, nature, timing, and associated characteristics.* American Journal of Roentgenology, vol. 221, no. 3, pages 344–353, 2023. (Cité en page 72.)

[Roberts 2022] Sidney Roberts, Aditya Desai, Enrico Checcucci, Stefano Puliatti, Mark Taratkin, Karl-Friedrich Kowalewski, I Rivero, D Veneziano, R Autorino, F Porpiglia*et al.* *" Augmented reality" applications in urology : a systematic review.* Minerva urology and nephrology, vol. 74, no. 5, pages 528–537, 2022. (Cité en page 94.)

[Robinson 2004] TN Robinson et GV Stiegmann. *Minimally invasive surgery.* Endoscopy, vol. 36, no. 01, pages 48–51, 2004. (Cité en page 25.)

[Robu 2018] Maria R Robu, João Ramalhinho, Stephen Thompson, Kurinchi Gurusamy, Brian Davidson, David Hawkes, Danail Stoyanov et Matthew J Clarkson. *Global rigid registration of CT to video in laparoscopic liver surgery.* International Journal of Computer Assisted Radiology and Surgery, vol. 13, pages 947–956, 2018. (Cité en page 172.)

[Rocco 2018] Ignacio Rocco, Mircea Cimpoi, Relja Arandjelović, Akihiko Torii, Tomas Pajdla et Josef Sivic. *Neighbourhood consensus networks.* Advances in neural information processing systems, vol. 31, 2018. (Cité en page 219.)

[Rodríguez-Sánchez 2017] Antonio Rodríguez-Sánchez, Daly Chea, George Azzopardi et Sebastian Stabinger. *A deep learning approach for detecting and correcting highlights in endoscopic images.* In 2017 seventh international conference on image processing theory, tools and applications (ipta), pages 1–6. IEEE, 2017. (Cité en page 132.)

[Rodríguez 2022] Juan J Gómez Rodríguez, José MM Montiel et Juan D Tardós. *Tracking monocular camera pose and deformation for slam inside the human*

body. In 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 5278–5285. IEEE, 2022. (Cité en page 223.)

[Rokh 2023] Babak Rokh, Ali Azarpeyvand et Alireza Khanteymoori. *A comprehensive survey on model quantization for deep neural networks in image classification*. ACM Transactions on Intelligent Systems and Technology, vol. 14, no. 6, pages 1–50, 2023. (Cité en page 243.)

[Ronneberger 2015] Olaf Ronneberger, Philipp Fischer et Thomas Brox. *U-net : Convolutional networks for biomedical image segmentation*. In Medical image computing and computer-assisted intervention–MICCAI 2015 : 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, pages 234–241. Springer, 2015. (Cité en pages 132, 146, 240 et 262.)

[Rosenblatt 1958] Frank Rosenblatt. *The perceptron : a probabilistic model for information storage and organization in the brain*. Psychological review, vol. 65, no. 6, page 386, 1958. (Cité en page 98.)

[Rosevear 2012] Henry M Rosevear, Paul T Gellhaus, Andrew J Lightfoot, Timothy P Kresowik, Fadi N Joudi et Chad R Tracy. *Utility of the RENAL nephrometry scoring system in the real world : predicting surgeon operative preference and complication risk*. BJU international, vol. 109, no. 5, pages 700–705, 2012. (Cité en page 72.)

[Rossi 2018] Sabrina H Rossi, Tobias Klatte, Juliet Usher-Smith et Grant D Stewart. *Epidemiology and screening for renal cancer*. World journal of urology, vol. 36, pages 1341–1353, 2018. (Cité en page 70.)

[Rosten 2006] Edward Rosten et Tom Drummond. *Machine learning for high-speed corner detection*. In Computer Vision–ECCV 2006 : 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9, pages 430–443. Springer, 2006. (Cité en pages 99, 145 et 217.)

[Rublee 2011] Ethan Rublee, Vincent Rabaud, Kurt Konolige et Gary Bradski. *ORB : An efficient alternative to SIFT or SURF*. In 2011 International conference on computer vision, pages 2564–2571. Ieee, 2011. (Cité en pages 99, 146 et 222.)

[Ruppert 2012] Guilherme Cesar Soares Ruppert, Leonardo Oliveira Reis, Paulo Henrique Junqueira Amorim, Thiago Franco de Moraes et Jorge Vicente Lopes da Silva. *Touchless gesture user interface for interactive image visualization in urological surgery*. World journal of urology, vol. 30, pages 687–691, 2012. (Cité en page 95.)

[Sadoughi 2019] Nima Sadoughi, Satheesh Krishna, David B Macdonald, Robert Chatelain, Trevor A Flood, Matthew DF McInnes et Nicola Schieda. *Diagnostic accuracy of attenuation difference and iodine concentration thresholds at rapid-kilovoltage-switching dual-energy CT for detection of enhancement in renal masses*. American Journal of Roentgenology, vol. 213, no. 3, pages 619–625, 2019. (Cité en page 71.)

[Saint-Pierre 2011] Charles-Auguste Saint-Pierre, Jonathan Boisvert, Guy Grimard et Farida Cheriet. *Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images*. Machine Vision and Applications, vol. 22, pages 171–180, 2011. (Cité en page 132.)

[Saklad 1941] Meyer Saklad. *Grading of patients for surgical procedures*. In The Journal of the American Society of Anesthesiologists, volume 2, pages 281–284. The American Society of Anesthesiologists, 1941. (Cité en page 16.)

[Sandler 2018] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov et Liang-Chieh Chen. *Mobilenetv2 : Inverted residuals and linear bottlenecks*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4510–4520, 2018. (Cité en pages 240 et 262.)

[Sanfilippo 2014] Kristen M Sanfilippo, Kathleen M McTigue, Christian J Fidler, James D Neaton, Yuefang Chang, Linda F Fried, Simin Liu et Lewis H Kuller. *Hypertension and obesity and the risk of kidney cancer in 2 large cohorts of US men and women*. Hypertension, vol. 63, no. 5, pages 934–941, 2014. (Cité en page 69.)

[Santé publique France 2023] Santé publique France. *Principaux résultats de l'enquête nationale de prévalence 2022 des infections nosocomiales et des traitements anti-infectieux en établissement de santé*, Mai 2023. (Cité en page 12.)

[Sarlin 2020] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz et Andrew Rabinovich. *Superglue : Learning feature matching with graph neural networks*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 4938–4947, 2020. (Cité en pages 146, 217, 219 et 271.)

[Sattler 2009] Torsten Sattler, Bastian Leibe et Leif Kobbelt. *SCRAMSAC : Improving RANSAC's efficiency with a spatial consistency filter*. In 2009 IEEE 12th International Conference on Computer Vision, pages 2090–2097. IEEE, 2009. (Cité en page 218.)

[Scelo 2018] Ghislaine Scelo, Peng Li, Estelle Chanudet et David C Muller. *Variability of sex disparities in cancer incidence over 30 years : the striking case of kidney cancer*. European urology focus, vol. 4, no. 4, pages 586–590, 2018. (Cité en page 69.)

[Scharstein 2002] Daniel Scharstein et Richard Szeliski. *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*. International journal of computer vision, vol. 47, pages 7–42, 2002. (Cité en page 157.)

[Schena 2016] Bruce Michael Schena, Roman L. Devengenzo, Scott Luke, David Martin, Thomas G. Cooper et Thomas Brown. *Multi-Port Surgical Robotic System Architecture*, 2016. (Cité en pages 33, 34 et 35.)

[Schiavina 2020] Riccardo Schiavina, Andrea Mari, Lorenzo Bianchi, Daniele Amparore, Alessandro Antonelli, Walter Artibani, Eugenio Brunocilla, Umberto Capitanio, Cristian Fiori, Fabrizio Di Maida*et al.* *Predicting positive surgical margins in partial nephrectomy : A prospective multicentre observational*

study (the RECORd 2 project). European Journal of Surgical Oncology, vol. 46, no. 7, pages 1353–1359, 2020. (Cité en page 76.)

[Schiavina 2021a] Riccardo Schiavina, Lorenzo Bianchi, Francesco Chessa, Umberto Barbaresi, Laura Cercenelli, Simone Lodi, Caterina Gaudiano, Barbara Bortolani, Andrea Angiolini, Federico Mineo Bianchi*et al. Augmented reality to guide selective clamping and tumor dissection during robot-assisted partial nephrectomy : a preliminary experience.* Clinical genitourinary cancer, vol. 19, no. 3, pages e149–e155, 2021. (Cité en pages 94, 95 et 176.)

[Schiavina 2021b] Riccardo Schiavina, Lorenzo Bianchi, Simone Lodi, Laura Cercenelli, Francesco Chessa, Barbara Bortolani, Caterina Gaudiano, Carlo Casablanca, Matteo Droghetti, Angelo Porreca*et al. Real-time augmented reality three-dimensional guided robotic radical prostatectomy : preliminary experience and evaluation of the impact on surgical planning.* European urology focus, vol. 7, no. 6, pages 1260–1267, 2021. (Cité en pages 171 et 172.)

[Schieda 2015] Nicola Schieda, Christian B Van der Pol, Bardia Moosavi, Matthew DF McInnes, Kien T Mai et Trevor A Flood. *Intracellular lipid in papillary renal cell carcinoma (pRCC) : T2 weighted (T2W) MRI and pathologic correlation.* European radiology, vol. 25, pages 2134–2142, 2015. (Cité en page 71.)

[Schieda 2017] Nicola Schieda, Satheesh Krishna, Matthew DF McInnes, Bardia Moosavi, Abdulmohsen Alrashed, Robert Moreland et Evan S Siegelman. *Utility of MRI to differentiate clear cell renal cell carcinoma adrenal metastases from adrenal adenomas.* American Journal of Roentgenology, vol. 209, no. 3, pages W152–W159, 2017. (Cité en page 71.)

[Schmidt 2021] Adam Schmidt et Septimiu E Salcudean. *Real-time rotated convolutional descriptor for surgical environments.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 279–289. Springer, 2021. (Cité en page 220.)

[Schmidt 2022a] Adam Schmidt, Omid Mohareri, Simon DiMaio et Septimiu E Salcudean. *Fast graph refinement and implicit neural representation for tissue tracking.* In 2022 International Conference on Robotics and Automation (ICRA), pages 1281–1288. IEEE, 2022. (Cité en page 222.)

[Schmidt 2022b] Adam Schmidt, Omid Mohareri, Simon DiMaio et Septimiu E Salcudean. *Recurrent implicit neural graph for deformable tracking in endoscopic videos.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 478–488. Springer, 2022. (Cité en page 222.)

[Schmidt 2023] Adam Schmidt, Omid Mohareri, Simon DiMaio et Septimiu E Salcudean. *SENDD : Sparse efficient neural depth and deformation for tissue tracking.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 238–248. Springer, 2023. (Cité en page 222.)

[Schmidt 2024a] Adam Schmidt, Omid Mohareri, Simon DiMaio, Michael C Yip et Septimiu E Salcudean. *Tracking and mapping in medical computer vision : A review.* Medical Image Analysis, page 103131, 2024. (Cité en pages 219, 220, 221 et 223.)

[Schmidt 2024b] Adam Schmidt, Omid Mohareri, Simon P DiMaio et Septimiu E Salcudean. *Surgical tattoos in infrared : A dataset for quantifying tissue tracking and mapping.* IEEE Transactions on Medical Imaging, vol. 43, no. 7, pages 2634–2645, 2024. (Cité en page 226.)

[Schmit 2010] Grant D Schmit, Thomas D Atwell, Matthew R Callstrom, Michael A Farrell, Bradley C Leibovich, David E Patterson, George K Chow, Michael L Blute et J William Charboneau. *Percutaneous cryoablation of renal masses superior to 3 cm : efficacy and safety in treatment of 108 patients.* Journal of endourology, vol. 24, no. 8, pages 1255–1262, 2010. (Cité en page 77.)

[Schneider 2020] C Schneider, S Thompson, J Totz, Y Song, M Allam, MH Sodergren, AE Desjardins, D Barratt, S Ourselin, K Gurusamy*et al. Comparison of manual and semi-automatic registration in augmented reality image-guided liver surgery : a clinical feasibility study.* Surgical endoscopy, vol. 34, pages 4702–4711, 2020. (Cité en page 172.)

[Schonberger 2016] Johannes L Schonberger et Jan-Michael Frahm. *Structure-from-motion revisited.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4104–4113, 2016. (Cité en page 220.)

[Schönberger 2017] Johannes L Schönberger, True Price, Torsten Sattler, Jan-Michael Frahm et Marc Pollefeys. *A vote-and-verify strategy for fast spatial verification in image retrieval.* In Computer Vision–ACCV 2016 : 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part I 13, pages 321–337. Springer, 2017. (Cité en page 218.)

[Schüle 2022] Johannes Schüle, Jonathan Haag, Peter Somers, Carina Veil, Christina Tarín et Oliver Sawodny. *A model-based simultaneous localization and mapping approach for deformable bodies.* In 2022 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), pages 607–612. IEEE, 2022. (Cité en pages 219 et 223.)

[SCHWAB 2017] Cedric SCHWAB. *Detection of user touch on controller handle*, 2017. (Cité en page 32.)

[Schweiger 2012] Florian Schweiger, Georg Schroth, Michael Eichhorn, Anas Al-Nuaimi, Burak Cizmeci, Michael Fahrmair et Eckehard Steinbach. *Fully automatic and frame-accurate video synchronization using bitrate sequences.* IEEE transactions on multimedia, vol. 15, no. 1, pages 1–14, 2012. (Cité en page 129.)

[Scosyrev 2014a] Emil Scosyrev, Edward M Messing, Richard Sylvester, Steven Campbell et Hendrik Van Poppel. *Renal function after nephron-sparing surgery versus radical nephrectomy : results from EORTC randomized trial*

30904. European urology, vol. 65, no. 2, pages 372–377, 2014. (Cité en page 75.)

[Scosyrev 2014b] Emil Scosyrev, Edward M Messing, Richard Sylvester, Steven Campbell et Hendrik Van Poppel. *Renal function after nephron-sparing surgery versus radical nephrectomy : results from EORTC randomized trial 30904*. European urology, vol. 65, no. 2, pages 372–377, 2014. (Cité en page 79.)

[Sederberg 1986] Thomas W Sederberg et Scott R Parry. *Free-form deformation of solid geometric models*. In Proceedings of the 13th annual conference on Computer graphics and interactive techniques, pages 151–160, 1986. (Cité en pages 95 et 186.)

[Segura 2018] Álvaro Segura, Javier Barandiaran, Aitor Moreno, Iñigo Barandiaran et Julián Flórez. *Improved virtual reality perception with calibrated stereo and variable focus for industrial use*. International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 12, pages 95–103, 2018. (Cité en pages 130 et 131.)

[Selka 2012] F Selka, S Nicolau, A Bessaid, L Soler et J Marescaux. *Performance evaluation of simultaneous RGB analysis for feature detection and tracking in endoscopic images*. In Medical Image Understanding and Analysis, pages 249–254, 2012. (Cité en page 256.)

[Semm 1983] K Semm. *Endoscopic appendectomy*. Endoscopy, vol. 15, no. 02, pages 59–64, 1983. (Cité en page 23.)

[Sengupta 2021] Agniva Sengupta et Adrien Bartoli. *Colonoscopic 3D reconstruction by tubular non-rigid structure-from-motion*. International Journal of Computer Assisted Radiology and Surgery, vol. 16, no. 7, pages 1237–1241, 2021. (Cité en page 222.)

[Servais 2024] Elliot L Servais, Laila Rashidi, Priyanshi Porwal, Mark Garibaldi et Andrew J Hung. *Novel force feedback technology improves suturing in robotic-assisted surgery : a pre-clinical study*. Surgical Endoscopy, pages 1–10, 2024. (Cité en page 37.)

[Service d'Ingénierie Biomédicale 2022] Hôpitaux Universitaires de Genève (HUG) Service d'Ingénierie Biomédicale. *État de l'art de la chirurgie robotique*, 2022. (Cité en page 28.)

[Service d'Ingénierie Biomédicale 2024] Hôpitaux Universitaires de Genève (HUG) Service d'Ingénierie Biomédicale. *État de l'art de la chirurgie robotique*, 2024. (Cité en page 28.)

[Seshamani 2006] Sharmishtaa Seshamani, William Lau et Gregory Hager. *Real-time endoscopic mosaicking*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 355–363. Springer, 2006. (Cité en page 221.)

[Shafer 2006] David C. Shafer et Dennis C. Leiner. *Stereoscopic endoscope*, 2006. (Cité en pages 36 et 37.)

[Shah 2016] Paras H Shah, Daniel M Moreira, Zhamshid Okhunov, Vinay R Patel, Sameer Chopra, Aria A Razmaria, Manaf Alom, Arvin K George, Oksana Yaskiv, Michael J Schwartz*et al.* *Positive surgical margins increase risk of recurrence after partial nephrectomy for high risk renal tumors.* The Journal of urology, vol. 196, no. 2, pages 327–334, 2016. (Cité en page 73.)

[Shah 2017] Paras H Shah, Daniel M Moreira, Vinay R Patel, Geoffrey Gaunay, Arvin K George, Manaf Alom, Zachary Kozel, Oksana Yaskiv, Simon J Hall, Michael J Schwartz*et al.* *Partial nephrectomy is associated with higher risk of relapse compared with radical nephrectomy for clinical stage T1 renal cell carcinoma pathologically up staged to T3a.* The Journal of urology, vol. 198, no. 2, pages 289–296, 2017. (Cité en page 76.)

[Shakil 2006] Omer Shakil. *An efficient video alignment approach for non-overlapping sequences with free camera movement.* In 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, volume 2, pages II–II. IEEE, 2006. (Cité en page 129.)

[Shankar 2013a] Sukrit Shankar, Joan Lasenby et Anil Kokaram. *Synchronization of user-generated videos through trajectory correspondence and a refinement procedure.* In Proceedings of the 10th European Conference on Visual Media Production, pages 1–10, 2013. (Cité en page 129.)

[Shankar 2013b] Sukrit Shankar, Joan Lasenby et Anil Kokaram. *Warping trajectories for video synchronization.* In Proceedings of the 4th ACM/IEEE international workshop on Analysis and retrieval of tracked events and motion in imagery stream, pages 41–48, 2013. (Cité en page 129.)

[Shapiro 2020] Daniel D Shapiro, Shane A Wells, Sara L Best, Sean P Hedican, Timothy J Ziemlewicz, Meghan G Lubner, James Louis Hinshaw, Fred T Lee Jr, David F Jarrard, Kyle A Richards*et al.* *Comparing outcomes for patients with clinical T1b renal cell carcinoma treated with either percutaneous microwave ablation or surgery.* Urology, vol. 135, pages 88–94, 2020. (Cité en page 77.)

[Sharifian 2025a] Rasoul Sharifian et Adrien Bartoli. *SurgIPC : a convex image perspective correction method to boost surgical keypoint matching.* International Journal of Computer Assisted Radiology and Surgery, pages 1–12, 2025. (Cité en page 231.)

[Sharifian 2025b] Rasoul Sharifian, Navid Rabbani et Adrien Bartoli. *The RoDEM benchmark : evaluating the robustness of monocular single-shot depth estimation methods in minimally-invasive surgery.* International Journal of Computer Assisted Radiology and Surgery, pages 1–15, 2025. (Cité en pages 109 et 171.)

[Shen 2017] Liang Shen, Xiaobing Feng, Yuan Zhang, Min Shi, Dengming Zhu et Zhaoqi Wang. *Stroboscope based synchronization of full frame CCD sensors.* Sensors, vol. 17, no. 4, page 799, 2017. (Cité en page 127.)

[Shi 1994] Jianbo Shi*et al.* *Good features to track.* In 1994 Proceedings of IEEE conference on computer vision and pattern recognition, pages 593–600. IEEE, 1994. (Cité en page 145.)

[Shi 2020] Taoping Shi, Qingbo Huang, Kan Liu, Songliang Du, Yang Fan, Luojia Yang, Cheng Peng, Dan Shen, Zhongxin Wang, Yu Gao*et al.* *Robot-assisted cavectomy versus thrombectomy for level II inferior vena cava thrombus : decision-making scheme and multi-institutional analysis.* European urology, vol. 78, no. 4, pages 592–602, 2020. (Cité en page 79.)

[Shi 2022] Yan Shi, Jun-Xiong Cai, Yoli Shavit, Tai-Jiang Mu, Wensen Feng et Kai Zhang. *Clustergnn : Cluster-based coarse-to-fine graph neural network for efficient feature matching.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 12517–12526, 2022. (Cité en page 219.)

[Shirazian 2019] Pourya Shirazian, Mahdi Azizian et Daniel Proksch. *Multi-modal visualization in computer-assisted tele-operated surgery*, 2019. (Cité en page 34.)

[Shrestha 2006] Prarthana Shrestha, Hans Weda, Mauro Barbieri et Dragan Sekulovski. *Synchronization of multiple video recordings based on still camera flashes.* In Proceedings of the 14th ACM international conference on Multimedia, pages 137–140, 2006. (Cité en page 129.)

[Silva 2020] Alessandro R Silva et Mylène CQ Farias. *Perceptual quality assessment of 3D videos with stereoscopic degradations.* Multimedia Tools and Applications, vol. 79, no. 1, pages 1603–1623, 2020. (Cité en page 131.)

[Silverman 2019] Stuart G Silverman, Ivan Pedrosa, James H Ellis, Nicole M Hindman, Nicola Schieda, Andrew D Smith, Erick M Remer, Atul B Shinagare, Nicole E Curci, Steven S Raman*et al.* *Bosniak classification of cystic renal masses, version 2019 : an update proposal and needs assessment.* Radiology, vol. 292, no. 2, pages 475–488, 2019. (Cité en page 72.)

[Simhan 2011] Jay Simhan, Marc C Smaldone, Kevin J Tsai, Daniel J Canter, Tianyu Li, Alexander Kutikov, Rosalia Viterbo, David YT Chen, Richard E Greenberg et Robert G Uzzo. *Objective measures of renal mass anatomic complexity predict rates of major complications following partial nephrectomy.* European urology, vol. 60, no. 4, pages 724–730, 2011. (Cité en page 81.)

[Simonyan 2014a] Karen Simonyan, Andrea Vedaldi et Andrew Zisserman. *Learning local feature descriptors using convex optimisation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 8, pages 1573–1585, 2014. (Cité en page 145.)

[Simonyan 2014b] Karen Simonyan et Andrew Zisserman. *Very deep convolutional networks for large-scale image recognition.* arXiv preprint arXiv :1409.1556, 2014. (Cité en pages 128 et 183.)

[Simpfendörfer 2016] Tobias Simpfendörfer, Claudia Gasch, Gencay Hatiboglu, Michael Müller, Lena Maier-Hein, Markus Hohenfellner et Dogu Teber. *Intraoperative computed tomography imaging for navigated laparoscopic renal surgery : first clinical experience.* Journal of endourology, vol. 30, no. 10, pages 1105–1111, 2016. (Cité en page 95.)

[Singla 2017] Rohit Singla, Philip Edgcumbe, Philip Pratt, Christopher Nguan et Robert Rohling. *Intra-operative ultrasound-based augmented reality guidance for laparoscopic surgery.* Healthcare technology letters, vol. 4, no. 5, pages 204–209, 2017. (Cité en page 96.)

[Sinha 2004] Sudipta N Sinha et Marc Pollefeys. *Synchronization and calibration of camera networks from silhouettes.* In Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., volume 1, pages 116–119. IEEE, 2004. (Cité en page 128.)

[Sinha 2010] Sudipta N Sinha et Marc Pollefeys. *Camera network calibration and synchronization from silhouettes in archived video.* International journal of computer vision, vol. 87, pages 266–283, 2010. (Cité en page 129.)

[Siva 2020] Shankar Siva, Rohann JM Correa, Andrew Warner, Michael Staehler, Rodney J Ellis, Lee Ponsky, Irving D Kaplan, Anand Mahadevan, William Chu, Senthilkumar Gandhidasan*et al. Stereotactic ablative radiotherapy for T1b primary renal cell carcinoma : A report from the International Radiosurgery Oncology Consortium for Kidney (IROCK).* International Journal of Radiation Oncology* Biology* Physics, vol. 108, no. 4, pages 941–949, 2020. (Cité en page 77.)

[Smelyanskiy 2009] Mikhail Smelyanskiy, David Holmes, Jatin Chhugani, Alan Larson, Douglas M Carmean, Dennis Hanson, Pradeep Dubey, Kurt Augustine, Daehyun Kim, Alan Kyker*et al. Mapping high-fidelity volume rendering for medical imaging to CPU, GPU and many-core architectures.* IEEE transactions on visualization and computer graphics, vol. 15, no. 6, pages 1563–1570, 2009. (Cité en page 44.)

[Smeulders 2013] Arnold WM Smeulders, Dung M Chu, Rita Cucchiara, Simone Calderara, Afshin Dehghan et Mubarak Shah. *Visual tracking : An experimental survey.* IEEE transactions on pattern analysis and machine intelligence, vol. 36, pages 1442–1468, 2013. (Cité en page 228.)

[Smid 2019] Matej Smid et Jiri Matas. *Rolling shutter camera synchronization with sub-millisecond accuracy.* arXiv preprint arXiv :1902.11084, 2019. (Cité en page 129.)

[Sobhy 2019] Soha Sobhy, David Arroyo-Manzano, Nilaani Murugesu, Gayathri Karthikeyan, Vinoth Kumar, Inderjeet Kaur, Evita Fernandez, Sirisha Rao Gundabattula, Ana Pilar Betran, Khalid Khan*et al. Maternal and perinatal mortality and complications associated with caesarean section in low-income and middle-income countries : a systematic review and meta-analysis.* The Lancet, vol. 393, no. 10184, pages 1973–1982, 2019. (Cité en page 16.)

[Song 2017] Jingwei Song, Jun Wang, Liang Zhao, Shoudong Huang et Gamini Dissanayake. *Dynamic reconstruction of deformable soft-tissue with stereo scope in minimal invasive surgery.* IEEE Robotics and Automation Letters, vol. 3, no. 1, pages 155–162, 2017. (Cité en page 171.)

[Song 2018] Jingwei Song, Jun Wang, Liang Zhao, Shoudong Huang et Gamini Dissanayake. *Mis-slam : Real-time large-scale dense deformable slam system in minimal invasive surgery based on heterogeneous computing.* IEEE Robotics and Automation Letters, vol. 3, no. 4, pages 4068–4075, 2018. (Cité en page 223.)

[Song 2024] Yunzhou Song, Jiahui Lei, Ziyun Wang, Lingjie Liu et Kostas Daniilidis. *Track everything everywhere fast and robustly.* In European Conference on Computer Vision, pages 343–359. Springer, 2024. (Cité en page 225.)

[Sousa 2015] Ricardo M Sousa, Martin Wäny, Pedro Santos et Morgado Dias. *Multi-camera synchronization core implemented on USB3 based FPGA platform.* In Image Sensors and Imaging Systems 2015, volume 9403, pages 32–39. SPIE, 2015. (Cité en page 127.)

[Soviany 2022] Petru Soviany, Radu Tudor Ionescu, Paolo Rota et Nicu Sebe. *Curriculum learning : A survey.* International Journal of Computer Vision, vol. 130, no. 6, pages 1526–1565, 2022. (Cité en page 202.)

[Spencer 2004] Lisa Spencer et Mubarak Shah. *Temporal synchronization from camera motion.* In Proceedings of Asian Conference on Computer Vision, pages 515–520. Citeseer, 2004. (Cité en page 129.)

[Spreafico 2021] Filippo Spreafico, Conrad V Fernandez, Jesper Brok, Kayo Nakata, Gordan Vujanic, James I Geller, Manfred Gessler, Mariana Maschietto, Sam Behjati, Angela Polanco*et al. Wilms tumour.* Nature Reviews Disease Primers, vol. 7, no. 1, page 75, 2021. (Cité en page 68.)

[Srivastava 2020] Arnav Srivastava, Zorimar Rivera-Núñez, Sinae Kim, Joshua Sterling, Nicholas J Farber, Kushan D Radadia, Hiren V Patel, Parth K Modi, Sharad Goyal, Rahul Parikh*et al. Impact of pathologic lymph node-positive renal cell carcinoma on survival in patients without metastasis : Evidence in support of expanding the definition of stage IV kidney cancer.* Cancer, vol. 126, no. 13, pages 2991–3001, 2020. (Cité en page 77.)

[St John 2020] Ace St John, Ilaria Caturegli, Natalia S Kubicki et Stephen M Kavic. *The rise of minimally invasive surgery : 16 year analysis of the progressive replacement of open surgery with laparoscopy.* JSLS : Journal of the Society of Laparoscopic & Robotic Surgeons, vol. 24, no. 4, pages e2020–00076, 2020. (Cité en page 24.)

[Stehlé 2023] Thomas Stehlé, Félix Wei, Séverine Brabant, Alain Luciani, Philippe Grimbert, Dominique Prié, Edouard Reizine, Antoine Durrbach, Sébastien Mulé, Anne Hulin*et al. Glomerular filtration rate measured based on iomeprol clearance assessed at CT urography in living kidney donor candidates.* Radiology, vol. 309, no. 3, page e230567, 2023. (Cité en page 71.)

[Steinberg 2004] Andrew P Steinberg, Antonio Finelli, Mihir M Desai, Sidney C Abreu, Anup P Ramani, Massimiliano Spaliviero, Lisa Rybicki, Jihad Kaouk, Andrew C Novick et Inderbir S Gill. *Laparoscopic radical nephrectomy for large (greater than 7 cm, T2) renal tumors.* The Journal of urology, vol. 172, no. 6, pages 2172–2176, 2004. (Cité en page 76.)

[Stewart 2015] Gregory Stewart, Lewis Johnson, Halemane Ganesh, Daniel Davenport, Woodson Smelser, Paul Crispen et Ramakrishna Venkatesh. *Stone size limits the use of Hounsfield units for prediction of calcium oxalate stone composition.* Urology, vol. 85, no. 2, pages 292–295, 2015. (Cité en page 77.)

[Stoyanov 2005a] Danail Stoyanov, Ara Darzi et Guang-Zhong Yang. *Laparoscope self-calibration for robotic assisted minimally invasive surgery.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 114–121. Springer, 2005. (Cité en page 107.)

[Stoyanov 2005b] Danail Stoyanov, George P Mylonas, Fani Deligianni, Ara Darzi et Guang Zhong Yang. *Soft-tissue motion tracking and structure estimation for robotic assisted MIS procedures.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 139–146. Springer, 2005. (Cité en page 226.)

[Stoyanov 2010] Danail Stoyanov, Marco Visentini Scarzanella, Philip Pratt et Guang-Zhong Yang. *Real-time stereo reconstruction in robotically assisted minimally invasive surgery.* In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010 : 13th International Conference, Beijing, China, September 20-24, 2010, Proceedings, Part I 13, pages 275–282. Springer, 2010. (Cité en pages 97 et 226.)

[Su 2009] Li-Ming Su, Balazs P Vagvolgyi, Rahul Agarwal, Carol E Reiley, Russell H Taylor et Gregory D Hager. *Augmented reality during robot-assisted laparoscopic partial nephrectomy : toward real-time 3D-CT to stereoscopic video registration.* Urology, vol. 73, no. 4, pages 896–900, 2009. (Cité en pages 97 et 176.)

[Su 2021] Yinqiang Su, Jinghong Liu, Fang Xu, Xueming Zhang et Yujia Zuo. *A novel anti-drift visual object tracking algorithm based on sparse response and adaptive spatial-temporal context-aware.* Remote Sensing, vol. 13, page 4672, 2021. (Cité en page 229.)

[Suárez 2020] Iago Suárez, Ghesn Sfeir, José M Buenaposada et Luis Baumela. *BEBLID : Boosted efficient binary local image descriptor.* Pattern recognition letters, vol. 133, pages 366–372, 2020. (Cité en page 145.)

[Sufrin 1989] G Sufrin, S Chasan, A Golio et GP Murphy. *Paraneoplastic and serologic syndromes of renal adenocarcinoma.* In Seminars in urology, volume 7, pages 158–171, 1989. (Cité en page 73.)

[Sumner 2007] Robert W Sumner, Johannes Schmid et Mark Pauly. *Embedded deformation for shape manipulation.* In ACM siggraph 2007 papers, pages 80–es. 2007. (Cité en page 223.)

[Sun 2018] Deqing Sun, Xiaodong Yang, Ming-Yu Liu et Jan Kautz. *Pwc-net : Cnns for optical flow using pyramid, warping, and cost volume*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 8934–8943, 2018. (Cité en page 222.)

[Sun 2021] Jiaming Sun, Zehong Shen, Yuang Wang, Hujun Bao et Xiaowei Zhou. *LoFTR : Detector-free local feature matching with transformers*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 8922–8931, 2021. (Cité en pages 146 et 217.)

[Sun 2023] Tianze Sun, Xiwang He et Zhonghai Li. *Digital twin in healthcare : Recent updates and challenges*. Digital Health, vol. 9, page 20552076221149651, 2023. (Cité en page 44.)

[Supervisely Team 2025] Supervisely Team. *Supervisely*, 2025. (Cité en page 238.)

[Suresh 2023] Ashwinram Suresh. *Surgical system haptic feedback systems*, 2023. (Cité en pages 37 et 38.)

[Surgical 2024] Intuitive Surgical. *Intuitive announces third quarter earnings*, 2024. (Cité en page 29.)

[Surgical 2025a] Intuitive Surgical. *Da Vinci 5 System User Manual*, 2025. (Cité en pages 31 et 33.)

[Surgical 2025b] Intuitive Surgical. *da Vinci Xi System User Manual*, 2025. (Cité en page 103.)

[Sutskever 2013] Ilya Sutskever, James Martens, George Dahl et Geoffrey Hinton. *On the importance of initialization and momentum in deep learning*. In International conference on machine learning, pages 1139–1147. pmlr, 2013. (Cité en page 262.)

[Szegedy 2017] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke et Alexander Alemi. *Inception-v4, inception-resnet and the impact of residual connections on learning*. In Proceedings of the AAAI conference on artificial intelligence, volume 31, 2017. (Cité en page 128.)

[Tahbaz 2018] Rana Tahbaz, Marianne Schmid et Axel S Merseburger. *Prevention of kidney cancer incidence and recurrence : lifestyle, medication and nutrition*. Current opinion in urology, vol. 28, no. 1, pages 62–79, 2018. (Cité en page 69.)

[Takahashi 2016] Kosuke Takahashi, Shohei Nobuhara et Takashi Matsuyama. *Mirror-based camera pose estimation using an orthogonality constraint*. IPSJ Transactions on Computer Vision and Applications, vol. 8, pages 11–19, 2016. (Cité en page 183.)

[Talenfeld 2018] Adam D Talenfeld, Renee L Gennarelli, Elena B Elkin, Coral L Atoria, Jeremy C Durack, William C Huang et Sharon W Kwan. *Percutaneous ablation versus partial and radical nephrectomy for T1a renal cancer : a population-based analysis*. Annals of internal medicine, vol. 169, no. 2, pages 69–77, 2018. (Cité en pages 73 et 74.)

[Tan 2013] Puay Hoon Tan, Liang Cheng, Nathalie Rioux-Leclercq, Maria J Merino, George Netto, Victor E Reuter, Steven S Shen, David J Grignon, Rodolfo Montironi, Lars Egevad*et al.* *Renal tumors : diagnostic and prognostic biomarkers.* The American journal of surgical pathology, vol. 37, no. 10, pages 1518–1531, 2013. (Cité en page 73.)

[Tan 2019a] Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard et Quoc V Le. *Mnasnet : Platform-aware neural architecture search for mobile.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 2820–2828, 2019. (Cité en page 240.)

[Tan 2019b] Mingxing Tan et Quoc Le. *Efficientnet : Rethinking model scaling for convolutional neural networks.* In International conference on machine learning, pages 6105–6114. PMLR, 2019. (Cité en pages 240 et 262.)

[Tan 2021] Jiaqi Tan, Weijie Lin, Angel X Chang et Manolis Savva. *Mirror3D : Depth refinement for mirror surfaces.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 15990–15999, 2021. (Cité en page 133.)

[Tan 2022] Xin Tan, Jiaying Lin, Ke Xu, Pan Chen, Lizhuang Ma et Rynson WH Lau. *Mirror detection with the visual chirality cue.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 3, pages 3492–3504, 2022. (Cité en page 132.)

[Tanaka 2021] Hajime Tanaka, Xiaobo Ding, Yunlin Ye, Yanbo Wang, Rebecca A Campbell, Molly E DeWitt-Foy, Chalairat Suk-Ouichai, Ryan D Ward, Erick M Remer, Jianbo Li*et al.* *Infiltrative renal masses : clinical significance and fidelity of documentation.* European Urology Oncology, vol. 4, no. 2, pages 264–273, 2021. (Cité en page 78.)

[Tang 2017] Rui Tang, Longfei Ma, Canhong Xiang, Xuedong Wang, Ang Li, Hongen Liao et Jiahong Dong. *Augmented reality navigation in open surgery for hilar cholangiocarcinoma resection with hemihepatectomy using video-based in situ three-dimensional anatomical modeling : a case report.* Medicine, vol. 96, no. 37, page e8083, 2017. (Cité en page 171.)

[Tang 2021] Rui Tang, Wei Yang, Yucheng Hou, Lihan Yu, Guangdong Wu, Xuan Tong, Jun Yan et Qian Lu. *Augmented reality-assisted pancreaticoduodenectomy with superior mesenteric vein resection and reconstruction.* Gastroenterology research and practice, vol. 2021, no. 1, page 9621323, 2021. (Cité en page 175.)

[Teber 2009] Dogu Teber, Selcuk Guven, Tobias Simpfendörfer, Mathias Baumhauer, Esref Oguz Güven, Faruk Yencilek, Ali Serdar Gözen et Jens Rassweiler. *Augmented reality : a new tool to improve surgical accuracy during laparoscopic partial nephrectomy ? Preliminary in vitro and in vivo results.* European urology, vol. 56, no. 2, pages 332–338, 2009. (Cité en pages 94 et 95.)

[Teluob 2019] Guillaume Teluob, Lilian Calvet, Yamid Espinel, Laurent Savareux, Marlene Guandalino, Anne Ravel, Nicolas Bourdel, Pauline Chauvet, Laurent Guy, Pascal Chabrot*et al.* *Preliminary trial of augmented reality performed on a regular and a robot-assisted laparoscopic partial nephrectomies.* Videourology, vol. 33, no. 3, 2019. (Cité en pages 51, 52, 94 et 246.)

[The Lancet Commission on Global Surgery 2015] The Lancet Commission on Global Surgery. *La chirurgie dans le monde à l'horizon 2030 : données et solutions pour la santé, le bien-être, et le développement économique*, 2015. (Cité en page 15.)

[Thompson 2015] R Houston Thompson, Tom Atwell, Grant Schmit, Christine M Lohse, A Nicholas Kurup, Adam Weisbrod, Sarah P Psutka, Suzanne B Stewart, Matthew R Callstrom, John C Cheville*et al.* *Comparison of partial nephrectomy and percutaneous ablation for cT1 renal masses.* European urology, vol. 67, no. 2, pages 252–259, 2015. (Cité en page 74.)

[Thompson 2018] Stephen Thompson, Crispin Schneider, Michele Bosi, Kurinchi Gurusamy, Sébastien Ourselin, Brian Davidson, David Hawkes et Matthew J Clarkson. *In vivo estimation of target registration errors during augmented reality laparoscopic surgery.* International journal of computer assisted radiology and surgery, vol. 13, pages 865–874, 2018. (Cité en page 171.)

[Thomsen 1998] Henrik S Thomsen et William H Bush Jr. *Adverse effects of contrast media : incidence, prevention and management.* Drug safety, vol. 19, no. 4, pages 313–324, 1998. (Cité en page 70.)

[Tian 2023] Yating Tian, Hongwen Zhang, Yebin Liu et Limin Wang. *Recovering 3d human mesh from monocular images : A survey.* IEEE transactions on pattern analysis and machine intelligence, 2023. (Cité en page 178.)

[Toft 2020] Carl Toft, Daniyar Turmukhambetov, Torsten Sattler, Fredrik Kahl et Gabriel J Brostow. *Single-image depth prediction makes feature matching easier.* In Computer Vision–ECCV 2020 : 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVI 16, pages 473–492. Springer, 2020. (Cité en page 230.)

[Tola 2009] Engin Tola, Vincent Lepetit et Pascal Fua. *Daisy : An efficient dense descriptor applied to wide-baseline stereo.* IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 5, pages 815–830, 2009. (Cité en pages 137 et 145.)

[Tomasi 1991] Carlo Tomasi et Takeo Kanade. *Detection and tracking of point.* Int J Comput Vis, vol. 9, no. 137-154, page 3, 1991. (Cité en page 221.)

[Tombari 2014] Federico Tombari et Luigi Di Stefano. *Interest points via maximal self-dissimilarities.* In Asian Conference on Computer Vision, pages 586–600. Springer, 2014. (Cité en page 145.)

[Torr 2000] Philip HS Torr et Andrew Zisserman. *MLESAC : A new robust estimator with application to estimating image geometry.* Computer vision and image understanding, vol. 78, no. 1, pages 138–156, 2000. (Cité en page 218.)

[Torr 2002] Philip Hilaire Torr, Slawomir J Nasuto et John Mark Bishop. *Napsac : High noise, high dimensional robust estimation-it's in the bag.* In British Machine Vision Conference (BMVC), volume 2, page 3, 2002. (Cité en page 218.)

[Torres 2003] Fernando Torres, Jesús Angulo et Francisco Ortiz. *Automatic detection of specular reflectance in colour images using the MS diagram.* In International Conference on Computer Analysis of Images and Patterns, pages 132–139. Springer, 2003. (Cité en page 132.)

[Torresani 2008] Lorenzo Torresani, Aaron Hertzmann et Chris Bregler. *Nonrigid structure-from-motion : Estimating shape and motion with hierarchical priors.* IEEE transactions on pattern analysis and machine intelligence, vol. 30, no. 5, pages 878–892, 2008. (Cité en page 222.)

[Tosi 2024] Fabio Tosi, Youmin Zhang, Ziren Gong, Erik Sandström, Stefano Mattoccia, Martin R Oswald et Matteo Poggi. *How nerfs and 3d gaussian splatting are reshaping slam : a survey.* arXiv preprint arXiv :2402.13255, vol. 4, page 1, 2024. (Cité en page 190.)

[Tourdias 2025] T. Tourdias et V. Lebon. *Imagerie par résonance magnétique.* In Les fondamentaux de l'imagerie médicale. Collège National des Enseignants de Biophysique et de Médecine Nucléaire, 2025. (Cité en pages 42 et 43.)

[Touvron 2022] Hugo Touvron, Matthieu Cord et Hervé Jégou. *Deit iii : Revenge of the vit.* In European conference on computer vision, pages 516–533. Springer, 2022. (Cité en page 225.)

[Tresadern 2003] Philip A Tresadern et Ian Reid. *Synchronizing Image Sequences of Non-Rigid Objects.* In BMVC, volume 2, 2003. (Cité en page 129.)

[Tretschk 2023] Edith Tretschk, Navami Kairanda, Mallikarjun BR, Rishabh Dabral, Adam Kortylewski, Bernhard Egger, Marc Habermann, Pascal Fua, Christian Theobalt et Vladislav Golyanik. *State of the art in dense monocular non-rigid 3D reconstruction.* In Computer Graphics Forum, volume 42, pages 485–520. Wiley Online Library, 2023. (Cité en page 222.)

[Tripathi 2017] Ankit Tripathi, Benu Changmai, Shrukul Habib, Nagaratna B Chittaragi et Shashidhar G Koolagudi. *Normalized videosnapping : A non-linear video synchronization approach.* In 2017 Tenth International Conference on Contemporary Computing (IC3), pages 1–6. IEEE, 2017. (Cité en page 128.)

[Troccaz 2024] Jocelyne Troccaz. *Robotique interventionnelle : promesses, réalités et défis.* Bulletin de l'Académie Nationale de Médecine, 2024. (Cité en page 27.)

[Trzcinski 2013] Tomasz Trzcinski, Mario Christoudias, Pascal Fua et Vincent Lepetit. *Boosting binary keypoint descriptors.* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2874–2881, 2013. (Cité en page 145.)

[Tumanyan 2024] Narek Tumanyan, Assaf Singer, Shai Bagon et Tali Dekel. *Dino-tracker : Taming dino for self-supervised point tracking in a single video.* In

European Conference on Computer Vision, pages 367–385. Springer, 2024. (Cité en page 225.)

[Tunali 2018] Emre Tunali, Sinan Öz et Mustafa Eral. *Enhancing Correlation Filter based Trackers with Size Adaptivity and Drift Prevention.* In VISIGRAPP (5 : VISAPP), pages 472–480, 2018. (Cité en page 229.)

[Tuytelaars 2004] Tinne Tuytelaars et Luc Van Gool. *Synchronizing video sequences.* In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., volume 1, pages I–I. IEEE, 2004. (Cité en page 129.)

[Tyszkiewicz 2020] Michał Tyszkiewicz, Pascal Fua et Eduard Trulls. *DISK : Learning local features with policy gradient.* Advances in Neural Information Processing Systems, vol. 33, pages 14254–14265, 2020. (Cité en page 146.)

[Uhlig 2019] Johannes Uhlig, Arne Strauss, Gerta Rücker, Ali Seif Amir Hosseini, Joachim Lotz, Lutz Trojan, Hyun S Kim et Annemarie Uhlig. *Partial nephrectomy versus ablative techniques for small renal masses : a systematic review and network meta-analysis.* European radiology, vol. 29, pages 1293–1307, 2019. (Cité en page 74.)

[Ukimura 2008] Osamu Ukimura et Inderbir S Gill. *Imaging-assisted endoscopic surgery : Cleveland Clinic experience.* Journal of endourology, vol. 22, no. 4, pages 803–810, 2008. (Cité en page 94.)

[Ukrainitz 2006] Yaron Ukrainitz et Michal Irani. *Aligning sequences and actions by maximizing space-time correlations.* In Computer Vision–ECCV 2006 : 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part III 9, pages 538–550. Springer, 2006. (Cité en page 129.)

[Uppal 2010] Talat Uppal et Ritu Mogra. *RBC motion and the basis of ultrasound Doppler instrumentation.* Australasian journal of ultrasound in medicine, vol. 13, no. 1, pages 32–34, 2010. (Cité en page 47.)

[Usher-Smith 2022] Juliet A Usher-Smith, Angela Godoy, Sarah W Burge, Simon Burbidge, Jon Cartledge, Philip AJ Crosbie, Claire Eckert, Fiona Farquhar, David Hammond, Neil Hancock*et al. The Yorkshire Kidney Screening Trial (YKST) : protocol for a feasibility study of adding non-contrast abdominal CT scanning to screen for kidney cancer and other abdominal pathology within a trial of community-based CT screening for lung cancer.* BMJ open, vol. 12, no. 9, page e063018, 2022. (Cité en page 70.)

[Ushizaki 2006] Manabu Ushizaki, Takayuki Okatani et Koichiro Deguchi. *Video synchronization based on co-occurrence of appearance changes in video sequences.* In 18th International Conference on Pattern Recognition (ICPR'06), volume 3, pages 71–74. IEEE, 2006. (Cité en page 129.)

[Uzosike 2018] Akachimere C Uzosike, Hiten D Patel, Ridwan Alam, Zeyad R Schwen, Mohit Gupta, Michael A Gorin, Michael H Johnson, Heather Gausepohl, Mark F Riffon, Bruce J Trock*et al. Growth kinetics of small renal*

masses on active surveillance : variability and results from the DISSRM registry. The Journal of urology, vol. 199, no. 3, pages 641–648, 2018. (Cité en page 75.)

[Vagvolgyi 2008] Balazs Vagvolgyi, Li-Ming Su, Russell Taylor et Gregory D Hager. *Video to CT registration for image overlay on solid organs.* Proc. Augmented Reality in Medical Imaging and Augmented Reality in Computer-Aided Surgery (AMIARCS), pages 78–86, 2008. (Cité en page 98.)

[Van der Molen 2018a] Aart J Van der Molen, Peter Reimer, Ilona A Dekkers, Georg Bongartz, Marie-France Bellin, Michele Bertolotto, Olivier Clement, Gertraud Heinz-Peer, Fulvio Stacul, Judith AW Webb*et al. Post-contrast acute kidney injury–Part 1 : Definition, clinical features, incidence, role of contrast medium and risk factors : Recommendations for updated ESUR Contrast Medium Safety Committee guidelines.* European radiology, vol. 28, pages 2845–2855, 2018. (Cité en page 70.)

[Van der Molen 2018b] Aart J Van der Molen, Peter Reimer, Ilona A Dekkers, Georg Bongartz, Marie-France Bellin, Michele Bertolotto, Olivier Clement, Gertraud Heinz-Peer, Fulvio Stacul, Judith AW Webb*et al. Post-contrast acute kidney injury. Part 2 : risk stratification, role of hydration and other prophylactic measures, patients taking metformin and chronic dialysis patients : recommendations for updated ESUR Contrast Medium Safety Committee guidelines.* European radiology, vol. 28, pages 2856–2869, 2018. (Cité en page 70.)

[Van Poppel 2011] Hendrik Van Poppel, Luigi Da Pozzo, Walter Albrecht, Vsevolod Matveev, Aldo Bono, Andrzej Borkowski, Marc Colombel, Laurence Klotz, Eila Skinner, Thomas Keane*et al. A prospective, randomised EORTC intergroup phase 3 study comparing the oncologic outcome of elective nephron-sparing surgery and radical nephrectomy for low-stage renal cell carcinoma.* European urology, vol. 59, no. 4, pages 543–552, 2011. (Cité en pages 73 et 75.)

[Vasconcelos 2019] Francisco Vasconcelos, Evangelos Mazomenos, John Kelly et Danail Stoyanov. *RCM-SLAM : Visual localisation and mapping under remote centre of motion constraints.* In 2019 International conference on robotics and automation (ICRA), pages 9278–9284. IEEE, 2019. (Cité en page 223.)

[Veccia 2020] Alessandro Veccia, Alessandro Antonelli, Robert G Uzzo, Giacomo Novara, Alexander Kutikov, Vincenzo Ficarra, Claudio Simeone, Vincenzo Mirone, Lance J Hampton, Ithaar Derweesh*et al. Predictive value of nephrometry scores in nephron-sparing surgery : a systematic review and meta-analysis.* European urology focus, vol. 6, no. 3, pages 490–504, 2020. (Cité en page 72.)

[Verhoest 2016] Grégory Verhoest, Jean-Philippe Couapel, Emmanuel Oger, Nathalie Rioux-Leclercq, Géraldine Pignot, Jean-Jacques Patard, Axel Bex, Paul Panayotopoulos, Pierre Bigot, Viktor Eret*et al. Safety and feasibility of laparoscopic nephrectomy for big tumors : A retrospective multicentric study.*

Clinical Genitourinary Cancer, vol. 14, no. 4, pages e335–e340, 2016. (Cité en page 77.)

[Verma 2018] Kamlesh Verma, Avnish Kumar et Debashis Ghosh. *Robust Stabilised Visual Tracker for Vehicle Tracking.* Defence Science Journal, vol. 68, 2018. (Cité en page 229.)

[Veyres 2014] Dr. P. Veyres. *La Qualité de l'Air et la Maîtrise de la Contamination Aéroportée*, 2014. (Cité en page 14.)

[Visentini-Scarzanella 2017] Marco Visentini-Scarzanella, Takamasa Sugiura, Toshimitsu Kaneko et Shinichiro Koto. *Deep monocular 3D reconstruction for assisted navigation in bronchoscopy.* International journal of computer assisted radiology and surgery, vol. 12, no. 7, pages 1089–1099, 2017. (Cité en page 226.)

[Volpe 2015] Alessandro Volpe, Michael L Blute, Vincenzo Ficarra, Inderbir S Gill, Alexander Kutikov, Francesco Porpiglia, Craig Rogers, Karim A Touijer, Hendrik Van Poppel et R Houston Thompson. *Renal ischemia and function after partial nephrectomy : a collaborative review of the literature.* European urology, vol. 68, no. 1, pages 61–74, 2015. (Cité en pages 79 et 81.)

[Von Rundstedt 2017] Friedrich-Carl Von Rundstedt, Jason M Scovell, Smriti Agrawal, Jacques Zaneveld et Richard E Link. *Utility of patient-specific silicone renal models for planning and rehearsal of complex tumour resections prior to robot-assisted laparoscopic partial nephrectomy.* BJU international, vol. 119, no. 4, pages 598–604, 2017. (Cité en page 45.)

[Vries 2014] Joey De Vries. *LearnOpenGL*, 2014. tutorial website. (Cité en page 185.)

[Wake 2019] Nicole Wake, Andrew B Rosenkrantz, Richard Huang, Katalina U Park, James S Wysock, Samir S Taneja, William C Huang, Daniel K Sodickson et Hersh Chandarana. *Patient-specific 3D printed and augmented reality kidney and prostate cancer models : impact on patient education.* 3D printing in medicine, vol. 5, pages 1–8, 2019. (Cité en page 46.)

[Wang 2004] Zhou Wang, Alan C Bovik, Hamid R Sheikh et Eero P Simoncelli. *Image quality assessment : from error visibility to structural similarity.* IEEE transactions on image processing, vol. 13, no. 4, pages 600–612, 2004. (Cité en page 146.)

[Wang 2014a] Jiaji Wang, Zhenjiang Miao et Qinghua Liang. *Synchronization of cameras from human motion using feature points.* In 2014 12th International Conference on Signal Processing (ICSP), pages 1281–1284. IEEE, 2014. (Cité en page 128.)

[Wang 2014b] Oliver Wang, Christopher Schroers, Henning Zimmer, Markus Gross et Alexander Sorkine-Hornung. *Videosnapping : Interactive synchronization of multiple videos.* ACM Transactions on Graphics (TOG), vol. 33, no. 4, pages 1–10, 2014. (Cité en page 129.)

[Wang 2015] Dongwen Wang, Bin Zhang, Xiaobin Yuan, Xuhui Zhang et Chen Liu. *Preoperative planning and real-time assisted navigation by three-dimensional individual digital model in partial nephrectomy with three-dimensional laparoscopic system.* International journal of computer assisted radiology and surgery, vol. 10, pages 1461–1468, 2015. (Cité en page 176.)

[Wang 2016a] Shu Wang, Shaoting Zhang, Wei Liu et Dimitris N Metaxas. *Visual Tracking via Reliable Memories.* arXiv preprint arXiv :1602.01887, 2016. (Cité en page 229.)

[Wang 2016b] Xue Wang, Jianbo Shi, Hyun Soo Park et Qing Wang. *Motion-based temporal alignment of independently moving cameras.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 11, pages 2344–2354, 2016. (Cité en page 129.)

[Wang 2017] Yubin Wang, Jinkai Shao, Xin Ma, Qingshan Du, Huijie Gong et Xu Zhang. *Robotic and open partial nephrectomy for complex renal tumors : a matched-pair comparison with a long-term follow-up.* World journal of urology, vol. 35, pages 73–80, 2017. (Cité en page 76.)

[Wang 2019] Qiang Wang, Li Zhang, Luca Bertinetto, Weiming Hu et Philip HS Torr. *Fast online object tracking and segmentation : A unifying approach.* In Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition, pages 1328–1338, 2019. (Cité en page 99.)

[Wang 2021a] Gang Wang et Yufei Chen. *Robust feature matching using guided local outlier factor.* Pattern Recognition, vol. 117, page 107986, 2021. (Cité en page 218.)

[Wang 2021b] Xiaoxia Wang, Ping Li, Yuchun Lv, Huifeng Xue, Tianxiang Xu, Yongzhao Du et Peizhong Liu. *Integration of global and local features for specular reflection inpainting in colposcopic images.* Journal of Healthcare Engineering, vol. 2021, no. 1, page 5401308, 2021. (Cité en page 132.)

[Wang 2023] Yiming Wang, Qian Huang, Chuanxu Jiang, Jiwen Liu, Mingzhou Shang et Zhuang Miao. *Video stabilization : A comprehensive survey.* Neurocomputing, vol. 516, pages 205–230, 2023. (Cité en page 216.)

[Wang 2024a] Lei Wang, Yanqun Na et Gang Zhu. *Application of holographic imaging in partial nephrectomy : a literature review.* Holistic Integrative Oncology, vol. 3, no. 1, page 6, 2024. (Cité en page 94.)

[Wang 2024b] Shuzhe Wang, Vincent Leroy, Yohann Cabon, Boris Chidlovskii et Jerome Revaud. *Dust3r : Geometric 3d vision made easy.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20697–20709, 2024. (Cité en page 190.)

[Wang 2024c] Yifan Wang, Xingyi He, Sida Peng, Dongli Tan et Xiaowei Zhou. *Efficient LoFTR : Semi-dense local feature matching with sparse-like speed.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 21666–21675, 2024. (Cité en page 146.)

[Wang 2025] Jianyuan Wang, Minghao Chen, Nikita Karaev, Andrea Vedaldi, Christian Rupprecht et David Novotny. *Vggt : Visual geometry grounded transformer*. In Proceedings of the Computer Vision and Pattern Recognition Conference, pages 5294–5306, 2025. (Cité en page 190.)

[Warren 2024] Alex Warren, Ke Xu, Jiaying Lin, Gary KL Tam et Rynson WH Lau. *Effective video mirror detection with inconsistent motion cues*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 17244–17252, 2024. (Cité en page 133.)

[Watanabe 2016] Jun Watanabe, Mitsuyoshi Ota, Yusuke Suwa, Atsushi Ishibe, Hidenobu Masui et Kaoru Nagahori. *Real-Time Indocyanine Green Fluorescence Imaging–Guided Complete Mesocolic Excision in Laparoscopic Flexural Colon Cancer Surgery*. Diseases of the Colon & Rectum, vol. 59, no. 7, pages 701–705, 2016. (Cité en page 48.)

[Watson 2024] Jason Paul Watson. *Distortion correction of endoscopic images in computer-assisted tele-operated surgery*, 2024. (Cité en page 37.)

[Wauthy 2023] Pierre Wauthy. Histoire de la chirurgie. 2023. (Cité en pages 8 et 9.)

[Wedge 2005] Daniel Wedge, Peter Kovesi et Du Huynh. *Trajectory based video sequence synchronization*. In Digital Image Computing : Techniques and Applications (DICTA'05), pages 13–13. IEEE, 2005. (Cité en page 129.)

[Wedge 2006] Daniel Wedge, Du Huynh et Peter Kovesi. *Motion guided video sequence synchronization*. In Asian Conference on Computer Vision, pages 832–841. Springer, 2006. (Cité en page 129.)

[Wedge 2007] Daniel Wedge, Du Huynh et Peter Kovesi. *Using space-time interest points for video sequence synchronization*. In Using Space-Time Interest Points for Video Sequence Synchronization, pages 190–194. MVA Conference Committee, 2007. (Cité en page 129.)

[Welch 2011] Ashley J Welch, Martin JC Van Gemert*et al.* Optical-thermal response of laser-irradiated tissue, volume 2. Springer, 2011. (Cité en page 123.)

[Wen 2025] Bowen Wen, Matthew Trepte, Joseph Aribido, Jan Kautz, Orazio Gallo et Stan Birchfield. *FoundationStereo : Zero-Shot Stereo Matching*. arXiv preprint arXiv :2501.09898, 2025. (Cité en pages 109, 157 et 191.)

[Westerman 2020] Mary E Westerman, Daniel D Shapiro, Christopher G Wood et Jose A Karam. *Neoadjuvant therapy for locally advanced renal cell carcinoma*. Urologic Clinics, vol. 47, no. 3, pages 329–343, 2020. (Cité en page 77.)

[Whitehead 2005] Anthony Whitehead, Robert Laganiere et Prosenjit Bose. *Temporal synchronization of video sequences in theory and in practice*. In 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1, volume 2, pages 132–137. IEEE, 2005. (Cité en pages 128 et 129.)

[Widya 2019] Aji Resindra Widya, Yusuke Monno, Masatoshi Okutomi, Sho Suzuki, Takuji Gotoda et Kenji Miki. *Whole stomach 3D reconstruction and frame*

localization from monocular endoscope video. IEEE Journal of Translational Engineering in Health and Medicine, vol. 7, pages 1–10, 2019. (Cité en page 219.)

[Wieschollek 2017] Patrick Wieschollek, Ido Freeman et Hendrik PA Lensch. *Learning robust video synchronization without annotations*. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), pages 92–100. IEEE, 2017. (Cité en page 129.)

[Wild 2016] Esther Wild, Dogu Teber, Daniel Schmid, Tobias Simpfendörfer, Michael Müller, Ann-Christin Baranski, Hannes Kenngott, Klaus Kopka et Lena Maier-Hein. *Robust augmented reality guidance with fluorescent markers in laparoscopic surgery*. International journal of computer assisted radiology and surgery, vol. 11, pages 899–907, 2016. (Cité en page 96.)

[Wiles 2021] Olivia Wiles, Sebastien Ehrhardt et Andrew Zisserman. *Co-attention for conditioned image matching*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 15920–15929, 2021. (Cité en page 146.)

[Wilson 2011] Mark R Wilson, Jamie M Poolton, Neha Malhotra, Karen Ngo, Elizabeth Bright et Rich SW Masters. *Development and validation of a surgical workload measure : the surgery task load index (SURG-TLX)*. World journal of surgery, vol. 35, pages 1961–1969, 2011. (Cité en pages 92 et 273.)

[Winkler 2013] Stefan Winkler et Dongbo Min. *Stereo/multiview picture quality : Overview and recent advances*. Signal Processing : Image Communication, vol. 28, no. 10, pages 1358–1373, 2013. (Cité en page 130.)

[Wold 1987] Svante Wold, Kim Esbensen et Paul Geladi. *Principal component analysis*. Chemometrics and intelligent laboratory systems, vol. 2, no. 1-3, pages 37–52, 1987. (Cité en page 97.)

[Wong 2020] Martin CS Wong, Junjie Huang, Jingxuan Wang, Paul SF Chan, Veeleah Lok, Xiao Chen, Colette Leung, Harry HX Wang, Xiang Qian Lao et Zhi-Jie Zheng. *Global, regional and time-trend prevalence of central obesity : a systematic review and meta-analysis of 13.2 million subjects*. European journal of epidemiology, vol. 35, pages 673–683, 2020. (Cité en page 68.)

[Woo 2018] Sanghyun Woo, Jongchan Park, Joon-Young Lee et In So Kweon. *Cbam : Convolutional block attention module*. In Proceedings of the European conference on computer vision (ECCV), pages 3–19, 2018. (Cité en page 132.)

[Wood 2018] Erika L Wood, Mehrad Adibi, Wei Qiao, Jessica Brandt, Miao Zhang, Pheroze Tamboli, Surena F Matin, Christopher G Wood et Jose A Karam. *Local tumor bed recurrence following partial nephrectomy in patients with small renal masses*. The Journal of urology, vol. 199, no. 2, pages 393–400, 2018. (Cité en pages 73 et 74.)

[Wu 2010] C Wu, B Jaramaz et SG Narasimhan. *A full geometric and photometric calibration method for oblique-viewing endoscopes*. Computer Aided Surgery, vol. 15, no. 1-3, pages 19–31, 2010. (Cité en page 110.)

[Wu 2020] Hao Wu, Patrick Judd, Xiaojie Zhang, Mikhail Isaev et Paulius Micikevicius. *Integer quantization for deep learning inference : Principles and empirical evaluation*. arXiv preprint arXiv :2004.09602, 2020. (Cité en page 243.)

[Wu 2023a] Xiwen Wu, Dehui Wang, Nan Xiang, Mingxin Pan, Fucang Jia, Jian Yang et Chihua Fang. *Augmented reality-assisted navigation system contributes to better intraoperative and short-time outcomes of laparoscopic pancreaticoduodenectomy : a retrospective cohort study*. International Journal of Surgery, vol. 109, no. 9, pages 2598–2607, 2023. (Cité en page 171.)

[Wu 2023b] Zhongqi Wu, Jianwei Guo, Chuanqing Zhuang, Jun Xiao, Dong-Ming Yan et Xiaopeng Zhang. *Joint specular highlight detection and removal in single images via Unet-Transformer*. Computational visual media, vol. 9, no. 1, pages 141–154, 2023. (Cité en page 132.)

[Xia 2019] Wenyao Xia, Elvis CS Chen, Stephen E Pautler et Terry M Peters. *A global optimization method for specular highlight removal from a single image*. IEEE Access, vol. 7, pages 125976–125990, 2019. (Cité en page 132.)

[Xia 2022] Yifan Xia et Jiayi Ma. *Locality-guided global-preserving optimization for robust feature matching*. IEEE Transactions on Image Processing, vol. 31, pages 5093–5108, 2022. (Cité en page 218.)

[Xiao 2024] Yuxi Xiao, Qianqian Wang, Shangzhan Zhang, Nan Xue, Sida Peng, Yujun Shen et Xiaowei Zhou. *Spatialtracker : Tracking any 2d pixels in 3d space*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20406–20417, 2024. (Cité en page 224.)

[Xiaolei 2025] Shi Xiaolei, Guo Fei, Zhi Chao, Xiao Guangan, Zhao Lin, Wang Yang, Zhang Wei, Xiao Chengwu, Wu Zhenjie, Wang Linhui*et al.* *Artificial intelligence based augmented reality navigation in minimally invasive partial nephrectomy*. Urology, 2025. (Cité en page 98.)

[Xie 2014] Qi Xie, Yong Liu, Rong Xiong et Jian Chu. *Real-time accurate ball trajectory estimation with "asynchronous" stereo camera system for humanoid Ping-Pong robot*. In 2014 IEEE International Conference on Robotics and Automation (ICRA), pages 6212–6217. IEEE, 2014. (Cité en page 129.)

[Xie 2024a] Weixing Xie, Junfeng Yao, Xianpeng Cao, Qiqin Lin, Zerui Tang, Xiao Dong et Xiaohu Guo. *Surgicalgaussian : Deformable 3d gaussians for high-fidelity surgical scene reconstruction*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 617–627. Springer, 2024. (Cité en page 190.)

[Xie 2024b] Zhifeng Xie, Sen Wang, Qiucheng Yu, Xin Tan et Yuan Xie. *CSFwinformer : Cross-space-frequency window transformer for mirror detection*. IEEE Transactions on Image Processing, 2024. (Cité en page 132.)

[Xu 2005] Shu-Chang Xu, Xiuzi Ye, Yin Wu et Sanyuan Zhang. *Highlight detection and removal based on chromaticity.* In Image Analysis and Recognition : Second International Conference, ICIAR 2005, Toronto, Canada, September 28-30, 2005. Proceedings 2, pages 199–206. Springer, 2005. (Cité en page 132.)

[Xu 2023] Jiayi Xu, Gaspard Thevenon, Timothee Chabat, Matthew McCormick, Forrest Li, Tom Birdsong, Ken Martin, Yueh Lee et Stephen Aylward. *Interactive, in-browser cinematic volume rendering of medical images.* Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization, vol. 11, no. 4, pages 1019–1026, 2023. (Cité en page 44.)

[Xu 2024a] Ke Xu, Tsun Wai Siu et Rynson WH Lau. *ZOOM : learning video mirror detection with extremely-weak supervision.* In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 6315–6323, 2024. (Cité en page 133.)

[Xu 2024b] Mengya Xu, Ziqi Guo, An Wang, Long Bai et Hongliang Ren. *A review of 3d reconstruction techniques for deformable tissues in robotic surgery.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 157–167. Springer, 2024. (Cité en page 190.)

[Yamaguchi 2004] Tetsuzo Yamaguchi, Masahiko Nakamoto, Yoshinobu Sato, Kozo Konishi, Makoto Hashizume, Nobuhiko Sugano, Hideki Yoshikawa et Shinichi Tamura. *Development of a camera model and calibration procedure for oblique-viewing endoscopes.* Computer Aided Surgery, vol. 9, no. 5, pages 203–214, 2004. (Cité en page 110.)

[Yanagisawa 2022] Takafumi Yanagisawa, Keiichiro Mori, Tatsushi Kawada, Reza Sari Motlagh, Hadi Mostafaei, Fahad Quhal, Ekaterina Laukhtina, Pawel Rajwa, Abdulmajeed Aydh, Frederik König*et al. Differential efficacy of ablation therapy versus partial nephrectomy between clinical T1a and T1b renal tumors : a systematic review and meta-analysis.* In Urologic Oncology : Seminars and Original Investigations, volume 40, pages 315–330. Elsevier, 2022. (Cité en page 77.)

[Yang 2014] Guanyu Yang, Jinjin Gu, Yang Chen, Wangyan Liu, Lijun Tang, Huazhong Shu et Christine Toumoulin. *Automatic kidney segmentation in CT images based on multi-atlas image registration.* In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 5538–5541. IEEE, 2014. (Cité en page 45.)

[Yang 2019] Xin Yang, Haiyang Mei, Ke Xu, Xiaopeng Wei, Baocai Yin et Rynson WH Lau. *Where is my mirror ?* In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 8809–8818, 2019. (Cité en page 132.)

[Yang 2024a] Lihe Yang, Bingyi Kang, Zilong Huang, Xiaogang Xu, Jiashi Feng et Hengshuang Zhao. *Depth anything : Unleashing the power of large-scale unla-*

beled data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10371–10381, 2024. (Cité en page 109.)

[Yang 2024b] Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng et Hengshuang Zhao. *Depth anything v2*. Advances in Neural Information Processing Systems, vol. 37, pages 21875–21911, 2024. (Cité en page 109.)

[Yano 2018] Motoyo Yano, Kathryn J Fowler, Santip Srisuwan, Amber Salter et Cary L Siegel. *Quantitative multiparametric MR analysis of small renal lesions : correlation with surgical pathology*. Abdominal Radiology, vol. 43, pages 3390–3399, 2018. (Cité en pages 72 et 74.)

[Yao 2010] Rui Yao, Yilun Wu, Wei Yang, Xiaolin Lin, Shidan Chen et Su Zhang. *Specular reflection detection on gastroscopic images*. In 2010 4th International Conference on Bioinformatics and Biomedical Engineering, pages 1–4. IEEE, 2010. (Cité en page 132.)

[Ye 2016] Menglong Ye, Stamatia Giannarou, Alexander Meining et Guang-Zhong Yang. *Online tracking and retargeting with applications to optical biopsy in gastrointestinal endoscopic examinations*. Medical image analysis, vol. 30, pages 144–157, 2016. (Cité en page 226.)

[Ye 2017] Menglong Ye, Edward Johns, Ankur Handa, Lin Zhang, Philip Pratt et Guang-Zhong Yang. *Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery*. arXiv preprint arXiv :1705.08260, 2017. (Cité en page 226.)

[Yeong 2021] De Jong Yeong, Gustavo Velasco-Hernandez, John Barry et Joseph Walsh. *Sensor and sensor fusion technology in autonomous vehicles : A review*. Sensors, vol. 21, no. 6, page 2140, 2021. (Cité en page 104.)

[Yi 2016] Kwang Moo Yi, Eduard Trulls, Vincent Lepetit et Pascal Fua. *Lift : Learned invariant feature transform*. In European conference on computer vision, pages 467–483. Springer, 2016. (Cité en page 217.)

[Yi 2018] Kwang Moo Yi, Eduard Trulls, Yuki Ono, Vincent Lepetit, Mathieu Salzmann et Pascal Fua. *Learning to find good correspondences*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2666–2674, 2018. (Cité en page 219.)

[Yin 2022] Liqiang Yin, Ruize Han, Wei Feng et Song Wang. *Self-supervised human pose based multi-camera video synchronization*. In Proceedings of the 30th ACM International Conference on Multimedia, pages 1739–1748, 2022. (Cité en page 128.)

[Yip 2012] Michael C Yip, David G Lowe, Septimiu E Salcudean, Robert N Rohling et Christopher Y Nguan. *Tissue tracking and registration for image-guided surgery*. IEEE transactions on medical imaging, vol. 31, no. 11, pages 2169–2182, 2012. (Cité en page 221.)

[Zadeh 2023] Sabrina Madad Zadeh, Tom François, Aurélie Comptour, Michel Canis, Nicolas Bourdel et Adrien Bartoli. *SurgAI3. 8K : a labeled dataset of*

gynecologic organs in laparoscopy with application to automatic augmented reality surgical guidance. Journal of Minimally Invasive Gynecology, vol. 30, no. 5, pages 397–405, 2023. (Cité en page 244.)

[Zaid 2017] Harras B Zaid, William P Parker, Nida S Safdar, Boris Gershman, Patricia J Erwin, M Hassan Murad, Stephen A Boorjian, Brian A Costello, R Houston Thompson et Bradley C Leibovich. *Outcomes following complete surgical metastasectomy for patients with metastatic renal cell carcinoma : a systematic review and meta-analysis.* The Journal of Urology, vol. 197, no. 1, pages 44–49, 2017. (Cité en page 77.)

[Zampokas 2022] Giorgos Zampokas, Georgia Peleka, Konstantinos Tsiolis, Angeliki Topalidou-Kyniazopoulou, Ioannis Mariolis et Dimitrios Tzovaras. *Real-time stereo reconstruction of intraoperative scene and registration to preoperative 3D models for augmenting surgeons' view during RAMIS.* Medical Physics, vol. 49, no. 10, pages 6517–6526, 2022. (Cité en page 97.)

[Zeisl 2012] Bernhard Zeisl, Kevin Köser et Marc Pollefeys. *Viewpoint invariant matching via developable surfaces.* In Computer Vision–ECCV 2012. Workshops and Demonstrations : Florence, Italy, October 7-13, 2012, Proceedings, Part II 12, pages 62–71. Springer, 2012. (Cité en page 230.)

[Zeisl 2013] Bernhard Zeisl, Kevin Koser et Marc Pollefeys. *Automatic registration of RGB-D scans via salient directions.* In Proceedings of the IEEE international conference on computer vision, pages 2808–2815, 2013. (Cité en page 230.)

[Zeng 2021] Shaohua Zeng, Yu Zhou, Min Wang, Hui Bao, Yanqun Na et Tiejun Pan. *Holographic reconstruction technology used for intraoperative real-time navigation in robot-assisted partial nephrectomy in patients with renal tumors : a single center study.* Translational Andrology and Urology, vol. 10, no. 8, page 3386, 2021. (Cité en page 94.)

[Zeng 2025] Yunfan Zeng, Guangyan Cai et Shuang Zhao. *A Survey on Physics-based Differentiable Rendering.* arXiv preprint arXiv :2504.01402, 2025. (Cité en page 226.)

[Zhang 1999] Zhengyou Zhang. *Flexible camera calibration by viewing a plane from unknown orientations.* In Proceedings of the seventh ieee international conference on computer vision, volume 1, pages 666–673. Ieee, 1999. (Cité en page 110.)

[Zhang 2002] Zhengyou Zhang. *A flexible new technique for camera calibration.* IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 11, pages 1330–1334, 2002. (Cité en page 107.)

[Zhang 2013] Tianzhu Zhang, Bernard Ghanem, Si Liu et Narendra Ahuja. *Robust visual tracking via structured multi-task sparse learning.* International journal of computer vision, vol. 101, pages 367–383, 2013. (Cité en page 229.)

[Zhang 2015] Tianzhu Zhang, Si Liu, Narendra Ahuja, Ming-Hsuan Yang et Bernard Ghanem. *Robust visual tracking via consistent low-rank sparse learning.* In-

ternational Journal of Computer Vision, vol. 111, pages 171–190, 2015. (Cité en page 229.)

[Zhang 2016] Zhiling Zhang, Juping Zhao, Wen Dong, Eric Remer, Jianbo Li, Sevag Demirjian, Joseph Zabell et Steven C Campbell. *Acute kidney injury after partial nephrectomy : role of parenchymal mass reduction and ischemia and impact on subsequent functional recovery.* European urology, vol. 69, no. 4, pages 745–752, 2016. (Cité en pages 79 et 81.)

[Zhang 2017] Shaojing Zhang, Zhankui Jia, Zechen Yan et Jinjian Yang. *Consumption of fruits and vegetables and risk of renal cell carcinoma : a meta-analysis of observational studies.* Oncotarget, vol. 8, no. 17, page 27892, 2017. (Cité en page 69.)

[Zhang 2019a] Jiahui Zhang, Dawei Sun, Zixin Luo, Anbang Yao, Lei Zhou, Tianwei Shen, Yurong Chen, Long Quan et Hongen Liao. *Learning two-view correspondences and geometry using order-aware network.* In Proceedings of the IEEE/CVF international conference on computer vision, pages 5845–5854, 2019. (Cité en page 219.)

[Zhang 2019b] Xiaohui Zhang, Junchen Wang, Tianmiao Wang, Xuquan Ji, Yu Shen, Zhen Sun et Xuebin Zhang. *A markerless automatic deformable registration framework for augmented reality navigation of laparoscopy partial nephrectomy.* International journal of computer assisted radiology and surgery, vol. 14, pages 1285–1294, 2019. (Cité en pages 97, 177 et 192.)

[Zhang 2020a] Shuai Zhang, Liang Zhao, Shoudong Huang, Menglong Ye et Qi Hao. *A template-based 3D reconstruction of colon structures and textures from stereo colonoscopic images.* IEEE Transactions on Medical Robotics and Bionics, vol. 3, no. 1, pages 85–95, 2020. (Cité en page 226.)

[Zhang 2020b] Xiaohui Zhang, Tianmiao Wang, Xuebin Zhang, Yinghao Zhang et Junchen Wang. *Assessment and application of the coherent point drift algorithm to augmented reality surgical navigation for laparoscopic partial nephrectomy.* International journal of computer assisted radiology and surgery, vol. 15, no. 6, pages 989–999, 2020. (Cité en page 98.)

[Zhang 2021a] Haijie Zhang, Fu Yin, Liyang Yang, Anqi Qi, Weiwei Cui, Shanshan Yang et Ge Wen. *Computed Tomography Image under Three-Dimensional Reconstruction Algorithm Based in Diagnosis of Renal Tumors and Retroperitoneal Laparoscopic Partial Nephrectomy.* Journal of Healthcare Engineering, vol. 2021, no. 1, page 3066930, 2021. (Cité en page 94.)

[Zhang 2021b] Hongwen Zhang, Yating Tian, Xinchi Zhou, Wanli Ouyang, Yebin Liu, Limin Wang et Zhenan Sun. *Pymaf : 3d human pose and shape regression with pyramidal mesh alignment feedback loop.* In Proceedings of the IEEE/CVF international conference on computer vision, pages 11446–11456, 2021. (Cité en page 178.)

[Zhang 2021c] Shuai Zhang, Liang Zhao, Shoudong Huang, Ruibin Ma, Boni Hu et Qi Hao. *3D reconstruction of deformable colon structures based on preopera-*

tive model and deep neural network. In 2021 IEEE International Conference on Robotics and Automation (ICRA), pages 1875–1881. IEEE, 2021. (Cité en pages 222 et 226.)

[Zhang 2022a] Kai Zhang, Lei Wang, Ying Sun, Wei Wang, Shuaiying Hao, Hongbo Li, Jingjing Lu, Derek A O'Reilly, Yanqun Na et Gang Zhu. *Combination of holographic imaging with robotic partial nephrectomy for renal hilar tumor treatment*. International Urology and Nephrology, vol. 54, no. 8, pages 1837–1844, 2022. (Cité en page 94.)

[Zhang 2022b] Xiaohui Zhang, Elsa-Marie Otoo, Yubo Fan, Chunjing Tao, Tianmiao Wang et Kawal Rhode. *Autostereoscopic 3D augmented reality navigation for laparoscopic surgery : a preliminary assessment*. IEEE Transactions on Biomedical Engineering, vol. 70, no. 4, pages 1413–1421, 2022. (Cité en page 131.)

[Zhang 2023a] Hongwen Zhang, Yating Tian, Yuxiang Zhang, Mengcheng Li, Liang An, Zhenan Sun et Yebin Liu. *Pymaf-x : Towards well-aligned full-body model regression from monocular images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 10, pages 12287–12303, 2023. (Cité en page 178.)

[Zhang 2023b] Xiaohui Zhang, Xuquan Ji, Junchen Wang, Yubo Fan et Chunjing Tao. *Renal surface reconstruction and segmentation for image-guided surgical navigation of laparoscopic partial nephrectomy*. Biomedical Engineering Letters, vol. 13, no. 2, pages 165–174, 2023. (Cité en page 97.)

[Zhang 2024a] Han Zhang, Xiaoxiao Ran et Wujie Zhou. *Self-knowledge distillation-based staged extraction and multiview collection network for RGB-D mirror segmentation*. IEEE Signal Processing Letters, 2024. (Cité en page 133.)

[Zhang 2024b] Hongyu Zhang, Fukai Li, Mengyuan Jing, Huaze Xi, Yali Zheng et Jianli Liu. *Nomogram combining pre-operative clinical characteristics and spectral CT parameters for predicting the WHO/ISUP pathological grading in clear cell renal cell carcinoma*. Abdominal Radiology, vol. 49, no. 4, pages 1185–1193, 2024. (Cité en page 71.)

[Zhang 2025] Bowei Zhang, Lei Ke, Adam W Harley et Katerina Fragkiadaki. *TAPIP3D : Tracking Any Point in Persistent 3D Geometry*. arXiv preprint arXiv :2504.14717, 2025. (Cité en pages 224 et 257.)

[Zhao 2009] Qi Zhao et Yan Qiu Chen. *High-precision synchronization of video cameras using a single binary light source*. Journal of Electronic Imaging, vol. 18, no. 4, pages 040501–040501, 2009. (Cité en page 127.)

[Zhao 2019] Chen Zhao, Zhiguo Cao, Chi Li, Xin Li et Jiaqi Yang. *NM-Net : Mining reliable neighbors for robust feature correspondences*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 215–224, 2019. (Cité en page 219.)

[Zhao 2022a] Shang Zhao, Ce Wang, Qiyuan Wang, Yanzhe Liu et S Kevin Zhou. *3D endoscopic depth estimation using 3D surface-aware constraints.* arXiv preprint arXiv :2203.02131, 2022. (Cité en page 219.)

[Zhao 2022b] Xiaoming Zhao, Xingming Wu, Jinyu Miao, Weihai Chen, Peter CY Chen et Zhengguo Li. *Alike : Accurate and lightweight keypoint detection and descriptor extraction.* IEEE Transactions on Multimedia, vol. 25, pages 3101–3112, 2022. (Cité en page 145.)

[Zhao 2022c] Zichen Zhao et Jin Gu. *Open surgery in the era of minimally invasive surgery.* Chinese Journal of Cancer Research, vol. 34, no. 1, page 63, 2022. (Cité en pages 17 et 18.)

[Zhao 2023a] Wenyi Zhao, Catherine J. Mohr et Simon P. DiMaio. *Stereo imaging system with automatic disparity adjustment for displaying close range objects*, 2023. (Cité en page 37.)

[Zhao 2023b] Xiaoming Zhao, Xingming Wu, Weihai Chen, Peter CY Chen, Qingsong Xu et Zhengguo Li. *Aliked : A lighter keypoint and descriptor extraction network via deformable transformation.* IEEE Transactions on Instrumentation and Measurement, vol. 72, pages 1–16, 2023. (Cité en page 145.)

[Zheng 2019] Yongchang Zheng, Huayu Yang, Henan Wang, Kai Kang, Wenkai Zhang, Gang Ma et Shunda Du. *Fluorescence-guided surgery in cancer treatment : current status and future perspectives.* Annals of Translational Medicine, vol. 7, no. Suppl 1, 2019. (Cité en page 48.)

[Zheng 2025] Yuchao Zheng, Jianing Zhang, Guochen Ning et Hongen Liao. *SurGSplat : Progressive Geometry-Constrained Gaussian Splatting for Surgical Scene Reconstruction.* arXiv preprint arXiv :2506.05935, 2025. (Cité en page 190.)

[Zholus 2025] Artem Zholus, Carl Doersch, Yi Yang, Skanda Koppula, Viorica Patraucean, Xu Owen He, Ignacio Rocco, Mehdi SM Sajjadi, Sarath Chandar et Ross Goroshin. *TAPNext : Tracking Any Point (TAP) as Next Token Prediction.* arXiv preprint arXiv :2504.05579, 2025. (Cité en page 225.)

[Zhong 2017] Fu-Ping Zhong, Yu-Jun Zhang, Ying Liu et Shu-Bing Zou. *Prognostic impact of surgical margin in patients with hepatocellular carcinoma : a meta-analysis.* Medicine, vol. 96, no. 37, page e8043, 2017. (Cité en page 169.)

[Zhou 2018] Qian-Yi Zhou, Jaesik Park et Vladlen Koltun. *A Modern Library for 3 D Data Processing.* In 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 110–113, 2018. (Cité en page 184.)

[Zhou 2019a] Haoyin Zhou et Jayender Jagadeesan. *Real-time dense reconstruction of tissue surface from stereo optical video.* IEEE transactions on medical imaging, vol. 39, no. 2, pages 400–412, 2019. (Cité en page 190.)

[Zhou 2019b] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh et Jianming Liang. *Unet++ : Redesigning skip connections to exploit multiscale features in image segmentation.* IEEE transactions on medical imaging, vol. 39, no. 6, pages 1856–1867, 2019. (Cité en page 262.)

[Zhou 2021] Haoyin Zhou et Jagadeesan Jayender. *Emdq-slam : Real-time high-resolution reconstruction of soft tissue surface from stereo laparoscopy videos.* In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 331–340. Springer, 2021. (Cité en page 223.)

[Zhou 2023] Wujie Zhou, Yuqi Cai, Liting Zhang, Weiqing Yan et Lu Yu. *UTLNet : Uncertainty-aware transformer localization network for RGB-depth mirror segmentation.* IEEE Transactions on Multimedia, vol. 26, pages 4564–4574, 2023. (Cité en page 133.)

[Zhou 2024] Wujie Zhou, Yuqi Cai et Fangfang Qiang. *Morphology-guided network via knowledge distillation for RGB-D mirror segmentation.* IEEE Transactions on Intelligent Transportation Systems, 2024. (Cité en page 133.)

[Zhou 2025] Rulin Zhou, Wenlong He, An Wang, Qiqi Yao, Haijun Hu, Jiankun Wang*et al.* *Endo-TTAP : Robust Endoscopic Tissue Tracking via Multi-Facet Guided Attention and Hybrid Flow-point Supervision.* arXiv preprint arXiv :2503.22394, 2025. (Cité en page 225.)

[Zhu 2017] Jun-Yan Zhu, Taesung Park, Phillip Isola et Alexei A Efros. *Unpaired image-to-image translation using cycle-consistent adversarial networks.* In Proceedings of the IEEE international conference on computer vision, pages 2223–2232, 2017. (Cité en page 223.)

[Zhu 2024] Meikai Zhu, Zhifeng Liu, Yongheng Zhou, Zhiwen Jiang, Shouzhen Chen, Wenfu Wang, Benkang Shi et Yaofeng Zhu. *Effects of neoadjuvant VEGF-TKI treatment on surgery for renal cell carcinoma : A systematic review and meta-analysis.* Oncology Letters, vol. 27, no. 4, page 162, 2024. (Cité en page 77.)

[Ziegler 2012] Andrew Ziegler, Eric Christiansen, David Kriegman et Serge Belongie. *Locally uniform comparison image descriptor.* Advances in Neural Information Processing Systems, vol. 25, 2012. (Cité en page 145.)

[Zimmerman-Moreno 2006] Gali Zimmerman-Moreno et Hayit Greenspan. *Automatic detection of specular reflections in uterine cervix images.* In Medical Imaging 2006 : Image Processing, volume 6144, pages 2037–2045. SPIE, 2006. (Cité en page 132.)

[Zini 2008] Laurent Zini, Paul Perrotte, Claudio Jeldres, Umberto Capitanio, Daniel Pharand, Philippe Arjane, Steven Lapointe, Francesco Montorsi, Jean-Jacques Patard et Pierre I Karakiewicz. *Nephrectomy improves the survival of patients with locally advanced renal cell carcinoma.* BJU international, vol. 102, no. 11, pages 1610–1614, 2008. (Cité en page 77.)

[Znaor 2015] Ariana Znaor, Joannie Lortet-Tieulent, Mathieu Laversanne, Ahmedin Jemal et Freddie Bray. *International variations and trends in renal cell carcinoma incidence and mortality.* European urology, vol. 67, no. 3, pages 519–530, 2015. (Cité en page 69.)