Towards Live Monocular 3D Laparoscopy using Shading and Specularity Information

Toby Collins and Adrien Bartoli

ALCoV-ISIT
Université d'Auvergne, Clermont-Ferrand, France
Toby.Collins@gmail.com, Adrien.Bartoli@gmail.com

Abstract. We present steps towards the first real-time system for computing and visualising 3D surfaces viewed in live monocular laparoscopy video. Our method is based on estimating 3D shape using shading and specularity information, and seeks to push current Shape from Shading (SfS) boundaries towards practical, reliable reconstruction. We present an accurate method to model any laparoscope's light source, and a highly-parallelised SfS algorithm that outperforms the fastest current method. We give details of its GPU implementation that facilitates realtime performance of an average frame-rate of 23fps. Our system also incorporates live 3D visualisation with virtual stereoscopic synthesis. We have evaluated using real laparoscopic videos with ground-truth, and we present the successful in-vivo reconstruction of the human uterus. We however draw the conclusion that the shading cue alone is insufficient to reliably handle arbitrary laparoscopic images.

Fig. 1: Our current live monocular 3D laparoscopy prototype system. (Left): Two screens showing (white) the raw 2D video feed and (black) the reconstructed 3D with an active-shutter display. (Remaining images): System in use.

1 Introduction

An important computer vision task in Minimally Invasive Surgery (MIS) is to recover the 3D structure of organs and tissues viewed in laparoscopic images and videos. A general solution to this has important applications, including depth estimation and perception, enhanced intra-operative surgical guidance, motion estimation and compensation, pre-operative data registration and novel-view synthesis. Currently, state-of-the-art methods for acquiring 3D information differ along two main axes; (*i*) the sensor hardware used to estimate 3D, and (*ii*) the visual cue used to infer 3D shape. 3D reconstruction has been attempted previously by modifying traditional monocular laparoscopes; these include stereo laparoscopes [14, 4] and active 3D methods based on structured light [1] and Time-of-Flight cameras [12]. These simplify the reconstruction problem, yet come at the

price of additional intra-operative hardware. Furthermore, these have not been shown to work with a great degree of accuracy in practice. Stereo laparoscopes can also be used purely for 3D visualisation. However, resolving the difference between the camera's convergence angles and the user's eyes prove to be the limitation [8]. By contrast, monocular methods require no hardware modification and aim to estimate 3D shape from 2D data. Virtual stereo images can easily be visualised which match the convergence angle of the user's eyes. However, the 3D reconstruction problem is considerably more difficult, and remains an open challenge. Shape from Motion (SfM) is one monocular method gaining some ground. The most successful are based on realtime Simultaneous Localisation And Mapping (SLAM) [7, 3]. However, these require correspondence estimation; a very difficult task in laparoscopic images. Also SLAM assumes the 3D scene is either rigid or adheres to a very strict deformation model, which is mostly unrealistic during intervention. Shape from Shading (SfS) is another monocular method based on the relationship between 3D geometry, surface reflectance and scene illumination. It is a strong contender for monocular 3D laparoscopy since it (*i*) requires no correspondence, (*ii*) requires only a single input image and (*iii*) in laparoscopy the light conditions are highly controlled. However, SfS is a *weakly constrained problem*, and real conditions in laparoscopy often violate its core assumptions. Our overarching research goal is to answer the following two questions: *1. Is SfS a viable method for monocular 3D laparoscopy? 2. Is SfS sufficient on its own, or must it be complemented by other 3D cues?* In parallel to answering these questions, we have been developing a live (*i.e.* realtime) SfS-based 3D reconstruction/visualisation system (Fig. 1).

In this paper we extend the boundaries of SfS by improving some of the various modelling and computation aspects. We present a new way to accurately model a laparoscopic light source with what we call the *Nonparametric Light Model* (NLM). We show how to calibrate this easily and how it is incorporated into SfS. We also present a highly-parallelised shape estimation algorithm, which facilitates realtime 3D reconstruction. What we do not do is claim to have solved SfS for laparoscopy. There exist several open modelling, optimisation and practical challenges not addressed. We clearly state these in §1.2, to help guide other researches towards completing the problem.

1.1 Problem Statement

We briefly describe here the SfS image irradiance equation, which is the basis for all SfS methods. We then summarise its various instantiations from the state-of-the-art. Let us define a 3D surface S that is parameterized by the function $f(u,v) : \mathbb{R}^2 \rightarrow \mathbb{R}^3$. This maps some point (u,v) on the surface's domain into the camera's 3D coordinate frame. The Lambertian model predicts the *image irradiance* (the amount of light hitting the camera's CCD) according to:

$$g(\hat{I}(\psi(f(u,v)))) = \alpha(u,v)\mathcal{L}(f(u,v)) \cdot n(u,v) + \varepsilon \tag{1}$$

Here, the function $g : \mathbb{R} \rightarrow \mathbb{R}$ denotes the camera's response function, which converts pixel intensities into image irradiance. $\psi(f(u,v)) = (x,y)$ denotes

the 3D-to-2D camera projection process. $\hat{I}(x,y)$ denotes the measured pixel intensity at pixel (x,y). $\alpha(u,v)$ denotes the surface albedo and $n(u,v)$ denotes the surface normal. ε denotes pixel measurement noise. The illumination vector $\mathcal{L}(f(u,v)) = [l_x, l_y, l_z] : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ models the illumination as a directed ray of light. In classic SfS (u,v) spans a closed region in the input image: $\Omega \in \mathbb{R}^2$. The surface function $f(u,v)$ is then determined by the *depth function* $d(u,v) : \mathbb{R}^3 \rightarrow \mathbb{R}$. For perspective cameras, this is given by $f(u,v) = d(u,v)\mathbf{K}^{-1}(u,v,1)^\top$, where $\mathbf{K}$ is the matrix of camera intrinsics. The goal of SfS is to estimate $d(u,v)$ given intensity measurements at each pixel using (1). This is given by:

$$d^\star(u,v) = \underset{d(u,v)}{\arg\min} \int_\Omega (\alpha(u,v)\mathcal{L}(f(u,v)) \cdot n(u,v) - g(\hat{I}(u,v)))^2 dudv \qquad (2)$$

1.2 Solving SfS: State Of The Art

All SfS methods attempt to solve Problem (2), yet this is a highly non-trivial, often ill-posed problem. The specifics of an SfS method can be broken into three key components. These are: (*i*) **Modelling** the image formation process, (*ii*) making **scene assumptions** about the 3D environment and (*iii*) **3D computation** via optimisation. In Fig. (2) we present these three key SfS components. The table's second row summarises how they have been instantiated by recent works. In this paper we attempt to push forward the boundaries of SfS for 3D laparoscopy, but which permit effective realtime optimisation. Our core contributions are represented clearly in green.

	Modelling				Scene Assumptions		3D Computation	
	Camera Response Model	Illumination Model	Reflectance Model	Projection Model	Surface Albedo	Surface Discontinuities	Processing Time	Optimality
State Of The Art	Known & Constant	PLM & Constant	Lambertian	Perspective & Constant	Known & Constant	No Discontinuities	Not realtime*	Local&Global
Proposed Developments	Known & Constant	Accurate & Constant	Lambertian & specularities	Perspective & Constant	Small Improvement	No Discontinuities	Realtime	Local

Fig. 2: The three key components of SfS. Second row: current state-of-the-art. Third row: Our proposed developments specifically targeted at laparoscopy.

State-of-the-art methods require the camera response function to be known. There are no methods which can simultaneously adjust the response function in a video and perform SfS. This however is needed if the camera's exposure or shutter speed changes. For the illumination model, the Point Light Model (PLM) has been proposed for endoscopes [11, 18]. This uses the inverse-squared light falloff model. The illumination brightness must be kept constant throughout a video sequence. Regarding the reflectance model, all SfS methods applied to endoscopy have used Lambertian reflectance [13, 19, 18, 15, 9]. This makes the modelling problem less challenging, but is less accurate than for example the full BRDF. Regarding the camera's projection model, the perspective model was proposed in [10, 16]. However, no existing method can handle unknown and changing camera intrinsics in a video, caused by zooming, for example. The scene assumptions in SfS nearly always involve fixed, constant and known surface albedo. Some recent progress towards piecewise-constant albedo has been proposed in [2] using

learned natural image statistics. Since most SfS methods provide constraints on surface normals, they cannot handle discontinuous surfaces. Regarding the 3D computation, with the exception of Tsai and Shai (TS) [17] (noted by *), the SfS methods do not run in realtime on today's hardware. However TS uses models which are poor in laparoscopic conditions, such as an orthographic camera and the Distant Light Model (DLM). SfS methods can also be broken into locally or globally-optimal [19] solutions. Global solutions are however slow to compute.

2 Accurately Modelling a Laparoscope's Light Source

In this section we present the Nonparametric Light Model (NLM). In contrast to the DLM and PLM, which are poor approximations to the laparoscope's light source, one in fact cannot do better than the NLM. This is because it can model light functions of arbitrary complexity. We are however motivated by practical considerations regarding easy calibration. We have developed a one-time, fully automatic method that requires only multiple views of a planar calibration target. We use the NLM to compute $\mathcal{L}$ as a spatially-varying function de-

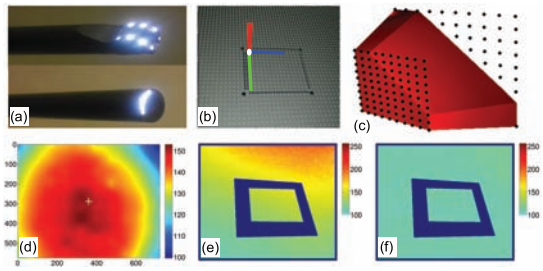

Fig. 3: Nonparametric Light Model: Calibration, modelling and accuracy.

fined at every 3D point within the operational volume of the laparoscope. It uses the premise that the illumination can be *locally modelled* at any 3D point $\mathbf{p} = (x, y, z)^\top$ by a local proximal light source. That is, $\mathcal{L}(\mathbf{p}) = [l_x, l_y, l_z]$, with $[l_x, l_y, l_z]$ being a local light vector of power $\|l_x, l_y, l_z\|_2$ and direction $[l_x, l_y, l_z] / \|l_x, l_y, l_z\|_2$. This model generalises both the DLM and PLM. The advantage of the NLM is that it can model well any of the various laparoscope light sources. Fig. 3(a) shows two examples we have modelled.

2.1 Calibrating the Nonparametric Light Model

To determine the function $\mathcal{L}$ the NLM must first be *calibrated*. We do this by using (1) and a ground-truth calibration object whose depths, normals and albedos are known. For convenience a planar calibration object is chosen with a tracking target printed on its surface. With this its 3D pose can be estimated robustly using well known methods [5] (Fig. 3(b)). Suppose the rigid transform $\mathbf{M} = [\mathbf{R}\,\mathbf{t}]$ maps the plane into the camera's coordinate frame. The surface function is then given by $f(u, v) = \left[\mathbf{R}_1\ \mathbf{R}_2\right][u, v]^\top + \mathbf{t}$ and $n(u, v) = \mathbf{R}_3$, where $\mathbf{R}_i$ denotes the i^{th} column of $\mathbf{R}$. Now that we have the surface function, at each pixel we obtain

a sparse collection of linear constraints on $\mathcal{L}(\mathbf{p})$ according to (1). Our method for calibration involves recording a video of the plane as it moves through the laparoscope's operational 3D volume. At each frame, we accumulate constraints at slices of the volume, and over the video we acquire enough data to learn $\mathcal{L}$. Importantly, the plane must be recorded at multiple orientations (a minimum of 3). This is because there are three unknowns (l_x, l_y and l_z) at any given point.

In practice we can expect a finite number of noisy measurements. Denote these by the set $\{\hat{\mathbf{I}}_i, \mathbf{n}_i, \mathbf{p}_i\}$. However, to estimate the continuous function $\mathcal{L}$, a unique solution can be found using regularised function approximation. We propose using the 3D Thin Plate Spline (TPS) which has several desirable properties for us. It is globally smooth, easily computable, and contains the least possible nonlinear component to achieve the function approximation. The 3D TPS defines a set of n 3D control points that we position throughout the volume (Fig. 3(c) shows control points at the front and back planes of the 3D volume). The TPS model takes the form: $\mathcal{L}(\mathbf{p}) = \mathbf{a}(\mathbf{p})[\mathbf{S}^\top, \mathbf{1}]^\top$, where $\mathbf{a}(\mathbf{p})$ is a vector which depends locally on $\mathbf{p}$. The function is parameterised by an $n \times 3$ control matrix $\mathbf{S}$. Calibrating the NLM involves determining $\mathbf{S}$. We pose this as minimising the following quadratic least squares objective function:

$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S}} \sum_i \left(\alpha_i \mathbf{a}_i [\mathbf{S}^\top, \mathbf{1}]^\top \mathbf{n}_i - g(\hat{I}_i)\right)^2 \tag{3}$$

Since $\mathbf{n}_i$ is expected to be noisy, we have a least squares problem with Errors In Variables. Thus we solve (3) using Total Least Squares. Once estimated, $\mathcal{L}(\mathbf{p})$ can be evaluated everywhere with $\mathcal{L}(x, y, z) = \mathbf{a}(\mathbf{p})[\hat{\mathbf{S}}^\top, \mathbf{1}]^\top$. Figs. 3(c,d) shows the NLM for the laparoscope in Figs. 3(a-bottom). A 1 minute video of the planar target was recorded (1,800 frames), whose depth ranged between 10mm to 110mm from the laparoscope tip. A TPS grid of $10 \times 10 \times 7$ was used to construct the NLM. Fig. 3(d) shows a cross-section of $\|\mathcal{L}\|_2$ at depth 15mm.

We can easily show the NLM is a far better model than the PLM. In Fig. 3(e) we show the predicted surface albedo, according to (1), for the plane in Fig. 3(b) assuming the PLM. The plane's true albedo is a constant $\alpha = 100$, but note that the albedo estimate towards the back is over double at the front, indicating a gross modelling error. By contrast, Fig. 3(f) shows a far better predicted albedo using the NLM. Note that this view was not used for training the model.

3 Parallelised 3D Depth Estimation

We return now to optimising (2), but now using our improved light model. For realtime performance our solution is inspired by TS. We compute a discretised version of (1), solved iteratively with depth estimates being updated fast, locally and in parallel. However, our key extensions are (*i*) to handle perspective cameras, (*ii*) to handle general light models (including our NLM), (*iii*) guaranteeing convergence and (*iv* computes smooth surfaces. We discretise (2) using the pixel grid and augment it with a smoothing prior. This takes the form:

$$\begin{aligned} d^\star(u,v) = \arg\min_{d(u,v)} \sum_\Omega (\alpha(u,v)\,\mathcal{L}(f(u,v)) \cdot n(u,v) - \hat{I}(u,v))^2 + \\ \lambda \sum_\Omega \left(\tfrac{d}{du} f(u,v) + \tfrac{d}{dv} f(u,v)\right)^2 \end{aligned} \tag{4}$$

λ denotes the smoothing weight, which we currently experimentally set. At the $(t+1)^{th}$ iteration the depth at (u,v) is updated by:

$$d^{t+1}(u,v) = \arg\min_{d} \left[\begin{array}{c} \left(\alpha(u,v)\,\mathcal{L}(f(u,v)) \cdot n(u,v;d) - \hat{I}(u,v)\right)^2 + \\ \lambda\left(\frac{d}{du} f(u,v) + \frac{d}{dv} f(u,v)\right) \end{array} \right] \tag{5}$$

where the local surface normal $n(u,v;d)$ is computed by finite differences:

$$n(u,v;d) = \frac{1}{z(d)} \left[\begin{array}{c} \left(d\mathbf{K}^{-1}(u,v,1)^{\top} - d^{t}(u+1,v)\,\mathbf{K}^{-1}(u+1,v,1)^{\top}\right) \times \\ \left(d\mathbf{K}^{-1}(u,v,1)^{\top} - d^{t}(u+1,v)\,\mathbf{K}^{-1}(u+1,v,1)^{\top}\right) \end{array} \right] \tag{6}$$

Here, $z(d)$ denotes the normalisation term such that $\|n(u,v;d)\|_2 = 1$. To optimise efficiently, we sample candidate depths about the current depth estimate in log-space via the rule: $d^{t+1}(u,v) \in \pm[d^{t}(u,v) + (1+\tau)^{a} - 1]$ with $\tau = 0.1$ and the integer a in the range: $a \in [0:10]$.

4 Algorithm and GPU Implementation

In this section we outline the details and some of the main design decisions we have made in implementing live 3D laparoscopy using the methods described in §2 and §3. Importantly, each input frame is processed independently, and thus does not rely on the 3D computed in previous or successive frames.

4.1 Hardware Configuration and Offline Processing

Our current hardware configuration is as follows: A Karl Storz laparoscope with video outputted via firewire at 720×576, a Windows 7 PC with an Intel Core2 Duo 3000MHz processor and NVidia Quadro FX 3800 card (with 192 CUDA cores). For 3D visualisation we use a single Acer 25 Inch LCD 3D display combined with 4 pairs of NVidia 3D active shutter glasses. The laparoscope's automatic exposure is turned off (typically we set it to a 1/500 sec exposure) and before running our reconstruction algorithm its camera is radiometrically and geometrically calibrated. After calibrating the NLM, we pre-compute the function $\mathcal{L}$ and store it as a $720 \times 576 \times 20$ quantised 3D frustum (requiring approximately 47MB of storage.) When processing live laparoscope videos, values from the NLM are then computed by 3D bilinear interpolation.

4.2 Online Processing

Input frames are captured from the laparoscope using the OpenCV libraries. We have found experimentally that the red channel best satisfies the constant albedo assumption for tissues in laparoscopy. We transfer the red channel to the GPU's DRAM. Subsequently all processing is done on the GPU. We report here the average absolute time to complete each process (in ms), and the process's percentage of the runtime budget.

Specular constraints and inpainting. (6.3ms, 14.2%) First, specularities are detected via combined saturation and lightness thresholding (we use thresholds of saturation 0.9 and lightness 0.95 respectively to detect specular pixels.) A rigorous modelling of specularities involves having physical reflectance models

of tissue, which can be hard to apply in general. We opt for a simpler strategy: Our trick with processing specular pixels is to modify their intensity to that which would be predicted by Lambertian reflectance, and then running our SfS method. The process works by constraining each specular region's centroid to have maximal Lambertian reflectance. This is a reasonable estimate at its true Lambertian reflectance, because for a laparoscope the viewing rays and light rays are approximately colinear at these points. We then smoothly propagate this through the specular region.

Removal of high-frequency content. (2.8ms, 6.4%) We make the assumption that high frequency changes in image structure are caused by artifacts such as vascular structure, and not by shading variation. These are removed by running a GPU-optimised 7×7 media filter over the image.

Depth map initialisation and estimation. (34.3ms, 77.3%) We initialise the depth map to be a fronto-parallel plane positioned at a 50mm from the laparoscope's end. We then run our SfS method until convergence is detected or if 30 iterations have passed. At each iteration, the pixel depths are updated in checkerboard fashion and in parallel on the GPU.

3D Visualisation. (0.93ms, 2.1%) Once the depth map has been estimated, we render two views of the 3D surface using stereoscopic OpenGL, and pass these to the FX 3800's quad buffers. Currently a parallel binocular setup is used, with cameras positioned either side of the real laparoscope camera model. Their parameters, including stereo baseline and 3D position are completely controllable.

5 Experimental Evaluation

5.1 Ex-vivo Experiments

To quantify the performance of our approach and to understand the general limitations of SfS for 3D laparoscopic reconstruction, we first present results for reconstructing ex-vivo a piglet's kidney. The analysis is presented in Figure 4. The kidney has been augmented with small white markers to provide the transform between the laparoscope's view and a Ground Truth (GT) surface, captured via a high-resolution structured light system (Fig. 4(l)). Three test laparoscope images are shown in Figs. 4(a,e,i). Our goal here is to understand three key aspects: the well-posedness of the problem, the accuracy of the reconstruction and to measure the sensitivity of our approach to initialisation. We achieve this by manually segmenting the kidney and running the SfS method, initialised by fronto-parallel depth maps positioned at increasing depths d_{init}, ranging from 10mm $< d_{init} <$ 100mm. Now, a well-posed problem is indicated by a unique global minimum of the cost function (4). By contrast, the sensitivity to initialisation is indicated by measuring, for each value of d_{init}, the post-optimisation RMS Error w.r.t true depth. We illustrate the results graphically in Figs. 4(b,f,j). In Fig. 4(b) one can see the same (correct) minimum is reached between 20mm $< d_{init} <$ 60mm mm, indicating reasonable resistance to coarse initialisation. The solutions for $d_{init} >$ 60mm and $d_{init} <$ 20mm have higher reconstruction error, however they also are marked by a higher cost function.

This suggests that for image Fig. 4(a), the problem is well-posed. The reconstructed depth map for $d_{init} = 50$mm is shown in Fig. 4(c), and a 3D render of the visible surface from a novel viewpoint is shown in Fig. 4(d). For images in Figs. 4(e,i) we have similar results. However, the depth maps contain an error. At the bottom right of the kidney, the reconstructed depths tend closer to the camera, suggesting the surface here is concave. This is incorrect. Thus, there exists a convex/concave ambiguity and a non-unique solution. Returning back to the question posed in §1: *Is SfS sufficient to resolve 3D depth from monocular laparoscope images?* The answer is not entirely, and another cue (such as motion) is needed to resolve the ambiguity.

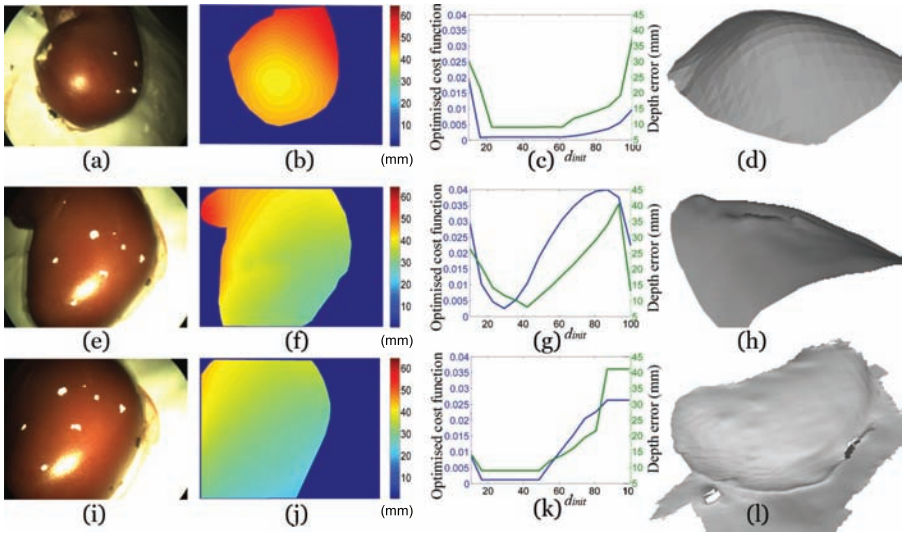

Fig. 4: Ex-vivo experimentation using a pig kidney with ground truth evaluation.

5.2 Successful In-vivo Reconstruction

§5.1 shows reasonable 3D can be obtained via our SfS method, although failures can arise due to pose ambiguities. Here we show an in-vivo video sequence where 3D has been successfully reconstructed. The sequence is of a human uterus comprising 340 frames lasting 14.8 seconds [6]. We are specifically motivated by the application of computer-assisted myoma removal, where intra-operative 3D reconstruction can aid pre-operative data registration. This allows for example, visual augmentation of the fist planned incision path. The results on this dataset are summarised in Fig. 5. In Fig. 5(a-d) we show four representative frames from the sequence. In Figs. 5(e-g) we show the automatic image preprocessing done before running SfS. Fig. 5(e) is a raw input frame, Fig. 5(f) is the red channel after high-frequency content removal with detected specularities, Fig. 5(g) shows the image after specular inpainting. The uterus remains approximately rigid in the first half of the sequence. Our key idea is to evaluate our SfS method by constructing *in-vivo* quasi-GT 3D data using rigid SfM on these frames. To do

this, a ROI was manually marked around the uterus (Fig. 5(h)) and a dense GT surface bound by the ROI was computed using standard manually-assisted multiview SfM methods. We show this in Fig. 5(i). We processed the video sequence using two methods: (1) Tsai and Shah (TS) and (2) our proposed method. Note that in TS absolute depth cannot be computed. To enable quantitative comparison, we compute the mean error in surface normals (in angular degrees). These scores are computed at each pixel bound within the visible part of the uterus.

Fig. 5(j) shows the reconstruction by our method for Fig. 5(h) and Fig. 5(k) shows it reconstructed by TS. Qualitatively, we ours appears superior to TS. In Fig. 5(l) we compare the reconstruction error of TS to ours (labelled as X-NLM). Here TS is noticeably outperformed by our method. The instrument occluded the uterus between frames 117-125, which explains the error spike. However, notices that since frames are processed independently, this has no adverse effect on successive frames. In Fig. 5(m,n) we show the full-image reconstruction using our method and TS respectively for the frame in Fig. 5(d). We note that no segmentation of the tool was performed, and the tissues surrounding the uterus violate the surface continuity and constant albedo assumptions. In spite of this, the uterus is recovered faithfully using our method. In Figs. 5(o,p,q) we show synthesised novel views, texture-mapped using the input image frames. In 5(q) the tool occludes the uterus, explaining the tool-shaped valley in the reconstruction. In Fig. 5(r) we show a basic application; we take 3D SfS surface and visually augment it with the pre-operative planned incision path (shown in green).

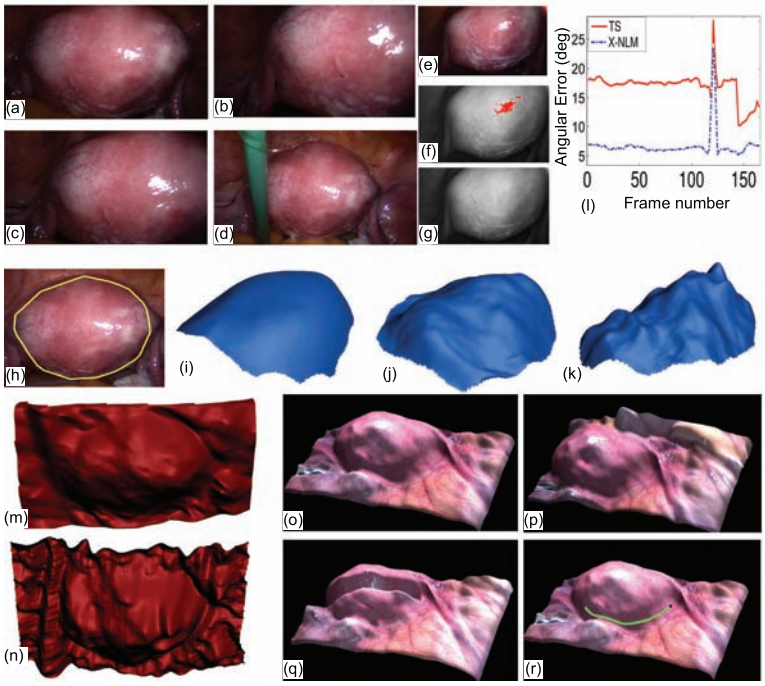

Fig. 5: Successful in-vivo 3D reconstruction of the human uterus.

6 Discussion and Conclusion

In this paper we have pushed forward SfS for live 3D monocular laparoscopy in various modelling and computational aspects. The contributions and outstanding challenges are stated clearly in Figure 2. The biggest limitations not addressed by our work are two-fold: (*i*) The surface albedo must be constant and known *a priori* and (*ii*) solution ambiguities due to the ill-conditioning of SfS. We believe these are tremendously difficult to resolve using the shading cue alone. Our direction for future research will be to take our live reconstruction framework and complement it with other 3D cues. For example using sparse realtime 3D estimates at tracked features or stereo laparoscopic images.

References

1. Ackerman, J. D., Keller, K., Fuchs, H. Surface reconstruction of abdominal organs using laparoscopic structured light for augmented reality. 3DICA (2002) 39-46
2. Barron, J. T., Malik, J.: High-frequency shape and albedo from shading. CVPR (2011) 2521-2528
3. M. Hu, G. Penney, P. Edwards, M. Figl, and D. Hawkes. 3D reconstruction of internal organ surfaces for minimal invasive surgery. MICCAI (2007) 68-77
4. Lau, W. W., Ramey, N. A., Corso, J. J., Thakor, N. V., Hager, G. D.: Stereo-Based Endoscopic Tracking of Cardiac Surface Deformation. MICCAI (2004) 494-501
5. Lepetit, V., Moreno-Noguer, F., Fua, P. EPnP: An Accurate O(n) Solution to the PnP Problem. IJCV 81 (2008) 155-166
6. Malti, A., Bartoli, A., Collins, T. Template-Based Conformal Shape-from-Motion from Registered Laparoscopic Images. MIUA (2011)
7. Mountney, P. and Yang, G.-Z. Motion compensated SLAM for image guided surgery. MICCAI (2010) 496-504
8. Mueller-Richter, U. D. A., Limberger, A., Weber, P., Ruprecht K. W., Spitzer, W., Schilling, M.: Possibilities and limitations of current stereo-endoscopy. Surgical Endoscopy 18 (2004) 942-947
9. Okatani, T., Deguchi, K.: Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. CVIU 66 (1997) 119-131
10. Prados, E., Faugeras, O.: Perspective Shape from Shading and Viscosity Solutions. ICCV (2003) 826-831
11. Prados, E., Faugeras, O.: Shape from Shading: a well-posed problem? CVPR (2005) 870-877
12. Penne, J., Höller, K., Stürmer, M., Schrauder, T., Schneider, A., Engelbrecht, R., Feußner, H., Schmauss, B.: Time-of-Flight 3-D Endoscopy. MICCAI (2009) 467-474
13. Quartucci, C. H., Tozzi, C. L.,: Towards 3D Reconstruction of Endoscope Images using Shape from Shading. SIBGRAPI (2000) 90-96
14. Stoyanov, D., Darzi, A., Yang, G.Z.: Dense 3D Depth Recovery for Soft Tissue Deformation During Robotic Assisted Laparoscopic Surgery. MICCAI (2004) 41-48
15. Tankus, A., Sochen, N., Yeshurun, Y.: Perspective SfS by fast marching. CVPR (2004) 43-49
16. Tankus, A., Sochen N., Yeshurun, Y.: Shape-from-Shading Under Perspective Projection. IJCV 63 (2007) 21-43
17. Tsai, P., Shah, M.: Shape From Shading Using Linear Approximation. Image and Vision Computing 12 (1994) 487-498
18. Wu, C., Narasimhan S.G., Jaramaz, B.: Shape-from-Shading under Near Point Lighting and Partial views for Orthopedic Endoscopy. PACV (2007)
19. Yeung, S. Y., Tsui, H. T., Yim, A.: Global SfS for an Endoscope Image. MICCAI (1999) 318-327