

# Generalized Thin-Plate Spline Warps

Adrien Bartoli<sup>1</sup>

Mathieu Perriollat<sup>2</sup>

Sylvie Chambon<sup>3</sup>

<sup>1</sup> Université d'Auvergne  
Clermont-Ferrand, France

<sup>2</sup> VI-Technology  
Grenoble, France

<sup>3</sup> LCPC  
Nantes, France

Corresponding author: Adrien Bartoli

`Adrien.Bartoli@gmail.com`

## Abstract

The Thin-Plate Spline warp has been shown to be a very effective parameterized model of the optic flow field between images of various types of deformable surfaces, such as a paper sheet being bent. Recent work has also used such warps for images of a smooth and rigid surface. Standard Thin-Plate Spline warps are however not rigid, in the sense that they do not comply with the epipolar geometry. They are also intrinsically affine, in the sense of the affine camera model, since they are not able to simply model the effect of perspective projection.

We propose three types of warps based on the Thin-Plate Spline. The first one is a rigid flexible warp. It describes the optic flow field induced by a smooth and rigid surface, and satisfies the affine epipolar geometry constraint. The second and third proposed warps extend the standard Thin-Plate Spline warp and the proposed rigid flexible warp to the perspective camera model. The properties of these warps are studied in details and a hierarchy is defined. Experimental results on simulated and real data are reported.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Previous Work</b>	<b>3</b>
<b>3</b>	<b>Notation and Preliminaries</b>	<b>5</b>
3.1	Notation . . . . .	5
3.2	TPS – The Thin-Plate Spline . . . . .	6
3.3	Rigid Surfaces . . . . .	9
<b>4</b>	<b>Warps with the Affine Camera Model</b>	<b>11</b>
4.1	DA-Warp – The Standard TPS-Warp . . . . .	11
4.1.1	Derivation . . . . .	11
4.1.2	Properties . . . . .	12
4.1.3	Projected Deformable Surface Interpretation . . . . .	12
4.1.4	Example . . . . .	14
4.2	RA-Warp – The Rigid Affine TPS Warp . . . . .	15
4.2.1	Derivation . . . . .	15
4.2.2	Properties . . . . .	17
4.2.3	Projected Rigid Surface Interpretation . . . . .	17
4.2.4	Example . . . . .	18
4.3	Behaviour with the Perspective Camera Model . . . . .	19
<b>5</b>	<b>Warps with the Perspective Camera Model</b>	<b>20</b>
5.1	RP-Warp – The Rigid Perspective TPS Warp . . . . .	21
5.1.1	Derivation . . . . .	21
5.1.2	Properties . . . . .	21
5.1.3	Example . . . . .	22
5.2	DP-Warp – The Deformable Perspective TPS Warp . . . . .	22
5.2.1	Derivation . . . . .	23
5.2.2	Properties . . . . .	24
5.2.3	Example . . . . .	24
<b>6</b>	<b>A Hierarchy of Warps</b>	<b>25</b>

<b>7</b>	<b>Estimation of the Warps</b>	<b>26</b>
7.1	General Points . . . . .	26
7.2	The DA-Warp . . . . .	27
7.3	The RA-Warp . . . . .	28
7.4	The RP-Warp . . . . .	29
7.5	The DP-Warp . . . . .	30
<b>8</b>	<b>Experimental Evaluation</b>	<b>30</b>
8.1	Simulated Data . . . . .	30
8.2	Real Data . . . . .	33
<b>9</b>	<b>Discussion</b>	<b>37</b>
<b>A</b>	<b>Asymptotic Internal Regularization Behaviour</b>	<b>40</b>
A.1	The DA-Warp . . . . .	40
A.2	The RA-Warp and the RP-Warp . . . . .	41
A.3	The DP-Warp . . . . .	42
<b>B</b>	<b>Relationship Between the Depth and the Transfer Errors</b>	<b>43</b>
B.1	Affine Imaging Model . . . . .	43
B.2	Perspective Imaging Model . . . . .	43
<b>C</b>	<b>Some Relationships Between the Sets of Warps</b>	<b>44</b>
C.1	$\mathbb{S}_{FA} \subset \mathbb{S}_{RA}$ and $\mathbb{S}_{FP} \subset \mathbb{S}_{RP}$ . . . . .	44
C.2	$(\mathbb{S}_{FP} \setminus \mathbb{S}_{FA}) \cap \mathbb{S}_{DA} = \emptyset$ . . . . .	45
C.3	$\mathbb{S}_{RA} = \mathbb{S}_{DA} \cap \mathbb{S}_{RP}$ . . . . .	45
<b>D</b>	<b>The 3D Coordinate Frame for Rigid Warps</b>	<b>46</b>

# 1 Introduction

Given two images of some scene surface, there exists an  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  function, called a *warp*, mapping a point from the *source image* to the corresponding point in the *target image*. For instance, two images of a rigid planar surface taken by a perspective camera are related by an homographic warp. For a non-planar, 3D scene, the warp is more complex since it depends on the surface depth. When the observed surface is deformable, the warp is even more complex. Examples of rigid scene models include the cases where the scene structure is piecewise planar as in (Yang et al., 2005) or nearly planar as in (Irani et al., 1999). Examples of deformable scene models include the flexible low-rank model as in (Bregler et al., 2000), the learned low-rank model in (Salzmann et al., 2007), the face model in (Banz and Vetter, 2003) and the paper model in (Perriollat and Bartoli, 2006).

Representing the warp using a parametric function requires prior assumptions about the observed scene structure. One common, fairly generic assumption is that the inter-image optic flow field is smooth.<sup>1</sup> This naturally leads to using the Thin-Plate Spline (TPS) as a building block for the warps. The TPS is a smooth, compact and convenient  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  function that minimizes the *bending energy* of the 2.5D surface it defines, while satisfying a set of constraints induced by data points. The standard TPS warp is very flexible in that it is controlled by centres that may be placed anywhere in the images. It is known to be an efficient approximation to many types of deformations (Bookstein, 1989). The standard TPS warp has recently been used as a simple parametric warp for images of a rigid and smooth 3D surface in (Wills and Belongie, 2004) and (Masson et al., 2005), for respectively wide-baseline image matching and object tracking.

There are, however, two main issues with the use of the standard TPS warp in this context, that have not been dealt with in the literature:

- **The standard TPS warp overfits affine images of rigid surfaces.** The standard TPS warp does not in general satisfy the rigidity constraint modeled by the affine epipolar geometry. In that sense it is ‘too flexible’ in affine imaging conditions and for rigid surfaces.
- **The standard TPS warp does not model perspective.** The standard TPS warp is intrinsically affine, in the sense of the affine camera model, since its formulation does not include a division. For instance, as mentioned in (Wills and Belongie, 2004), it is not able to ‘reproduce’ a simple homographic warp with a finite number of centres.<sup>2</sup>

<sup>1</sup>The inter-image optic flow field is smooth if the observed surface is itself smooth, opaque (so that points at the back of the surface are hidden from the cameras) and does not self-occlude in the two observed images. The effects of self-occlusions on the optic flow field were studied in (Gay-Bellile et al., 2009).

<sup>2</sup>Even though a large number of centres may be used to approximate perspective transformations arbitrarily well.

Henceforth, we call *DA-Warp* the standard TPS warp (for ‘Deformable Affine’). This paper brings contributions that address the two above mentioned issues, through the proposal of three new warps illustrated in figure 1. First, a specialization of the DA-Warp to rigid surfaces is introduced. This warp is called *RA-Warp* (for ‘Rigid Affine’) and is very similar to the DA-Warp with an epipolar constraint on the warp coefficients. This solves the first issue. Second, the extension of the RA-Warp to the perspective camera model is proposed. This warp, dubbed *RP-Warp* (for ‘Rigid Perspective’), naturally includes the *FP-Warp* (‘Flat Perspective’)<sup>3</sup> similarly to the RA-Warp including the *FA-Warp*. This solves the second issue for the case of rigid surfaces. Third, we introduce the *DP-Warp* (for ‘Deformable Perspective’) which is shown to be the perspective analogue of the DA-Warp. This solves the second issue for the case of deformable surfaces: perspective transformations can be approximated arbitrarily well by our DP-Warp with much fewer centres than for the DA-Warp. The derivation of these warps is made possible by a *feature-driven parameterization* of the TPS that we propose. The hierarchy and dependencies between these four types of warps – the DA-Warp and the three new types of warps – as well as the FA-Warp and the FP-Warp are studied in details. In order to derive warps independent of the intrinsic camera parameters, we consider uncalibrated cameras.

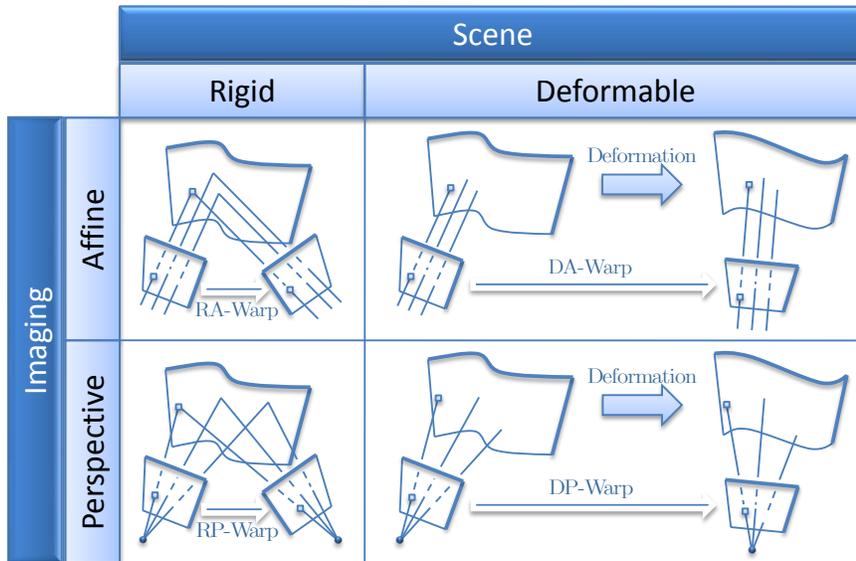


Figure 1: Summary showing the standard DA-Warp and the three new types of warps we propose (the RA-Warp, the RP-Warp and the DP-Warp) that deal with affine and perspective projections of rigid and deformable surfaces.

One of the reasons for the popularity of the TPS is that the 2.5D surface it defines interpolates a set of *centres* with minimal bending energy. In the DA-Warp, a TPS is used for each of the two coordinates.

<sup>3</sup>The FP-Warp is a 2D homographic warp with 8 parameters. Similarly, the FA-Warp (‘Flat Affine’) is a 2D affine warp with 6 parameters.

These two TPS functions share their centres. Their bending energy thus becomes empirical as it is expressed within the target image. In contrast, the bending energy of the proposed warps is related to a 2.5D or a 3D surface.

The paper is organized as follows. Previous work is reviewed in §2. We describe our notation in §3 followed by our feature-driven parameterization of the TPS. Affine and perspective projection warps are respectively presented in §§4 and 5. The hierarchy and relationships between the warps are studied in §6. Methods for estimating the warps are given in §7. Experimental results are reported in §8 and a discussion is provided in §9. The Appendices bring proofs for some of our results.

A set of images along with point correspondences is used to illustrate the algorithms throughout the paper. These are images of a poster. They were taken under various imaging conditions (affine and perspective) and with various deformations. We manually picked 206 point correspondences. We also generated a *warp visualization grid* by rectifying the images for which the poster is flat. The point correspondences and the warp visualization grid are shown in figure 2 for one image and the complete set of images is shown in figure 3. All the algorithms we use in this paper estimate the warps by minimizing a *transfer error*, defined as the discrepancy between the points transferred by the warp from the source to the target image, and the points measured in the target image. We report the RMSR (Root Mean Square Residual) in pixels, which is proportional to the transfer error and represents the extent to which the model fits the data. All the comparisons are done with  $l \in \{3, 9, 30\}$  deformation centres for the warps. The reason is as follows. First,  $l = 3$  is the minimum number of deformation centres for all warps. In that case they take very simple forms (for instance, the DA-Warp becomes an FA-Warp). Second,  $l = 9$  is a quite small number of centres (think of it as a  $(3 \times 3)$  grid for instance), that allows for some flexibility in the warps. Third,  $l = 30$  is the other extreme: it makes the warps very flexible.

## 2 Previous Work

Deformable image warps are used in different contexts, such as augmented reality (Pilet et al., 2008) and medical image registration (Bookstein, 1989). The DA-Warp *i.e.*, the standard TPS warp, is amongst the most popular parameterized warps, even though there is a great body of work on defining other warps, such as the FFD (Free-Form Deformations) (Sederberg and Parry, 1986), the MFFD (Multilevel FFD) (Lee et al., 1996), more recently the MLS (Moving Least Squares) (Schaefer et al., 2006) or the diffeomorphic warps in (Cootes et al., 2004).

The DA-Warp was introduced in the seminal paper (Bookstein, 1989), based on the TPS. The TPS was derived in (Duchon, 1976) as the function that interpolates or approximates data points while minimizing



Figure 2: (a) shows image FA1 of the poster overlaid with the 206 interest points that we manually entered on all the images of figure 3. (b) shows the warp visualization grid created with an homographic warping of a regular grid. This grid is later used for illustrating the computed warps.

Maximum zoom ( $\approx$ ‘Affine images’)			Minimum zoom (‘Perspective images’)		
Flat surface	Deformation 1	Deformation 2	Flat surface	Deformation 1	Deformation 2
					
FA1	DA11	DA21	FP1	DP11	DP21
					
FA2	DA12	DA22	FP2	DP12	DP22

Figure 3: Images of a poster that were taken for different imaging conditions and deformations, with their ‘name’ (F for Flat, D for Deformation, A for Affine and P for Perspective). The ‘affine images’ were obtained by strongly zooming on the poster, while the ‘perspective images’ were obtained by standing close to the poster. The effect of perspective is typically seen by foreshortening (see images FP1 and FP2, for instance). We selected source and target images from the images shown in this figure to illustrate the different warps throughout the paper. Every possible image pair formed by two of these images satisfies the assumption of smoothness of the image optic flow field we mentioned in the introduction: the surface is smooth, opaque and does not self-occlude.

the integral bending energy. It is worth of note that several papers use the integral bending energy for other purposes. For instance, it is used as a term in a compound cost function for 3D surface reconstruction in (Terzopoulos, 1983), and more recently, for 3D reconstruction of curved surfaces from a single view in (Prasad et al., 2006).

A number of papers focus on defining new estimation methods for the DA-Warp. (Bookstein, 1989) initially proposed a method relying on choosing point landmark correspondences as centres for the DA-Warp. (Lim and Yang, 2005) and (Bartoli and Zisserman, 2004) recently proposed two different strategies to estimate the DA-Warp with a direct method *i.e.*, intensity-based. (Chui and Rangarajan, 2003) simultaneously solve for the DA-Warp and match points using the principle of softassign, which uses the discrepancy between a source point transferred by the warp and a target point as a probability of matching. (Donato and Belongie, 2002) propose algorithms for the fast computation of the DA-Warp from point correspondences.

In contrast to existing work, we build on the TPS to derive new types of warps by taking into account the possible rigidity of the observed surface and the perspective camera model. Methods for estimating those warps from point correspondences are given.

### 3 Notation and Preliminaries

We give our notation, present the TPS and its feature-based parameterization, and explain how rigidity is modeled using the fundamental matrix.

#### 3.1 Notation

Scalars are in italics (*e.g.*,  $x$ ), vectors in bold right fonts (*e.g.*,  $\mathbf{q}$ ) and matrices in sans-serif and calligraphic fonts (*e.g.*,  $\mathbf{P}$  and  $\mathcal{E}$ ). The elements of a vector are written as in  $\mathbf{a}^\top = (a_1 \ a_2 \ a_3)$  where  $^\top$  is vector and matrix transpose. We do not make a difference between coordinate vectors and physical entities. The coordinates of a point in the source image are written with a 2-vector  $\mathbf{q}^\top = (x \ y)$ .  $\mathbb{R}^r$  and  $\mathbb{P}^r$  designate respectively the Euclidean and projective spaces of dimension  $r$ . We write  $d^2(\mathbf{q}, \mathbf{q}') = \|\mathbf{q} - \mathbf{q}'\|^2$  the Euclidean distance between two points  $\mathbf{q}$  and  $\mathbf{q}'$  with  $\|\cdot\|^2$  the vector two-norm and matrix Frobenius norm. Homogeneous coordinates are written as in  $\check{\mathbf{q}}^\top \sim (\check{q}_1 \ \check{q}_2 \ \check{q}_3)$ , where  $\sim$  means equality up to scale. Scaled homogeneous coordinates are written as in  $\tilde{\mathbf{q}}^\top = (\mathbf{q}^\top \ 1)$ . The homogeneous to affine coordinates function  $\psi$  is defined by  $\mathbf{q} = \psi(\check{\mathbf{q}})$ :

$$\psi(\check{\mathbf{q}}) \stackrel{\text{def}}{=} \frac{1}{\check{q}_3} \begin{pmatrix} \check{q}_1 \\ \check{q}_2 \end{pmatrix}.$$

The skew-symmetric  $(3 \times 3)$  cross-product matrix  $[\check{\mathbf{q}}]_{\times}$  is defined such that  $[\check{\mathbf{q}}]_{\times} \check{\mathbf{q}}' = \check{\mathbf{q}} \times \check{\mathbf{q}}'$ . Full column rank portrait matrix pseudo-inverse is defined by  $\mathbf{X}^{\dagger} = (\mathbf{X}^{\top} \mathbf{X})^{-1} \mathbf{X}^{\top}$ . Notation  $\mathbf{X}^{\dagger \top}$  designates the transpose of the pseudo-inverse of matrix  $\mathbf{X}$ .

We consider  $l$  centres with coordinates  $\mathbf{c}_k$  in the source image, with  $k = 1, \dots, l$ . They are gathered in an  $(l \times 2)$  matrix  $\mathbf{P}$  containing their  $x$  and  $y$  coordinates on its columns, and an  $(l \times 3)$  matrix  $\tilde{\mathbf{P}}$  with a third column of ones (*i.e.*,  $\tilde{\mathbf{P}} = (\mathbf{P} \ \mathbf{1})$ , with  $\mathbf{1}$  a vector of ones with the appropriate size). Matrix  $\check{\mathbf{P}}$  equals matrix  $\tilde{\mathbf{P}}$  with each row rescaled by some scalar factor *i.e.*,  $\check{\mathbf{P}} = \text{diag}(\mathbf{d}) \tilde{\mathbf{P}}$ . The centres in the target image are written  $\mathbf{c}'_k$ . Matrices  $\mathbf{P}'$ ,  $\tilde{\mathbf{P}}'$  and  $\check{\mathbf{P}}'$  are defined in the target image similarly as for the source image.

Warps in affine coordinates are written  $\mathcal{W}$ , while  $\check{\mathcal{W}}$  and  $\check{\check{\mathcal{W}}}$  are used for scaled homogeneous and homogeneous coordinates respectively. The sets of FA-Warps and FP-Warps (*i.e.*, affine and homographic warps respectively) are denoted  $\mathbb{S}_{\text{FA}}$  and  $\mathbb{S}_{\text{FP}}$ . For a  $(2 \times 3)$  FA-Warp matrix  $\mathbf{B}$  and a  $(3 \times 3)$  FP-Warp matrix  $\check{\mathbf{H}}$ , we have:

$$\mathcal{W}_{\text{FA}}(\mathbf{q}; \mathbf{B}) \stackrel{\text{def}}{=} \mathbf{B}\check{\mathbf{q}} \quad \text{and} \quad \check{\mathcal{W}}_{\text{FP}}(\mathbf{q}; \check{\mathbf{H}}) \stackrel{\text{def}}{\approx} \check{\mathbf{H}}\check{\check{\mathbf{q}}}.$$

### 3.2 TPS – The Thin-Plate Spline

The TPS is an  $\mathbb{R}^2 \rightarrow \mathbb{R}$  function driven by assigning target values  $\alpha_k$  to the 2D centres  $\mathbf{c}_k$  and enforcing several conditions: the TPS is the Radial Basis Function that minimizes the integral bending energy. The idea of using the Thin-Plate equation as an interpolation map is due to (Duchon, 1976). The standard  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  TPS-Warp, or DA-Warp with our naming conventions, is obtained by ‘stacking’ two TPS functions sharing their centres, as described in the next section.

**Standard parameterization.** Given the  $l$  centres, the TPS is usually parameterized by an  $l+3$  coefficient vector  $\mathbf{h}_{\alpha, \lambda}^{\top} = (\mathbf{w}^{\top} \ \mathbf{a}^{\top})$  and an *internal regularization weight*  $\lambda \in \mathbb{R}^+$ . These coefficients can be computed from the target vector  $\alpha$  as described in the next paragraph. There are  $l$  coefficients in  $\mathbf{w}$  and three coefficients in  $\mathbf{a}$ . The TPS is given by:

$$\omega(\mathbf{q}; \mathbf{h}_{\alpha, \lambda}) \stackrel{\text{def}}{=} \left( \sum_{k=1}^l w_k \rho(d^2(\mathbf{q}, \mathbf{c}_k)) \right) + \mathbf{a}^{\top} \check{\mathbf{q}}, \quad (1)$$

where  $\rho(d) \stackrel{\text{def}}{=} d \log(d)$  is the *TPS kernel function* for the squared distance. The coefficients in  $\mathbf{w}$  must satisfy  $\tilde{\mathbf{P}}^{\top} \mathbf{w} = \mathbf{0}$ . These three ‘side-conditions’ ensure that the TPS has square integrable second derivatives. It is convenient to define the  $(l+3)$ -vector  $\ell_{\mathbf{q}}$  as:

$$\ell_{\mathbf{q}}^{\top} \stackrel{\text{def}}{=} \left( \rho(d^2(\mathbf{q}, \mathbf{c}_1)) \ \cdots \ \rho(d^2(\mathbf{q}, \mathbf{c}_l)) \ \check{\mathbf{q}}^{\top} \right),$$

allowing the TPS (1) to be rewritten as a dot product:

$$\omega(\mathbf{q}; \mathbf{h}_{\alpha, \lambda}) = \ell_{\mathbf{q}}^{\top} \mathbf{h}_{\alpha, \lambda}. \quad (2)$$

**Standard estimation.** Applying the TPS (1) to the centre  $\mathbf{c}_r$  with target value  $\alpha_r$  gives:

$$\left( \sum_{k=1}^l w_k \rho(d^2(\mathbf{c}_r, \mathbf{c}_k)) \right) + \mathbf{a}^{\top} \tilde{\mathbf{c}}_r = \alpha_r.$$

Combining the equations obtained for all the  $l$  centres with the side-conditions  $\tilde{\mathbf{P}}^{\top} \mathbf{w} = \mathbf{0}$  in a single matrix equation gives:

$$\underbrace{\begin{pmatrix} \mathbf{K}_{\lambda} & \tilde{\mathbf{P}} \\ \tilde{\mathbf{P}}^{\top} & \mathbf{0} \end{pmatrix}}_{\mathcal{D}} \underbrace{\begin{pmatrix} \mathbf{w} \\ \mathbf{a} \end{pmatrix}}_{\mathbf{h}_{\alpha, \lambda}} = \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{0} \end{pmatrix} \quad \text{with} \quad K_{r,k} = \begin{cases} \lambda & r = k \\ \rho(d^2(\mathbf{c}_r, \mathbf{c}_k)) & \text{otherwise} \end{cases}$$

As can be seen,  $\lambda \mathbf{I}$  is added to the leading block  $\mathbf{K}_{\lambda}$  of the design matrix  $\mathcal{D}$ , and acts as an internal regularization. Solving for  $\mathbf{h}_{\alpha, \lambda}$  by inverting  $\mathcal{D}$  is the classical linear method for estimating the TPS coefficients due to (Bookstein, 1989). The coefficient vector  $\mathbf{h}_{\alpha, \lambda}$  is thus a nonlinear function of the internal regularization weight  $\lambda$  and a linear function of the target vector  $\boldsymbol{\alpha}$ . In practice, we set  $\lambda$  to some small value such as  $\lambda = 10^{-4}$ , to ensure that  $\mathbf{K}_{\lambda}$  and thus  $\mathcal{D}$  are full rank matrices.

**Feature-driven parameterization.** We express  $\mathbf{h}_{\alpha, \lambda}$  as a linear ‘back-projection’ of the target vector  $\boldsymbol{\alpha}$ . This is modeled by the matrix  $\mathcal{E}_{\lambda}$ , nonlinearly depending on  $\lambda$ , given by the  $l$  leading columns of  $\mathcal{D}^{-1}$ :

$$\mathbf{h}_{\alpha, \lambda} = \mathcal{E}_{\lambda} \boldsymbol{\alpha} \quad \text{with} \quad \mathcal{E}_{\lambda} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{K}_{\lambda}^{-1} \left( \mathbf{I} - \tilde{\mathbf{P}} \left( \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \tilde{\mathbf{P}} \right)^{-1} \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \right) \\ \left( \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \tilde{\mathbf{P}} \right)^{-1} \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \end{pmatrix}. \quad (3)$$

This parameterization of  $\mathbf{h}_{\alpha, \lambda}$  has the advantages to separate  $\lambda$  and  $\boldsymbol{\alpha}$  and to introduce units.<sup>4</sup> The side-conditions are naturally enforced by this parameterization: it can easily be verified that  $\tilde{\mathbf{P}}^{\top} \mathbf{w} = \left( \tilde{\mathbf{P}}^{\top} \mathbf{0} \right) \mathbf{h}_{\alpha, \lambda} = \left( \tilde{\mathbf{P}}^{\top} \mathbf{0} \right) \mathcal{E}_{\lambda} \boldsymbol{\alpha} = \mathbf{0}$ ,  $\forall \lambda \in \mathbb{R}$  and  $\forall \boldsymbol{\alpha} \in \mathbb{R}^l$ .

Incorporating the parameterization (3) into the TPS (2) we obtain what we call the *feature-driven*

<sup>4</sup>While  $\mathbf{h}_{\alpha, \lambda}$  has no obvious unit,  $\boldsymbol{\alpha}$  in general has (*e.g.*, pixels, meters).

parameterization  $\tau(\mathbf{q}; \boldsymbol{\alpha}, \lambda) = \omega(\mathbf{q}; \mathbf{h}_{\boldsymbol{\alpha}, \lambda})$  for the TPS:

$$\boxed{\tau(\mathbf{q}; \boldsymbol{\alpha}, \lambda) \stackrel{\text{def}}{=} \boldsymbol{\ell}_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \boldsymbol{\alpha}.} \quad (4)$$

**Properties.** Equation (4) is a feature-driven parameterization since  $\boldsymbol{\alpha}$  contains the target values for the centres. In practice, these can be the coordinates of image or 3D points. The following important properties are used later in the paper:

$$\tilde{\mathbf{q}}^{\top} = \boldsymbol{\ell}_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \tilde{\mathbf{P}} \quad \forall \mathbf{q} \in \mathbb{R}^2 \quad (5)$$

$$\tilde{\mathbf{q}}^{\top} \boldsymbol{\theta} = \boldsymbol{\ell}_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \tilde{\mathbf{P}} \boldsymbol{\theta} \quad \forall \mathbf{q} \in \mathbb{R}^2 \quad \forall \boldsymbol{\theta} \in \mathbb{R}^3. \quad (6)$$

This shows that an  $\mathbb{R}^2 \rightarrow \mathbb{R}$  affine transformation with parameters  $\boldsymbol{\theta}$  can be reformulated as a TPS with arbitrary centres. These properties stem from the fact that<sup>5</sup>  $\mathcal{E}_{\lambda} \tilde{\mathbf{P}} = (0 \ \mathbf{I})^{\top}$ , giving:

$$\tilde{\mathbf{q}}^{\top} \boldsymbol{\theta} = \boldsymbol{\ell}_{\mathbf{q}}^{\top} \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\theta} \end{pmatrix} = \boldsymbol{\ell}_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \tilde{\mathbf{P}} \boldsymbol{\theta}.$$

It is straightforward to see that these properties imply:

$$\boldsymbol{\ell}_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \mathbf{1} = 1. \quad (7)$$

The integral bending energy  $\kappa = \int_{\mathbb{R}^2} \|\nabla^2 \tau(\mathbf{q}; \boldsymbol{\alpha}, \lambda)\|^2 d\mathbf{q} = \mathbf{w}^{\top} \mathbf{K}_{\lambda} \mathbf{w}$  is also given by  $\kappa = \boldsymbol{\alpha}^{\top} \bar{\mathcal{E}}_{\lambda} \boldsymbol{\alpha}$ , where  $\bar{\mathcal{E}}_{\lambda}$  is the  $(l \times l)$  *bending energy matrix* given by amputating  $\mathcal{E}_{\lambda}$  of its last three rows:

$$\bar{\mathcal{E}}_{\lambda} \stackrel{\text{def}}{=} \mathbf{K}_{\lambda}^{-1} \left( \mathbf{I} - \tilde{\mathbf{P}} \left( \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \tilde{\mathbf{P}} \right)^{-1} \tilde{\mathbf{P}}^{\top} \mathbf{K}_{\lambda}^{-1} \right). \quad (8)$$

The bending energy matrix is symmetric and has rank  $l - 3$ . The eigenvectors corresponding to the  $l - 3$  nonzero eigenvalues are the *principal warps*, the corresponding eigenvalues indicating their bending energy (Bookstein, 1989).

The asymptotic internal regularization behaviour of the TPS is:

$$\lim_{\lambda \rightarrow +\infty} \tau(\mathbf{q}; \boldsymbol{\alpha}, \lambda) = \boldsymbol{\zeta}^{\top} \tilde{\mathbf{q}} \quad \text{with} \quad \boldsymbol{\zeta} \stackrel{\text{def}}{=} \tilde{\mathbf{P}}^{\dagger} \boldsymbol{\alpha}.$$

In other words, the TPS tends to an affine transformation with a 3-vector of coefficients given by  $\tilde{\mathbf{P}}^{\dagger} \boldsymbol{\alpha}$ .

<sup>5</sup>We have  $\mathcal{D}(0 \ \mathbf{I})^{\top} = (\tilde{\mathbf{P}}^{\top} \ \mathbf{0})^{\top}$ , thus  $\mathcal{D}^{-1}(\tilde{\mathbf{P}}^{\top} \ \mathbf{0})^{\top} = (0 \ \mathbf{I})^{\top}$  and  $\mathcal{E}_{\lambda} \tilde{\mathbf{P}} = (0 \ \mathbf{I})^{\top}$ .

Due to the 3 side-conditions, 3 is the minimal number of centres to define a TPS. This is also the minimal case for the warps derived in this paper, since they all use the TPS as a building block. In this minimal case, the TPS is an affine transformation with parameters  $\boldsymbol{\theta} = \tilde{\mathbf{P}}^{-1}\boldsymbol{\alpha}$ . Based on the fact that  $\tilde{\mathbf{P}}$  is of size  $(3 \times 3)$  in this particular case, we substitute  $\boldsymbol{\alpha} = \tilde{\mathbf{P}}\boldsymbol{\theta}$  for some  $\boldsymbol{\theta} \in \mathbb{R}^3$  in the TPS (4). Thanks to property (6), we get  $\tau(\mathbf{q}; \boldsymbol{\alpha}; \lambda) = \tilde{\mathbf{q}}^\top \boldsymbol{\theta}$ , giving:

$$\tau(\mathbf{q}; \boldsymbol{\alpha}; \lambda) = \tilde{\mathbf{q}}^\top \tilde{\mathbf{P}}^{-1} \boldsymbol{\alpha}. \quad (9)$$

We observe that this form of TPS is a particular case of the asymptotically regularized TPS. For all the warps studied in this paper, using the minimal number of centres leads to a special case of the asymptotically regularized warp.

### 3.3 Rigid Surfaces

We first describe how rigidity is modeled by the fundamental matrix. We then show how to extract a pair of projection matrices from the fundamental matrix.

**Rigidity.** The rigidity of the observed scene is modeled by the fundamental matrix that we write  $\mathcal{F}$  or  $\mathcal{A}$  for the perspective and affine camera models respectively. A pair of corresponding points  $\mathbf{q} \leftrightarrow \mathbf{q}'$  satisfies the rigidity constraint *i.e.*, is the projection of the same 3D point, if and only if it satisfies the fundamental equation  $\tilde{\mathbf{q}}'^\top \mathcal{F} \tilde{\mathbf{q}} = 0$ , holding in both the perspective and affine cases. Thus, a warp  $\mathcal{W}$  is rigid if and only if:

$$\check{\mathcal{W}}(\mathbf{q})^\top \mathcal{F} \tilde{\mathbf{q}} = 0 \quad \forall \mathbf{q} \in \mathbb{R}^2. \quad (10)$$

Parameterizing the affine fundamental matrix as:

$$\mathcal{A} \stackrel{\text{def}}{\sim} \begin{pmatrix} 0 & 0 & \mathbf{z} \\ 0 & 0 & \\ & \mathbf{J}^\top & \end{pmatrix} \quad \text{with} \quad \mathbf{J}^\top \stackrel{\text{def}}{=} (c \ d \ e) \quad \text{and} \quad \mathbf{z}^\top \stackrel{\text{def}}{=} (a \ b), \quad (11)$$

we rewrite the definition (10) of a rigid affine warp as:

$$\mathcal{W}(\mathbf{q})^\top \mathbf{z} + \tilde{\mathbf{q}}^\top \mathbf{J} = 0 \quad \forall \mathbf{q} \in \mathbb{R}^2. \quad (12)$$

The (perspective) fundamental matrix has 7 degrees of freedom and lies on a nontrivial algebraic variety in  $\mathbb{R}^9$ , written  $\mathbb{F}$ . The affine fundamental matrix has 4 degrees of freedom and is a point in  $\mathbb{P}^4$ , as is reported in for instance (Hartley and Zisserman, 2003).

**Fundamental and projection matrices.** The fundamental matrix is an implicit reconstruction of the two cameras. Following (Luong and Vieville, 1996), canonical cameras for both the affine and perspective camera models are obtained by setting the source ( $3 \times 4$ ) camera matrix to  $(\mathbf{I} \ \mathbf{0})$  and the target one to  $\check{\mathcal{G}}_{\mathcal{F}} = ([\check{\mathbf{e}}']_{\times} \ \mathcal{F} \ \check{\mathbf{e}}')$ , where the target epipole  $\mathbf{e}'$  is defined by  $\mathcal{F}^{\top} \check{\mathbf{e}}' = \mathbf{0}$ . In the affine case, we write  $\mathcal{S}_{\mathcal{A}}$  the first two rows of  $\check{\mathcal{G}}_{\mathcal{F}}$ , the third row being  $(0 \ 0 \ 1 \ 0)$ . Within this canonical reconstruction basis, a 3D point with depth  $\delta$  can be written:<sup>6</sup>

$$\check{\mathbf{Q}} \sim \begin{pmatrix} \mathbf{q} \\ 1 \\ \delta \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{q}} \\ \delta \end{pmatrix}.$$

Note that scaled homogeneous coordinates  $\tilde{\mathbf{Q}}$  (leading to proper affine cameras) are obtained by  $\tilde{\mathbf{Q}} = \mathcal{M}\check{\mathbf{Q}}$ , where matrix  $\mathcal{M}$  simply swaps the third and fourth coordinates. Note that  $\mathcal{M}\mathcal{M} \sim \mathbf{I}$ . Reprojecting a 3D point in the target camera  $\check{\mathcal{G}}_{\mathcal{F}}$  gives the transfer equation:

$$\tilde{\mathbf{q}}' \sim \bar{\mathcal{G}}_{\mathcal{F}}\tilde{\mathbf{q}} + \check{\mathbf{g}}_{\mathcal{F}}\delta, \quad (13)$$

with  $\bar{\mathcal{G}}_{\mathcal{F}}$  the first three columns of  $\check{\mathcal{G}}_{\mathcal{F}}$  and  $\check{\mathbf{g}}_{\mathcal{F}}$  the fourth one. In the affine case, the target camera matrix  $\mathcal{S}_{\mathcal{A}}$  has size  $(2 \times 4)$ . Defining  $\bar{\mathcal{S}}_{\mathcal{A}}$  and  $\mathbf{s}_{\mathcal{A}}$  similarly to  $\bar{\mathcal{G}}_{\mathcal{F}}$  and  $\check{\mathbf{g}}_{\mathcal{F}}$ , the transfer equation (13) specializes to:

$$\mathbf{q}' = \bar{\mathcal{S}}_{\mathcal{A}}\tilde{\mathbf{q}} + \mathbf{s}_{\mathcal{A}}\delta. \quad (14)$$

A short calculation shows that:

$$\mathbf{s}_{\mathcal{A}} = \mathbf{e}' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{i} \quad \text{and} \quad \bar{\mathcal{S}}_{\mathcal{A}} = \mathbf{i}\mathbf{j}^{\top}.$$

An important remark that can be made here is that the target canonical projection matrix is defined up to a 4 degrees of freedom ambiguity, since  $\check{\mathcal{G}}_{\mathcal{F}} = (\check{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}}\boldsymbol{\pi}^{\top} \ \nu\check{\mathbf{g}}_{\mathcal{F}})$  is also a valid projection matrix,  $\forall \boldsymbol{\pi} \in \mathbb{R}^3$  and  $\forall \nu \in \mathbb{R}^*$ . This changes the 3D coordinate frame that is being implicitly used by the warps. However,

---

<sup>6</sup> $\delta$  is actually the inverse of the depth relative to the source camera. If the camera is calibrated this is the ‘true’ inverse depth, otherwise this is the inverse projective depth. The advantages of this 3D point parameterization are that it is minimal (*i.e.*, it has 3 effective parameters in  $\mathbf{q}$  and  $\delta$ ) and still it handles points at infinity.

we demonstrate in Appendix D that this choice has no influence on our two proposed rigid warps.

## 4 Warps with the Affine Camera Model

We derive the warps corresponding to an affine camera model, namely the DA-Warp and the RA-Warp. The RA-Warp forms the subset of the DA-Warps satisfying the rigidity constraints.

### 4.1 DA-Warp – The Standard TPS-Warp

We derive the classical DA-Warp from the TPS and give some of its properties, and a novel derivation of the warp in terms of a deformable 3D surface acquired by affine cameras. This interpretation shows that the DA-Warp is intrinsically affine and that it overfits affine images of a rigid surface, as will be illustrated by an example.

#### 4.1.1 Derivation

The standard  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  TPS-Warp is obtained by stacking two  $\mathbb{R}^2 \rightarrow \mathbb{R}$  TPS functions sharing their centres and internal regularization weight. Let  $\mathbf{P}' = (\mathbf{x}' \quad \mathbf{y}')$ . Using (4), we get:

$$\begin{pmatrix} \tau(\mathbf{q}; \mathbf{x}', \lambda) \\ \tau(\mathbf{q}; \mathbf{y}', \lambda) \end{pmatrix} = \left( \boldsymbol{\ell}_{\mathbf{q}}^\top \mathcal{E}_\lambda \begin{pmatrix} \mathbf{x}' & \mathbf{y}' \end{pmatrix} \right)^\top.$$

The DA-Warp is thus defined as:

$$\boxed{\mathcal{W}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\text{DA}} \boldsymbol{\ell}_{\mathbf{q}} \quad \text{with} \quad \mathcal{M}_{\text{DA}}^\top \stackrel{\text{def}}{=} \mathcal{E}_\lambda \mathbf{P}'} \quad (15)$$

We call this warp the Deformable Affine Thin-Plate Spline Warp, or ‘DA-Warp’ since, as shown below, it models images of deformable surfaces and corresponds to an affine camera model. Thanks to property (7), we can write an homogeneous version of the DA-Warp as:

$$\tilde{\mathcal{W}}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda) \stackrel{\text{def}}{=} \tilde{\mathcal{M}}_{\text{DA}} \boldsymbol{\ell}_{\mathbf{q}} \quad \text{with} \quad \tilde{\mathcal{M}}_{\text{DA}}^\top \stackrel{\text{def}}{=} \mathcal{E}_\lambda \tilde{\mathbf{P}}'. \quad (16)$$

### 4.1.2 Properties

The set of DA-Warps is written  $\mathbb{S}_{\text{DA}}$ . It has  $2l$  degrees of freedom through  $\mathbf{P}' \in \mathbb{R}^{2l}$ . The asymptotic internal regularization behaviour is as follows:

$$\lim_{\lambda \rightarrow +\infty} \mathcal{W}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda) = \mathcal{L}_{\text{DA}} \tilde{\mathbf{q}}' \quad \text{with} \quad \mathcal{L}_{\text{DA}}^{\top} \stackrel{\text{def}}{=} \tilde{\mathbf{P}}^{\dagger} \mathbf{P}'.$$

In other words, a DA-Warp tends to an FA-Warp represented by the  $(2 \times 3)$  matrix  $\mathcal{L}_{\text{DA}}$ . More precisely,  $\mathcal{L}_{\text{DA}}$  represents the FA-Warp minimizing the transfer error.<sup>7</sup> A proof is given in Appendix A.1. In the minimal case of  $l = 3$  centres, the DA-Warp is an FA-Warp given by  $\mathcal{W}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda) = \mathbf{P}'^{\top} \tilde{\mathbf{P}}^{-\top} \tilde{\mathbf{q}} = \mathcal{W}_{\text{FA}}(\mathbf{q}; \mathbf{P}'^{\top} \tilde{\mathbf{P}}^{-\top})$ .

The DA-Warp was originally introduced in (Bookstein, 1989), where its bending energy  $\kappa_{\text{DA}}$  is defined as the sum of the integral bending energies over the two coordinates, giving  $\kappa_{\text{DA}} = \text{tr}(\mathbf{P}'^{\top} \bar{\mathcal{E}}_{\lambda} \mathbf{P}')$ , where  $\text{tr}$  is matrix trace.

### 4.1.3 Projected Deformable Surface Interpretation

We propose a geometric interpretation of the DA-Warp, as the warp induced by the observation of a deforming surface with an affine camera: a 3D surface is observed in the source image, moves and deforms, and is observed again in the target image. This is illustrated in figure 4.

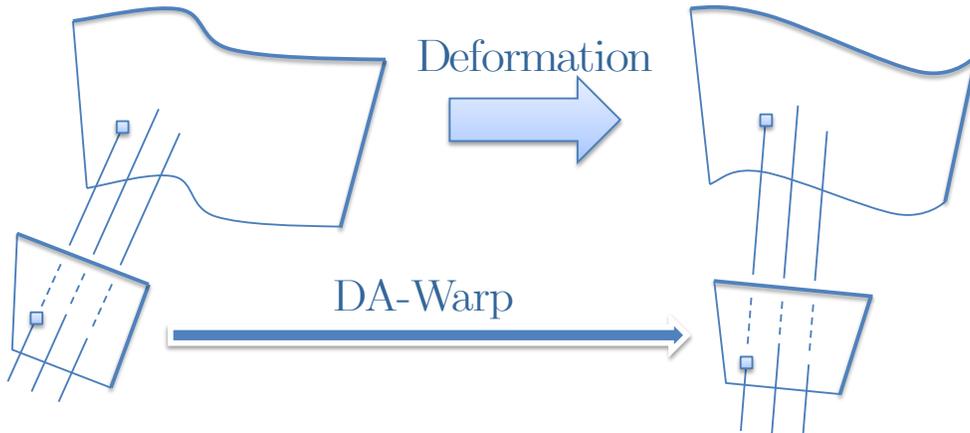


Figure 4: The DA-Warp relates the projection by two affine cameras of a deforming surface. Note that replacing the affine by the perspective projection subsequently leads to the proposed DP-Warp.

In order to model the surface depth, its displacement and deformation, we introduce an  $\mathbb{R}^2 \rightarrow \mathbb{R}^3$  map, parameterizing the target surface in a concise manner by the 3D coordinates  $\mathbf{C}'_k$  of the  $l$  centres. This map is built by stacking three TPS functions sharing their centres and internal regularization weight, similarly

<sup>7</sup>The transfer error is the discrepancy between the data points in the target image and the points transferred by the warp from the source image, see §7 for more details.

to the construction of the DA-Warp with two TPS functions. More formally, gathering the ‘3D centres’  $\mathbf{C}'_k$  in a single  $(l \times 3)$  matrix  $\mathbf{Z}^\top = (\mathbf{C}'_1 \ \cdots \ \mathbf{C}'_l)$ , the map is written:

$$\mathcal{R}(\mathbf{q}; \mathbf{Z}, \lambda) \stackrel{\text{def}}{=} \mathcal{M}_{3\text{D}} \ell_{\mathbf{q}} \quad \text{with} \quad \mathcal{M}_{3\text{D}}^\top \stackrel{\text{def}}{=} \mathcal{E}_\lambda \mathbf{Z}. \quad (17)$$

It will be seen that for the rigid warps, namely the RA-Warp and the RP-Warp, the 3D surface can be expressed as the graph of a function, or equivalently as a 2.5D surface or a Monge patch, with respect to the first camera. In the deformable case at hand, and also for the DP-Warp, such a representation is not possible, and one has to resort to the full  $\mathbb{R}^2 \rightarrow \mathbb{R}^3$  map  $\mathcal{R}$ . The reason for this is that  $\mathcal{R}$  contains the surface deformation that occurs between the two images. In other words, a 3D point as seen by the second camera cannot be parameterized as a point lying on the viewing ray of the corresponding point seen in the first camera, due to the surface deformation.

Reprojecting a 3D surface point in the target image gives:

$$\mathbf{q}' = \mathcal{S}_A \mathcal{M} \left( \mathbf{Z}^\top \mathcal{E}_\lambda^\top \ell_{\mathbf{q}} \right).$$

Using property (7) and defining  $\tilde{\mathbf{Z}} \stackrel{\text{def}}{=} (\mathbf{Z} \ \mathbf{1})$ , we get:

$$\mathbf{q}' = \mathcal{S}_A \mathcal{M} \begin{pmatrix} \mathbf{Z}^\top \mathcal{E}_\lambda^\top \\ \mathbf{1}^\top \mathcal{E}_\lambda^\top \end{pmatrix} \ell_{\mathbf{q}} = \mathcal{S}_A \mathcal{M} \tilde{\mathbf{Z}}^\top \mathcal{E}_\lambda^\top \ell_{\mathbf{q}},$$

that we identify with a DA-Warp (15), giving:

$$\mathbf{q}' = \mathcal{W}_{\text{DA}} \left( \mathbf{q}; \tilde{\mathbf{Z}} \mathcal{M} \mathcal{S}_A^\top, \lambda \right).$$

This shows that the centres  $\mathbf{c}'_k$  in the target image are the reprojection of the 3D centres  $\mathbf{C}'_k$  *i.e.*,  $\mathbf{P}' = \tilde{\mathbf{Z}} \mathcal{M} \mathcal{S}_A^\top$ . The DA-Warp therefore depends on the camera matrix  $\mathcal{S}_A$  and on the observed surface parameterized by the ‘3D centres’ in matrix  $\mathbf{Z}$ , both encapsulated in the target image centres.

This geometric interpretation does not only provide a strong intuition on the fact that the DA-Warp is intrinsically affine, but gives a setting for naturally deriving the DP-Warp, the perspective projection extension of the DA-Warp in §5.2. The 3D integral bending energy is obtained by summing over the three coordinate axes, giving  $\kappa_{3\text{D}} = \text{tr}(\mathbf{Z}^\top \bar{\mathcal{E}}_\lambda \mathbf{Z})$ . We note that due to the affine projection, the DA-Warp only depends on the projection of the 3D centres. Therefore, the DA-Warp minimizes an image, ‘projected’ integral bending energy. The DP-Warp however minimizes the 3D integral bending energy  $\kappa_{3\text{D}}$ .

#### 4.1.4 Example

We show two examples of the DA-Warp: one with perspective images and one with affine images of a rigid surface.

**Perspective images of a flat surface.** The DA-Warp is known to be a powerful means to represent the warp between two images of a smooth surfaces. We show an example illustrating that the DA-Warp fails to capture the perspective part of the transformation. The source image is FP1 and the target image is FP2, both showing the poster flat but with significant perspective effects. Figure 5 shows a DA-Warp estimated with various number of centres  $l \in \{3, 9, 30\}$ .



Figure 5: The DA-Warp estimated between two perspective images of a flat surface. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres improves the visual quality of the warp. However, using more centres degrades the smoothness of the warp. The fitting transfer error RMSR is given. For  $l = 3$  centres, the DA-Warp reduces to an FA-Warp, and the strong effect of perspective prevents the visualization grid transferred from the source image to fit the poster well.

Estimating an FP-Warp *i.e.*, an homography, gives an RMSR of 1.39 pixels, while estimating an FA-Warp *i.e.*, an affine transformation, gives an RMSR of 24.67 pixels. This confirms the significant perspective effect in these images. As can be seen, the DA-Warp does not model the perspective effect well, even though the ‘true’ warp can be well described by as few as 8 parameters by the FP-Warp.<sup>8</sup> Adding centres makes the flexibility of the DA-Warp to model the perspective. This however makes the warp ‘unstable’ since it has many more degrees of freedom than what is necessary.

**Affine images of a rigid surface.** The example we now show illustrates the fact that a DA-Warp applied to two affine images of a rigid smooth surface captures the rigidity. The source image is DA11 and the target image is DA12, showing the poster under the same deformation in an affine imaging configuration. Figure 6

<sup>8</sup>The 1.39 pixels RMSR comes from the noise in the manually entered point positions, unmodeled camera effects such as radial distortion and non-perfect planarity of the surface.

shows a DA-Warp estimated with various number of centres  $l \in \{3, 9, 30\}$ .

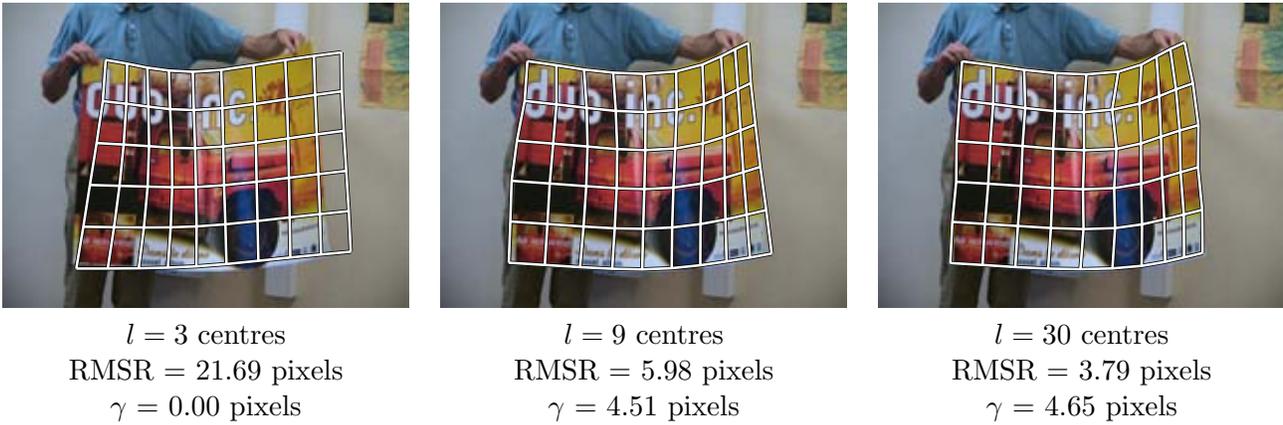


Figure 6: The DA-Warp estimated between two affine images of a rigid smooth surface. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres improves the visual quality of the warp. The fitting transfer error RMSR and a measure of non-rigidity  $\gamma$  are given.

Estimating the affine epipolar geometry gave a transfer error of 3.71 pixels, which shows to which level the studied pair of images complies with a rigid affine model. We computed a measure of non-rigidity  $\gamma$  of the warp, related to the definition of a rigid affine warp (12). The functional is defined as the discrepancy between the epipolar lines and the warp predicted points in the target image, over  $n$  points  $\mathbf{p}_j$  sampled in the source image:

$$\gamma^2(\mathcal{W}; \mathcal{F}) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{j=1}^n d_{\perp}^2(\mathcal{W}(\mathbf{p}_j), \mathcal{F}\mathbf{p}_j),$$

where  $d_{\perp}$  is the point to line distance,  $\mathcal{W}$  an  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  warp and  $\mathcal{F}$  a fundamental matrix. The values given in figure 6 for  $\gamma$  are small, and of the same order as the affine epipolar geometry RMSR. This suggests that the DA-Warp follows the epipolar lines, and thus complies with the rigidity constraint. The order of magnitude of  $\gamma$  tells us that the rigidity is up to the noise contained in the point positions.

## 4.2 RA-Warp – The Rigid Affine TPS Warp

The RA-Warp has very similar properties to the DA-Warp, but is a rigid warp, in the sense that it exactly fulfills the rigidity constraint modeled by the affine epipolar geometry.

### 4.2.1 Derivation

Our main result for this section is that a DA-Warp is a rigid warp if and only if the  $l$  centres defining the warp satisfy the epipolar constraint. Such a warp is called an RA-Warp. As will be seen, it has  $l + 4$  degrees of freedom.

The demonstration is as follows. Applying the rigid warp definition (12) to a DA-Warp (15) gives:

$$\ell_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} \mathbf{P}' \mathbf{i} + \tilde{\mathbf{q}}^{\top} \mathbf{j} = 0 \quad \forall \mathbf{q} \in \mathbb{R}^2.$$

Using property (6) and factorizing gives:

$$\begin{aligned} \ell_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} (\mathbf{P}' \mathbf{i} + \tilde{\mathbf{P}} \mathbf{j}) &= 0 \quad \forall \mathbf{q} \in \mathbb{R}^2 \\ \mathbf{P}' \mathbf{i} + \tilde{\mathbf{P}} \mathbf{j} &= \mathbf{0}_{(l \times 1)}, \end{aligned}$$

which is the epipolar constraint for all pairs of centres. This is the *rigidity constraint for the DA-Warp*. This means that the warp satisfies the epipolar geometry if and only if the centres themselves satisfy it.

An alternative proof is to assume that each pair of centres  $\mathbf{c}_k \leftrightarrow \mathbf{c}'_k$  satisfies the epipolar constraint and thus is the projection of a 3D point  $\mathbf{C}_k$  with  $\mathbf{C}_k^{\top} = (\mathbf{c}_k^{\top} \ \delta_k)$ . Reprojecting all the centres in the target image gives:

$$\mathbf{P}'^{\top} = \mathcal{S}_{\mathcal{A}} \begin{pmatrix} \tilde{\mathbf{P}}^{\top} \\ \boldsymbol{\delta}^{\top} \end{pmatrix}.$$

Substituting into the DA-Warp (15) gives:

$$\mathcal{W}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda) = \mathcal{S}_{\mathcal{A}} \begin{pmatrix} \tilde{\mathbf{P}}^{\top} \\ \boldsymbol{\delta}^{\top} \end{pmatrix} \mathcal{E}_{\lambda}^{\top} \ell_{\mathbf{q}}, \quad (18)$$

which is the projection of some 3D point  $(\ell_{\mathbf{q}}^{\top} \mathcal{E}_{\lambda} (\tilde{\mathbf{P}} \ \boldsymbol{\delta}))^{\top}$  by the target camera, thereby satisfying the rigidity constraint, which completes the proof.

The RA-Warp is thus a DA-Warp with centres satisfying the rigidity constraint. This is used in the next section to find a formulation for the RA-Warp based on the epipolar geometry and a depth map. We write the RA-Warp using reprojected points as centres for the DA-Warp:

$$\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \mathcal{W}_{\text{DA}}\left(\mathbf{q}; \begin{pmatrix} \tilde{\mathbf{P}} \\ \boldsymbol{\delta} \end{pmatrix} \mathcal{S}_{\mathcal{A}}^{\top}, \lambda\right).$$

The RA-Warp depends on the depth  $\boldsymbol{\delta}$  of the centres and on the target camera  $\mathcal{S}_{\mathcal{A}}$  which itself depends on the affine fundamental matrix  $\mathcal{A}$ . We define:

$$\boxed{\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\text{RA}} \ell_{\mathbf{q}} \quad \text{with} \quad \mathcal{M}_{\text{RA}}^{\top} \stackrel{\text{def}}{=} \mathcal{E}_{\lambda} \begin{pmatrix} \tilde{\mathbf{P}} \\ \boldsymbol{\delta} \end{pmatrix} \mathcal{S}_{\mathcal{A}}^{\top}.} \quad (19)$$

This definition of the RA-Warp can be made homogeneous by replacing the  $(2 \times 4)$  camera  $\mathcal{S}_{\mathcal{A}}$  by its  $(3 \times 4)$  equivalent  $\mathcal{G}_{\mathcal{A}}$  in the above equation, giving:

$$\tilde{\mathcal{W}}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) \stackrel{\text{def}}{=} \tilde{\mathcal{M}}_{\text{RA}} \ell_{\mathbf{q}} \quad \text{with} \quad \tilde{\mathcal{M}}_{\text{RA}}^{\top} \stackrel{\text{def}}{=} \mathcal{E}_{\lambda} \left( \tilde{\mathcal{P}} \boldsymbol{\delta} \right) \mathcal{G}_{\mathcal{A}}^{\top}. \quad (20)$$

#### 4.2.2 Properties

From the above derivation follows that the set of RA-Warps, denoted  $\mathbb{S}_{\text{RA}}$ , is a subset of  $\mathbb{S}_{\text{DA}}$ :

$$\mathbb{S}_{\text{RA}} \subset \mathbb{S}_{\text{DA}}.$$

The set of FA-Warps  $\mathbb{S}_{\text{FA}}$  is included into  $\mathbb{S}_{\text{RA}}$ , as demonstrated in Appendix C.1. The RA-Warp has  $l + 4$  degrees of freedom through  $(\boldsymbol{\delta}, \mathcal{A}) \in \mathbb{R}^l \times \mathbb{P}^4$ . We interpret the parameters  $\boldsymbol{\delta}$  as the depth of the centres with respect to the source camera.

The asymptotic internal regularization behaviour of the RA-Warp is derived directly from the one for the DA-Warp:

$$\begin{aligned} \lim_{\lambda \rightarrow +\infty} \mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) &= \lim_{\lambda \rightarrow +\infty} \mathcal{W}_{\text{DA}}\left(\mathbf{q}; \left(\tilde{\mathcal{P}} \boldsymbol{\delta}\right) \mathcal{S}_{\mathcal{A}}^{\top}, \lambda\right) \\ &= \mathcal{L}_{\text{RA}} \tilde{\mathbf{q}} \quad \text{with} \quad \mathcal{L}_{\text{RA}}^{\top} \stackrel{\text{def}}{=} \tilde{\mathcal{P}}^{\dagger} \left(\tilde{\mathcal{P}} \boldsymbol{\delta}\right) \mathcal{S}_{\mathcal{A}}^{\top}. \end{aligned}$$

In other words, an RA-Warp tends to an FA-Warp represented by the  $(2 \times 3)$  matrix  $\mathcal{L}_{\text{RA}}$ . We prove in Appendix A.2 that this FA-Warp is rigid and can be written as a plane-induced affine warp:

$$\mathcal{L}_{\text{RA}} = \bar{\mathcal{S}}_{\mathcal{A}} + \mathbf{s}_{\mathcal{A}} \boldsymbol{\delta}^{\top} \tilde{\mathcal{P}}^{\dagger}.$$

Note that  $\boldsymbol{\pi}^{\top} = \boldsymbol{\delta}^{\top} \tilde{\mathcal{P}}^{\dagger}$  is the reduced plane equation defined such that the transfer error is minimized, under the assumption that the point correspondences satisfy the rigidity constraint, as shown in Appendix B.1.

In the minimal  $l = 3$  centres case, we substitute the TPS (9) in the RA-Warp (21), which gives  $\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \bar{\mathcal{S}}_{\mathcal{A}} \tilde{\mathbf{q}} + \mathbf{s}_{\mathcal{A}} \boldsymbol{\delta}^{\top} \tilde{\mathcal{P}}^{-\top} \tilde{\mathbf{q}}$ . Factorizing gives  $\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \left(\bar{\mathcal{S}}_{\mathcal{A}} + \mathbf{s}_{\mathcal{A}} \boldsymbol{\delta}^{\top} \tilde{\mathcal{P}}^{-\top}\right) \tilde{\mathbf{q}}$ , which, as in the asymptotic regularization case, is a plane affinity.

#### 4.2.3 Projected Rigid Surface Interpretation

A geometric interpretation of the RA-Warp directly stems from its definition (19). The RA-Warp is induced by a surface defined as a 2.5D surface parameterized by a TPS mapping points from the source image to

their depths. This  $\mathbb{R}^2 \rightarrow \mathbb{R}$  TPS is of the form (4) and has the same source centres as the RA-Warp. This is derived by expanding the formulation (19) of the RA-Warp and using property (6):

$$\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \bar{\mathcal{S}}_{\mathcal{A}} \tilde{\mathbf{q}} + \mathbf{s}_{\mathcal{A}T}(\mathbf{q}; \boldsymbol{\delta}, \lambda). \quad (21)$$

This interpretation is illustrated in figure 7. We note that the RA-Warp explicitly depends on the target camera, contrarily to the DA-Warp.

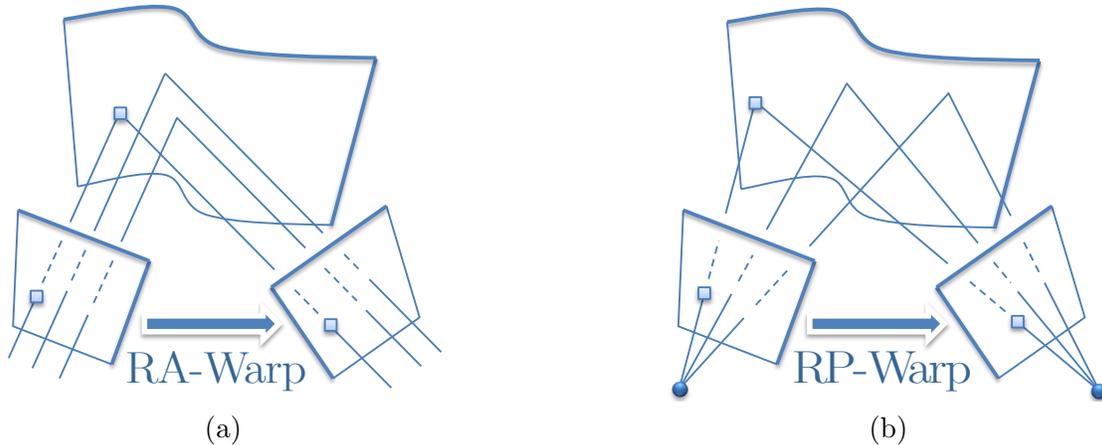


Figure 7: (a) The RA-Warp is interpreted as relating the projection by two affine cameras of a rigid and smooth surface. (b) Replacing the affine by the perspective projection subsequently leads to the RP-Warp.

#### 4.2.4 Example

We show an example of the RA-Warp applied to affine and perspective images of a rigid surface.

**Affine images of a rigid surface.** We show the result of estimating an RA-Warp on the same data as in §4.1.4 for a rigid smooth surface. Our goal is to compare the DA-Warp and the RA-Warp on this particular example. Figure 8 shows the result. The RMSR we obtained for the RA-Warp is slightly higher than for the DA-Warp. This difference is due to two factors. First, the approximation made when using the affine camera model onto which the RA-Warp is based. Second, the noise in the point positions that were manually specified. As previously mentioned, fitting the affine epipolar geometry gives a transfer error of 3.71 pixels. This provides a lower bound for the RMSR of any rigid affine warp. We recall that for a fixed number of centres, the RA-Warp has fewer degrees of freedom than the DA-Warp. For instance, for  $l = 30$  centres, the DA-Warp has  $2l = 60$  degrees of freedom while the RA-Warp has  $l + 4 = 34$  degrees of freedom. This shows that with nearly half the number of parameters of the DA-Warp, the RA-Warp has the same order of representational power (*i.e.*, ability to represent the image deformation), obtained by

explicitly taking advantage of the surface rigidity.



Figure 8: The RA-Warp estimated between two affine images of a rigid smooth surface. This figure matches figure 6 showing the DA-Warp estimated for the same images and centres. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres improves the visual quality of the warp.

**Perspective images of a rigid surface.** We estimated the RA-Warp on a pair of images showing the poster rigid but with significant perspective effect, so as to illustrate the behavior of the RA-Warp in this case. The source image we used is DP11 and the target image is DP12. The RMSR for the affine epipolar geometry is 13.77 pixels, but is 0.57 pixels for the (perspective) epipolar geometry. This shows that the images are rigidly linked, but do not comply very well with the affine camera model. Figure 9 shows the result for various number of centres  $l \in \{3, 9, 30\}$ . Increasing the number of centres does not improve the visual quality of the RA-Warp much. The reason is that the RMSR of the RA-Warp is bounded by 13.77 pixels, the RMSR of the affine epipolar geometry. The RMSR for the DA-Warp is 21.12 pixels, 7.23 pixels and 3.94 pixels for  $l = \{3, 9, 30\}$  centres respectively. This shows that by increasing the number of centres, and thus the flexibility of the DA-Warp, it can fit a perspective transformation, as we discuss in the next section. It will be seen that our RP-Warp has much less degrees of freedom than the DA-Warp and leads to a lower RMSR.

### 4.3 Behaviour with the Perspective Camera Model

As pointed out in (Wills and Belongie, 2004), the DA-Warp, and thus the RA-Warp, are not able to model the effect of perspective projection. This stems from the lack of a fraction in the warp formulations (15) and (19), that are linear combinations of deformation centres. We confirmed this with our geometric interpretation in terms of surfaces projected with an affine camera model. Below, we derive a more rigorous argument. Writing the perspective rigidity constraint (10) for a DA-Warp in homogeneous coordinates (16)



Figure 9: The RA-Warp estimated between two perspective images of a rigid smooth surface. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres does not improve much the visual quality of the warp.

gives:

$$\begin{aligned} \tilde{\mathcal{W}}_{\text{DA}}(\mathbf{q}; \mathbf{P}', \lambda)^\top \mathcal{F} \tilde{\mathbf{q}} &= 0 & \forall \mathbf{q} \in \mathbb{R}^2 \\ \ell_{\mathbf{q}}^\top \mathcal{E}_\lambda \tilde{\mathbf{P}}' \mathcal{F} \tilde{\mathbf{q}} &= 0 & \forall \mathbf{q} \in \mathbb{R}^2. \end{aligned}$$

This leads to  $\mathcal{E}_\lambda \tilde{\mathbf{P}}' \mathcal{F} = 0$ , meaning that the rigidity constraint implies that all the centres must lie on the target epipole  $\mathbf{e}'$ . The only eligible warp maps all points onto  $\mathbf{e}'$  and is thus degenerate.

We showed in §§4.1.4 and 4.2.4 on an example that the DA-Warp and the RA-Warp do indeed not behave well in the presence of perspective effects. Figures 5 and 18 show that however, with a ‘large’ number of centres, the DA-Warp approximates the perspective transformation to a large extent, even though the latter has few degrees of freedom. This is because the TPS is dense in the space of diffeomorphisms of the image domain. The space of diffeomorphisms contains perspective transformations, and therefore a TPS can approximate perspective transformations arbitrarily well, although this may require a large number of centres. The RA-Warp however cannot approximate a perspective transformation, as figure 9 shows. This is because the camera geometry the RA-Warp uses is affine and cannot be made more flexible by increasing the number of centres. Our goal is now to embed the perspective projection model into these two affine warps, the DA-Warp and the RA-Warp, to allow them to gracefully deal with perspective effects.

## 5 Warps with the Perspective Camera Model

We propose two warps: the RP-Warp and the DP-Warp, which introduce perspective projection in the affine warps of the previous section.

## 5.1 RP-Warp – The Rigid Perspective TPS Warp

The RP-Warp is an extension of the RA-Warp to perspective projection, as figure 7 illustrates.

### 5.1.1 Derivation

We derive the RP-Warp by introducing a perspective projection in the RA-Warp. The RP-Warp thus satisfies the epipolar geometry by construction. Following the reasoning in §4.2, we pick up a 3D point  $\mathbf{Q}$  on the scene surface, defined by an  $\mathbb{R}^2 \rightarrow \mathbb{R}$  TPS parameterized 2.5D surface, and reproject it in the target image, giving from equation (21):

$$\check{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}} + \check{\mathbf{g}}_{\mathcal{F}} \tau(\mathbf{q}; \boldsymbol{\delta}, \lambda). \quad (22)$$

Replacing  $\tau$  by its expression (4), applying property (6) to the first term and factorizing we get:

$$\check{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^{\top} \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}} \sim \left( \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^{\top} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \right) \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}},$$

and thus:

$$\check{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \stackrel{\text{def}}{\sim} \check{\mathcal{M}}_{\text{RP}} \boldsymbol{\ell}_{\mathbf{q}} \quad \text{with} \quad \check{\mathcal{M}}_{\text{RP}}^{\top} \stackrel{\text{def}}{\sim} \mathcal{E}_{\lambda} \left( \tilde{\mathbf{P}} \boldsymbol{\delta} \right) \mathcal{G}_{\mathcal{F}}^{\top}. \quad (23)$$

This is the homogeneous Rigid Perspective Thin-Plate Spline Warp. The homogeneous coordinates of the transferred image point are linear functions of  $\boldsymbol{\ell}_{\mathbf{q}}$ . The affine coordinates are obtained as ratios of linear functions through  $\mathcal{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \stackrel{\text{def}}{=} \psi(\check{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda))$ .

### 5.1.2 Properties

The set of RP-Warps, denoted  $\mathbb{S}_{\text{RP}}$ , is a superset of  $\mathbb{S}_{\text{RA}}$ :

$$\mathbb{S}_{\text{RA}} \subset \mathbb{S}_{\text{RP}}.$$

This is shown easily by choosing for  $\mathcal{F}$  an affine fundamental matrix.  $\mathbb{S}_{\text{RP}}$  is also a superset of  $\mathbb{S}_{\text{FP}}$ , as shown in Appendix C.1. An RP-Warp is guaranteed to be rigid since it implicitly projects 3D points, giving image points satisfying the epipolar constraint. It has  $l + 7$  degrees of freedom through  $(\boldsymbol{\delta}, \mathcal{F}) \in \mathbb{R}^l \times \mathbb{F}$ , where  $\mathbb{F}$  is the 7-dimensional space of fundamental matrices.

The asymptotic internal regularization behaviour is:

$$\lim_{\lambda \rightarrow +\infty} \check{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim \check{\mathcal{L}}_{\text{RP}} \tilde{\mathbf{q}} \quad \text{with} \quad \check{\mathcal{L}}_{\text{RP}}^{\top} \stackrel{\text{def}}{\sim} \tilde{\mathbf{P}}^{\dagger} \left( \tilde{\mathbf{P}} \boldsymbol{\delta} \right) \mathcal{G}_{\mathcal{F}}^{\top}. \quad (24)$$

An RP-Warp thus tends to an FP-Warp represented by the  $(3 \times 3)$  homogeneous homography matrix  $\check{\mathcal{L}}_{\text{RP}}$ . We prove in Appendix A.2 that this is a plane-induced rigid warp *i.e.*, that  $\check{\mathcal{L}}_{\text{RP}}$  factors as:

$$\check{\mathcal{L}}_{\text{RP}} \sim \bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \tilde{\mathbf{P}}^{\dagger}.$$

Note that  $\boldsymbol{\pi}^{\top} = \boldsymbol{\delta}^{\top} \tilde{\mathbf{P}}^{\dagger}$  is the reduced plane equation that minimizes a weighted algebraic transfer error, under the assumption that the point correspondences satisfy the rigidity constraint, as shown in Appendix B.2.

In the minimal  $l = 3$  centres case, we substitute the TPS (9) in the RP-Warp (22), which gives  $\check{\mathcal{W}}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \tilde{\mathbf{P}}^{-\top} \tilde{\mathbf{q}}$ . Factorizing gives  $\check{\mathcal{W}}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim \left( \bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \tilde{\mathbf{P}}^{-\top} \right) \tilde{\mathbf{q}}$ , which, as in the asymptotic regularization case, is a plane homography.

### 5.1.3 Example

We show how the RP-Warp deals with the perspective images of a rigid surface used to illustrate the RA-Warp in §4.2.4, with the same experimental setup. Contrarily to the RA-Warp, as the number of centres grows, the RMSR significantly decreases for the RP-Warp, while the visual quality of the warp improves. This shows that the RP-Warp models perspective well. For the particular minimal case of  $l = 3$  centres, the RP-Warp is an FP-Warp.

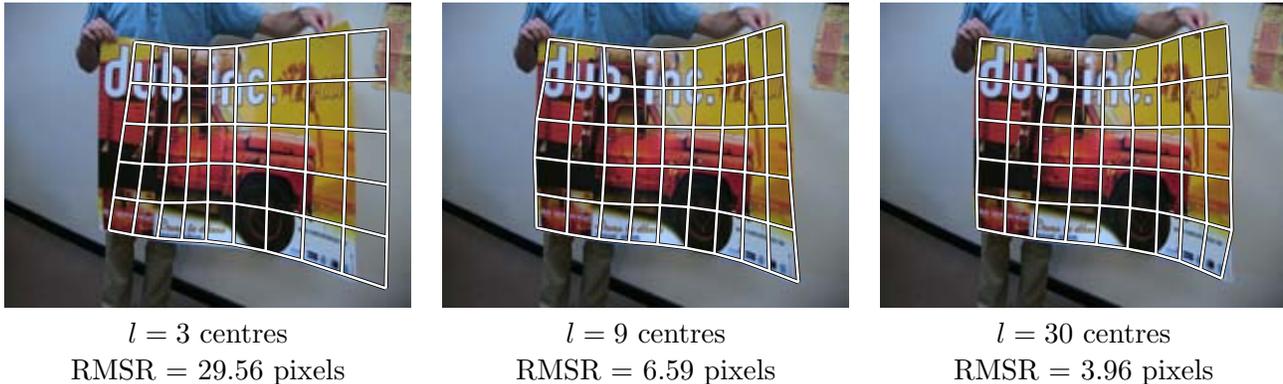


Figure 10: The RP-Warp estimated between two perspective images of a rigid smooth surface. The figure matches figure 9 showing the RA-Warp estimated for the same images and centres. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres does improve the visual quality of the warp. For  $l = 3$  centres, the RP-Warp reduces to a plane-induced homography. The visualization grid is distorted in that case since the source visualization grid itself comes from an image where the surface is deformed.

## 5.2 DP-Warp – The Deformable Perspective TPS Warp

The DP-Warp extends the DA-Warp by introducing perspective projection.

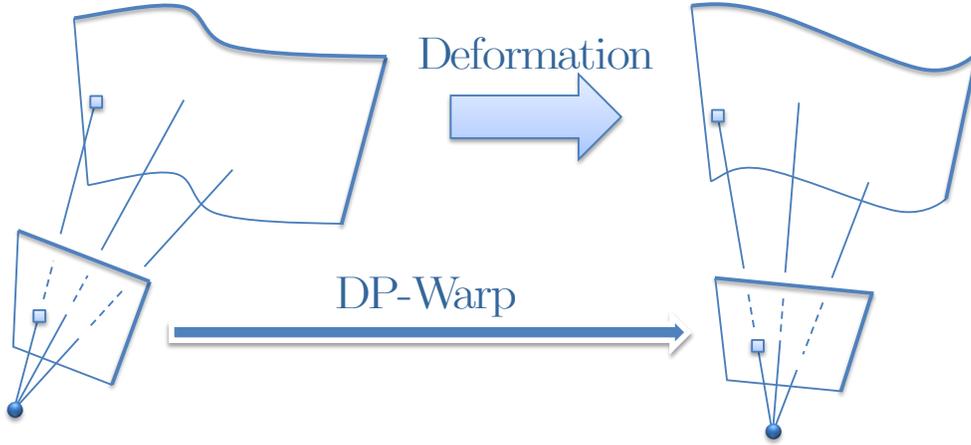


Figure 11: The DP-Warp relates the projection by two perspective cameras of a deforming surface. It is an extension of the classical DA-Warp to perspective projection.

### 5.2.1 Derivation

The DP-Warp forms a superset of all the other warps derived so far in this paper, including the standard DA-Warp. The RP-Warp is derived by introducing perspective projection in the RA-Warp, as figure 7 illustrates. We derive the DP-Warp from the DA-Warp using the same trick. We consider the deformable surface geometric interpretation shown in figure 4. The surface seen by the target camera is defined by an  $\mathbb{R}^2 \rightarrow \mathbb{R}^3$  map  $\mathcal{R}(\mathbf{q}; \mathbf{Z}, \lambda) = \mathbf{Z}^\top \mathcal{E}_\lambda^\top \ell_{\mathbf{q}}$  (see equation (17)). Following §4.1.3, we project a point on this surface to the target image, giving:

$$\check{\mathbf{q}}' \sim \check{\mathcal{G}}_{\mathcal{F}} \mathcal{M} \begin{pmatrix} \mathbf{Z}^\top \mathcal{E}_\lambda^\top \ell_{\mathbf{q}} \\ 1 \end{pmatrix}.$$

This process is illustrated in figure 11. Using property (7) and factorizing, we get:

$$\check{\mathbf{q}}' \sim \check{\mathcal{G}}_{\mathcal{F}} \mathcal{M} \tilde{\mathbf{Z}}^\top \mathcal{E}_\lambda^\top \ell_{\mathbf{q}}.$$

The centres  $\check{\mathbf{P}}' \sim \tilde{\mathbf{Z}} \mathcal{M} \check{\mathcal{G}}_{\mathcal{F}}^\top$  in the target image are the reprojection of the ‘3D centres’ in matrix  $\mathbf{Z}$ . The weights of the homogeneous coordinates in  $\check{\mathbf{P}}'$  are important: they model the perspective part of the DP-Warp, which we define as:

$$\check{\mathcal{W}}_{\text{DP}}(\mathbf{q}; \check{\mathbf{P}}', \lambda) \stackrel{\text{def}}{\sim} \check{\mathcal{M}}_{\text{DP}} \ell_{\mathbf{q}} \quad \text{with} \quad \check{\mathcal{M}}_{\text{DP}}^\top \stackrel{\text{def}}{\sim} \mathcal{E}_\lambda \check{\mathbf{P}}'. \quad (25)$$

The affine coordinates are obtained as ratios of linear functions through  $\mathcal{W}_{\text{DP}}(\mathbf{q}; \check{\mathbf{P}}', \lambda) \stackrel{\text{def}}{=} \psi(\check{\mathcal{W}}_{\text{DP}}(\mathbf{q}; \check{\mathbf{P}}', \lambda))$ . Similarly to the DA-Warp, the DP-Warp depends on the camera matrix  $\check{\mathcal{G}}_{\mathcal{F}}$  and the ‘3D centres’ in matrix  $\mathbf{Z}$ , through the target image centres.

### 5.2.2 Properties

The set of DP-Warps, denoted  $\mathbb{S}_{\text{DP}}$ , forms a superset of all the warps studied in this paper, and in particular of  $\mathbb{S}_{\text{RP}}$  and  $\mathbb{S}_{\text{DA}}$ :

$$\mathbb{S}_{\text{RP}} \subset \mathbb{S}_{\text{DP}} \quad \text{and} \quad \mathbb{S}_{\text{DA}} \subset \mathbb{S}_{\text{DP}}.$$

The DP-Warp has its parameters in the  $(l \times 3)$  matrix  $\check{\mathbf{P}}'$  defined up to scale and thus has  $3l - 1$  degrees of freedom.<sup>9</sup>

The asymptotic internal regularization behaviour is formulated below for all data points chosen as centres. A consequence is that the limiting warp we get is undetermined *i.e.*, it has some free parameters. Unsurprisingly, it actually has  $(3l - 1) - 2l = l - 1$  free parameters *i.e.*, the difference between the number of free parameters of the DP-Warp and the number of constraints given by interpolating the  $l$  centres:

$$\lim_{\lambda \rightarrow +\infty} \check{\mathcal{W}}_{\text{DP}}(\mathbf{q}; \check{\mathbf{P}}', \lambda) \sim \check{\mathcal{L}}_{\text{DP}} \tilde{\mathbf{q}} \quad \text{with} \quad \check{\mathcal{L}}_{\text{DP}}^{\top} \stackrel{\text{def}}{\sim} \tilde{\mathbf{P}}^{\dagger} \text{diag}(\mathbf{d}) \tilde{\mathbf{P}}'.$$

The  $(l \times 1)$  vector  $\mathbf{d}$ , defined up to scale, represents the  $l - 1$  free parameters of the limiting FP-Warp, represented by matrix  $\check{\mathcal{L}}_{\text{DP}}$ . This corresponds to the FP-Warp minimizing some algebraic transfer error. More details are given in Appendix A.3. In the minimal  $l = 3$  case, the DP-Warp is an FP-Warp given by  $\check{\mathcal{W}}_{\text{DP}}(\mathbf{q}; \check{\mathbf{P}}', \lambda) \sim \check{\mathbf{P}}'^{\top} \tilde{\mathbf{P}}^{-\top} \tilde{\mathbf{q}} \sim \check{\mathcal{W}}_{\text{FP}}(\mathbf{q}; \check{\mathbf{P}}'^{\top} \tilde{\mathbf{P}}^{-\top})$ .

The 3D integral bending energy  $\kappa_{3\text{D}}$  is identical to the one given for the DA-Warps in §4.1.3.

### 5.2.3 Example

We show the result of estimating a DP-Warp on the same data as in §4.1.4 for a flat surface. Our goal is to show that the DP-Warp adds the perspective part ‘missing’ to the DA-Warp. Figure 12 shows the result. What we first observe is that the DP-Warp models the actual warp very well, whatever the number of centres. In the particular case of  $l = 3$  centres, the DP-Warp becomes an FP-Warp, and has a 1.39 pixels RMSR. The DA-Warp cannot model the perspective and requires many more centres to achieve a ‘low’ RMSR. As the number of centre grows, the RMSR for the DP-Warp slightly decreases since the extra

<sup>9</sup>Consequently, and in the absence of external regularization, the DP-Warp cannot be estimated by choosing as centres all data points: each point correspondence giving two constraints, we end up with  $2l$  constraints, which is less than the  $3l - 1$  unknowns. Methods for estimating the DP-Warp are reported in §7.

degrees of freedom model the noise in the point position. The resulting warp however keeps a nice visual behavior.



Figure 12: The DP-Warp estimated between two perspective images of a flat surface. This figure matches figure 5 showing the DA-Warp estimated for the same images and centres. Only the target image is shown, overlaid with the warp visualization grid transferred from the source image with the estimated warp. As can be seen, increasing the number of centres causes overfitting.

## 6 A Hierarchy of Warps

The aim of this section is to define a hierarchy between the sets of standard DA-Warps  $\mathbb{S}_{DA}$ , of FA-Warps and FP-Warp  $\mathbb{S}_{FA}$  and  $\mathbb{S}_{FP}$ , and of the three types of warps we introduced,  $\mathbb{S}_{RA}$ ,  $\mathbb{S}_{RP}$  and  $\mathbb{S}_{DP}$ . The four most important warps were summarized in figure 1. The proofs of the statements below are given in Appendix C. The whole hierarchy is illustrated in figure 13. So far, we have established  $\mathbb{S}_{RA} \subset \mathbb{S}_{DA}$  and  $\mathbb{S}_{RA} \subset \mathbb{S}_{RP}$ . Intuitively, the common warps to  $\mathbb{S}_{DA}$  and  $\mathbb{S}_{RP}$  must be rigid and affine. More precisely, we have  $\mathbb{S}_{RA} = \mathbb{S}_{DA} \cap \mathbb{S}_{RP}$ . We have also established that  $\mathbb{S}_{DP}$  is a superset of all the other warps *i.e.*,  $\mathbb{S}_{RP} \subset \mathbb{S}_{DP}$  and  $\mathbb{S}_{DA} \subset \mathbb{S}_{DP}$ , and thus  $\mathbb{S}_{RA} \subset \mathbb{S}_{DP}$ . The set of DA-Warps  $\mathbb{S}_{DA}$  does not contain any flat perspective warp. More formally,  $(\mathbb{S}_{FP} \setminus \mathbb{S}_{FA}) \cap \mathbb{S}_{DA} = \emptyset$ , implying  $(\mathbb{S}_{FP} \setminus \mathbb{S}_{FA}) \cap \mathbb{S}_{RA} = \emptyset$ .

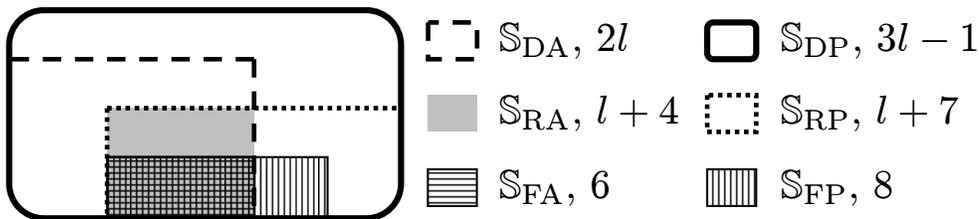


Figure 13: Hierarchical representation for the three proposed types of warps – the RA-Warp, the RP-Warp and the DP-Warp – along with the standard TPS warp, dubbed DA-Warp, and the flat warps FA-Warp and FP-Warp. ‘D’ stands for Deformable, ‘R’ for rigid, ‘F’ for flat, ‘A’ for Affine and ‘P’ for Perspective. The number of degrees of freedom for  $l$  centres is indicated in the legend for each warp.

We define what we call the *intrinsic dimension* of each warp as the number of  $\mathbb{R}^2 \rightarrow \mathbb{R}$  TPS functions

implied in its formulation, or equivalently the scalar factor multiplying  $l$  in its number of degrees of freedom, giving the following dimensions:

Structure	Camera model	Warp	Number of parameters	Intrinsic dimension
Flat	Affine	FA	6	0
	Perspective	FP	8	0
Rigid	Affine	RA	$l + 4$	1
	Perspective	RP	$l + 7$	1
Deformable	Affine	DA	$2l$	2
	Perspective	DP	$3l - 1$	3

The intrinsic dimension is an important cue: it says if the warp can be determined by choosing all data points as centres, and if extra constraints are left. More precisely, if the intrinsic dimension is 0, 1 or 2, then the warp is in general uniquely determined. Only the DP-Warp does not fulfill this condition. If the intrinsic dimension is 0 or 1, then the warp is uniquely determined and extra constraints are left. For the DA-Warp, once the warp is defined and interpolates the centres, then no other constraint is left, while for instance for the RP-Warp, the points are constrained to lie on the epipolar lines.

## 7 Estimation of the Warps

We propose methods for the estimation of each type of warps presented in this paper from point correspondences.

### 7.1 General Points

The  $m$  point correspondences are written  $\mathbf{q}_j \leftrightarrow \mathbf{q}'_j$ . The  $l$  centres in the source image in matrix  $\mathbf{P}$  are typically chosen on a fixed size regular grid. Broadly speaking, the unknowns are the corresponding centres in the target image.

All warps are estimated by minimizing the *transfer error*. It is defined by the discrepancy, measured by the Euclidean distance, between the data points in the target image, and the corresponding points transferred by the sought after warp from the source image:

$$\mathcal{T}(\mathcal{W}) \stackrel{\text{def}}{=} \sum_{j=1}^m d^2(\mathcal{W}(\mathbf{q}_j), \mathbf{q}'_j).$$

This gives the RMSR as  $\sqrt{\frac{1}{m}\mathcal{T}(\mathcal{W})}$ . In other words, the estimation methods we give do Empirical Risk Minimization (ERM), assuming that the source data points are the inputs and the target data points the

outputs of some learning problem. Assuming that the target data points are subject to an *i.i.d.* gaussian noise, the estimation methods can also be seen as Maximum Likelihood Estimators (MLE). We assume that the number of data points is larger than the number of centres:  $m \gg l$ . In practice, as we mentioned in §3.2, the minimal number of centres is  $l \geq 3$  for all warps. For most warps, the transfer error is Nonlinear Least Squares (NLS). Two-step methods and algebraic approximations are used to get an initial estimate through Linear Least Squares minimization (LLS), solved using the pseudo-inverse technique or Singular Value Decomposition (SVD), if the system is homogeneous, enforcing unit two-norm on the unknown vector, see for instance (Hartley and Zisserman, 2003, A5). We normalize the coordinates of the target points so that they lie in the interval  $-\sqrt{2} \dots \sqrt{2}$  and their mean is at the origin – this has been shown to bring the algebraic error close to the Euclidean one (Chojnacki et al., 2003). The initial estimate is refined by iteratively minimizing the NLS transfer error with the Levenberg-Marquardt algorithm, see *e.g.*, (Hartley and Zisserman, 2003, A6) and Sequential Quadratic Programming to deal with the possible constraint.

The estimation procedures we propose below for the rigid cases, that is to say for the RA-Warp and the RP-Warp, follow the classical Structure-from-Motion paradigm of first reconstructing the cameras (*i.e.*, the ‘motion’) and then the structure, see for instance (Hartley and Zisserman, 2003, II). The structure is usually computed as a set of sparse features by triangulation. In the warp case, structure estimation takes the form of estimating the warp coefficients  $\delta$ , giving the inverse projective depth of the centres, and therefore the sought 3D structure. As in usual Structure-from-Motion, this initial estimation step could be followed by the joint refinement of both structure and motion and by bundle adjustment.<sup>10</sup>

## 7.2 The DA-Warp

We minimize the transfer error over the coordinates of the centres in the target image in  $P'$  by solving:

$$\min_{P'} \sum_{j=1}^m d^2(\mathcal{W}_{\text{DA}}(\mathbf{q}_j; P', \lambda), \mathbf{q}'_j).$$

Writing  $\ell_{\mathbf{q}_j}$  as  $\ell_j$ , and replacing  $\mathcal{W}_{\text{DA}}$  by its expression (15), we get:

$$\text{(LLS)} \quad \min_{P'} \sum_{j=1}^m \left\| \ell_j^\top \mathcal{E}_\lambda P' - \mathbf{q}'_j^\top \right\|^2.$$

The minimal number of point correspondences is  $m \geq 3$ .

<sup>10</sup>Applying bundle adjustment to warps is not trivial since it requires one to define a ‘generative’ model, whereby the image points are predicted from some unknown structure that is to be estimated, and naturally generalizes to more than two images. Defining bundle adjustment for image warps thus falls out of the scope of this paper.

### 7.3 The RA-Warp

Our algorithm first estimates the affine fundamental matrix  $\mathcal{A}$  and then the warp coefficients  $\delta$ . The minimal number of point correspondences is  $m \geq 4$ .

1. **Estimation of the affine fundamental matrix  $\mathcal{A}$ .** We estimate the affine fundamental matrix by minimizing the point-to-line transfer error *i.e.*, the sum of squared distances between the epipolar lines and the data points in the target image:

$$\min_{\mathcal{A}} \sum_{j=1}^m d_{\perp}^2(\mathbf{q}'_j, \mathcal{A}\tilde{\mathbf{q}}_j).$$

Expanding the point-to-line orthogonal distance  $d_{\perp}$  and replacing  $\mathcal{A}$  by its expression (11), we get:

$$\min_{\mathbf{z}, \mathcal{J}} \frac{1}{\|\mathbf{z}\|^2} \sum_{j=1}^m \left( \mathcal{J}^{\top} \tilde{\mathbf{q}}_j + \mathbf{z}^{\top} \mathbf{q}'_j \right)^2.$$

The difficulty of this minimization problem stems from the leading factor. In order to get rid of it, we fix the arbitrary scale of the fundamental matrix using  $\|\mathbf{z}\|^2 = 1$ . This makes sense since  $\|\mathbf{z}\|^2$  cannot vanish since otherwise  $\mathcal{A}$  would become rank one. The problem turns into a constrained homogeneous LLS minimization:

$$\min_{\mathbf{z}, \mathcal{J}} \sum_{j=1}^m \left( \mathcal{J}^{\top} \tilde{\mathbf{q}}_j + \mathbf{z}^{\top} \mathbf{q}'_j \right)^2 \quad \text{s.t.} \quad \|\mathbf{z}\|^2 = 1.$$

We rewrite the problem in matrix form:  $\min_{\mathbf{z}, \mathcal{J}} \|\mathcal{I}\mathbf{z} + \mathcal{J}\mathcal{J}\|^2$  such that  $\|\mathbf{z}\|^2 = 1$ , where the rows of  $\mathcal{I}$  are  $\mathbf{q}'_j{}^{\top}$  and the rows of  $\mathcal{J}$  are  $\tilde{\mathbf{q}}_j{}^{\top}$ . Setting  $\mathcal{J} = -\mathcal{J}^{\dagger}\mathcal{I}\mathbf{z}$ , substituting in the previous equation and factorizing gives:

$$\text{(LLS)} \quad \min_{\mathbf{z}} \left\| \left( \mathbf{I} - \mathcal{J}\mathcal{J}^{\dagger} \right) \mathcal{I}\mathbf{z} \right\|^2 \quad \text{s.t.} \quad \|\mathbf{z}\|^2 = 1,$$

that we solve using the standard method based on the SVD. Note that this algorithm requires at least  $m \geq 4$  point correspondences. From the affine fundamental matrix, we extract the target projection matrix  $\mathcal{S}_{\mathcal{A}}$  as indicated in §3.3.

2. **Estimation of the warp coefficients  $\delta$ .** We minimize the transfer error by solving:

$$\text{(LLS)} \quad \min_{\delta} \sum_{j=1}^m \left\| \mathbf{s}_{\mathcal{A}}^{\top} \ell_j^{\top} \mathcal{E}_{\lambda}^{\top} \delta + \ell_j^{\top} \mathcal{E}_{\lambda} \tilde{\mathbf{P}} \bar{\mathcal{S}}_{\mathcal{A}}^{\top} - \mathbf{q}'_j{}^{\top} \right\|^2.$$

The minimal number of point correspondences is  $m \geq 4$ , due to the first step of the algorithm (the second step requires  $m \geq 3$  point correspondences only).

## 7.4 The RP-Warp

Similarly to the algorithm for the RA-Warp we use a two-step procedure:

1. **Estimation of the fundamental matrix  $\mathcal{F}$ .** We estimate  $\mathcal{F}$  by minimizing the point-to-line transfer error in the target image. This NLS problem is initialized by the solution obtained by the 8 point algorithm and is stated as:

$$(NLS) \quad \min_{\mathcal{F}} \sum_{j=1}^m d_{\perp}^2(\mathbf{q}'_j, \mathcal{F}\tilde{\mathbf{q}}_j).$$

From the fundamental matrix, we extract the target projection matrix  $\mathcal{G}_{\mathcal{F}}$  as indicated in §3.3.

2. **Estimation of the warp parameters  $\delta$ .** We minimize the following algebraic approximation to the transfer error:

$$\min_{\delta} \sum_{j=1}^m d_a^2(\check{\mathcal{W}}_{RP}(\mathbf{q}_j; \delta, \mathcal{F}, \lambda), \mathbf{q}'_j),$$

with  $d_a^2(\check{\mathbf{q}}, \mathbf{q}') = \|\mathbf{S}[\check{\mathbf{q}}]_{\times} \tilde{\mathbf{q}}'\|^2$  an algebraic distance between points  $\mathbf{q}$  and  $\mathbf{q}'$ , and  $\mathbf{S} = (\mathbf{I} \ \mathbf{0})$  simply selects the two first rows of the cross-product. The algebraic approximation yields an LLS minimization problem since the algebraic distance directly compares the homogeneous coordinate vectors, thereby avoiding the need for the division required by the perspective warp. Substituting  $d_a$  by its expression, and the RP-Warp by its homogeneous formulation (23), we get:

$$\min_{\delta} \sum_{j=1}^m \left\| \mathbf{S}[\tilde{\mathbf{q}}'_j]_{\times} \check{\mathcal{G}}_{\mathcal{F}}(\tilde{\mathbf{P}} \ \delta)^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j \right\|^2,$$

and as sought, after minor algebraic manipulations, we get:

$$(LLS) \quad \min_{\delta} \sum_{j=1}^m \left\| \mathbf{S}[\tilde{\mathbf{q}}'_j]_{\times} \check{\mathbf{g}}_{\mathcal{F}} \ell_j^{\top} \mathcal{E}_{\lambda} \delta + \mathbf{S}[\tilde{\mathbf{q}}'_j]_{\times} \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j \right\|^2.$$

The solution provides an initialization to minimize the transfer error:

$$(NLS) \quad \min_{\delta} \sum_{j=1}^m d^2(\mathcal{W}_{RP}(\mathbf{q}_j; \delta, \mathcal{F}, \lambda), \mathbf{q}'_j),$$

The minimal number of point correspondences is  $m \geq 7$ , due to the first step of the algorithm that estimates the fundamental matrix (the second step requires  $m \geq 3$  point correspondences only).

## 7.5 The DP-Warp

As for the RP-Warp, we minimize an algebraic approximation to the transfer error, over the homogeneous coordinates  $\check{\mathbf{P}}'$  of the centres in the target image:

$$\min_{\check{\mathbf{P}}'} \sum_{j=1}^m d_a^2 (\check{\mathcal{W}}_{\text{DP}} (\mathbf{q}_j; \check{\mathbf{P}}', \lambda), \mathbf{q}'_j) \quad \text{s.t.} \quad \|\check{\mathbf{P}}'\| = 1.$$

The scale of  $\check{\mathbf{P}}'$  is fixed by enforcing its norm to unity. Replacing  $d_a$  by its expression, and  $\check{\mathcal{W}}_{\text{DP}}$  by its homogeneous expression (25), we obtain, after some minor algebraic manipulations:

$$\text{(LLS)} \quad \min_{\check{\mathbf{P}}'} \sum_{j=1}^m \left\| \mathbf{S} [\check{\mathbf{q}}'_j]_{\times} \text{diag}_3 (\boldsymbol{\ell}_j^{\top} \boldsymbol{\mathcal{E}}_{\lambda}) \text{vect} (\check{\mathbf{P}}') \right\|^2 \quad \text{s.t.} \quad \|\check{\mathbf{P}}'\| = 1,$$

with  $\text{diag}_r (\mathbf{x})$  an  $r$  block diagonal matrix with  $\mathbf{x}$  the repeated block, and with  $\text{vect}$  row-wise matrix vectorization. The minimizer is used to initialize the iterative minimization of the transfer error:

$$\text{(NLS)} \quad \min_{\check{\mathbf{P}}'} \sum_{j=1}^m d^2 (\mathcal{W}_{\text{DP}} (\mathbf{q}_j; \check{\mathbf{P}}', \lambda), \mathbf{q}'_j) \quad \text{s.t.} \quad \|\check{\mathbf{P}}'\| = 1.$$

The minimal number of point correspondences is  $m \geq 4$ .

## 8 Experimental Evaluation

We report experimental results on simulated and real data. For all the results, the  $l$  centres were randomly drawn well-spread within the area of the source image covered by the data points.<sup>11</sup> The same centres are used for all the compared warps. In the legends, ‘EGA’ and ‘EGP’ respectively stand for the affine and perspective epipolar geometry.

### 8.1 Simulated Data

**Simulation setup.** We simulate a camera observing a smooth surface which may undergo deformations between the two snapshots. The surface is generated by linearly interpolating points on a plane and on a cylinder, creating a varying amount of surface bending. Points on the surface are projected to the images and corrupted by an additive Gaussian noise. The default simulation parameters are  $m = 100$  points,  $\sigma = 1$  pixel noise variance, a focal length of  $f = 700$  pixels and a scene to camera distance of 2.5. Together, the focal length and scene to camera distance are tuned to increase or decrease the perspective while keeping

<sup>11</sup>Other options are to choose the centres from a regular grid, as image keypoints or as data points. This does not make a noticeable difference in practice.

constant the size of the projected scene. The default parameters yield mild perspective effects. The default surface is slightly bent (half-way between the plane and the cylinder). The results we report are averages over 300 randomly drawn sets of centres, and several numbers of centres  $l \in \{3, 9, 30\}$  were tested. What will be observed is that, as expected, the larger the number of centres, the lower the transfer error for the four warps depending on the centres (the DA-Warp, the RA-Warp, the RP-Warp and the DP-Warp). Note that the rigid and deformable warps are respectively estimated from rigid and deformable datasets. A deformable dataset is obtained by choosing a different amount of surface bending between the two snapshots. Note that only the DA-Warp and the DP-Warp are estimated from deformable datasets. This is why for  $l = 3$  centres the RMSR for the DA-Warp and the FA-Warp do not match, as well as for the DP-Warp and the RP-Warp.

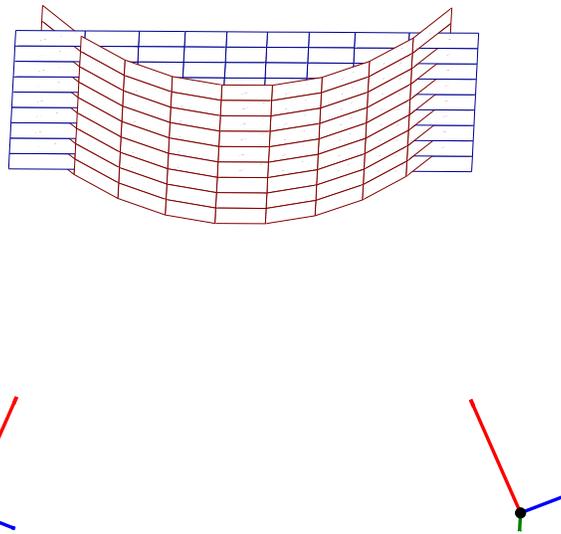


Figure 14: We simulate two cameras observing a smooth surface. This surface is created by linearly interpolating points on a plane and part of a cylinder. The interpolation weight thus controls the amount of surface bending.

**Influence of noise.** We evaluate the influence of the noise level on the estimated warps. Figure 15 shows the results we obtained for low and high amounts of noise. High amounts of noise are meant to show how the warps behave when the optic flow field gets less smooth.

The perspective epipolar geometry is the only ‘physical’ model for the simulated data. Its transfer error gracefully degrades as the noise increases. The affine epipolar geometry, the FA-Warp and the FP-Warp do not match the simulated data and have quite high transfer errors. They are not influenced by the added noise of low magnitude and mildly by higher magnitudes. As the number of centres increases, the DA-Warp, the DP-Warp and the RP-Warp model the data better and better, and get consequently more and more influenced by the added noise, even though they are sensitive to high amounts of added noise even in the few centres case. The transfer error for the RA-Warp nearly reaches its lower bound, given by the affine

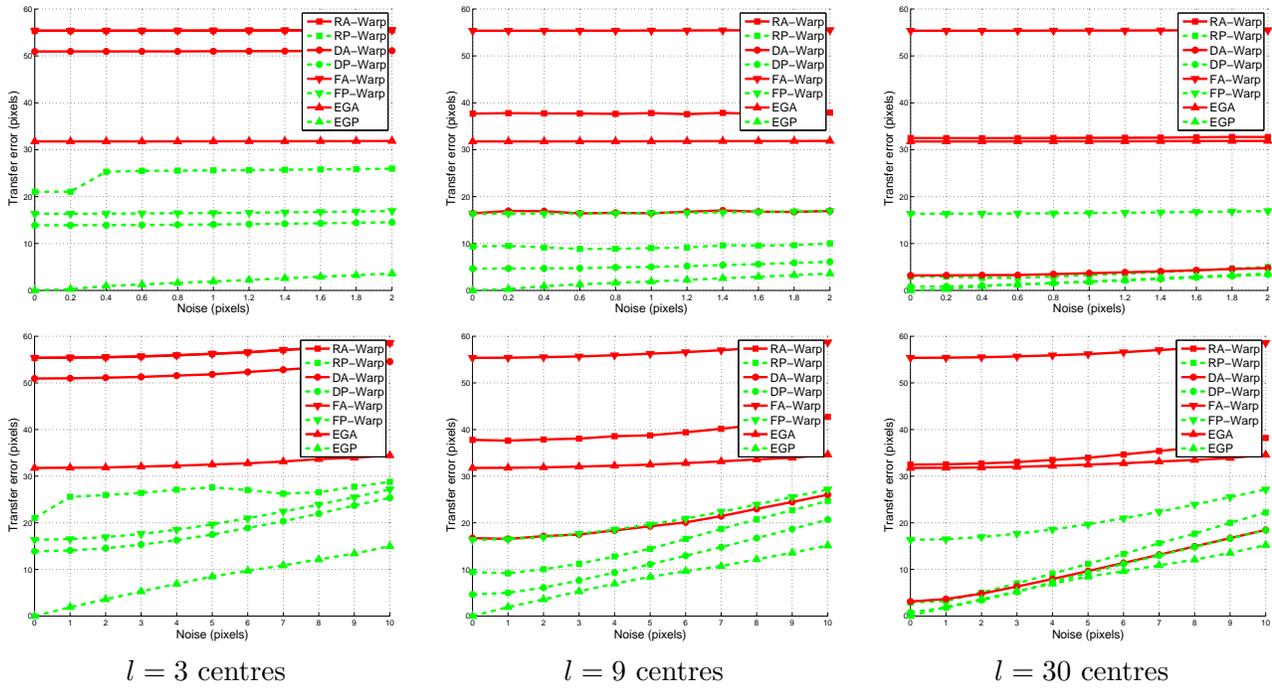


Figure 15: The influence on the estimated warps of the amount of noise added in the projected image points. Low and high amounts of noise are shown on the top and bottom graphs respectively.

epipolar geometry, and is slightly influenced by the low magnitude added noise for  $l = 30$  centres and clearly influenced by the high magnitude added noise whatever the number of centres.

To conclude, we can say that the added noise of a low magnitude significantly influences only those warps that have a relatively ‘low’ transfer error (the order of which being a few pixels) *i.e.*, that really models the data. The added noise of higher magnitude however has a clear influence on all the warps, for which the transfer error gracefully degrades as the noise increases.

**Influence of the amount of perspective.** We evaluate the influence of the amount of perspective on the estimated warps. Figure 16 shows the results we obtained.

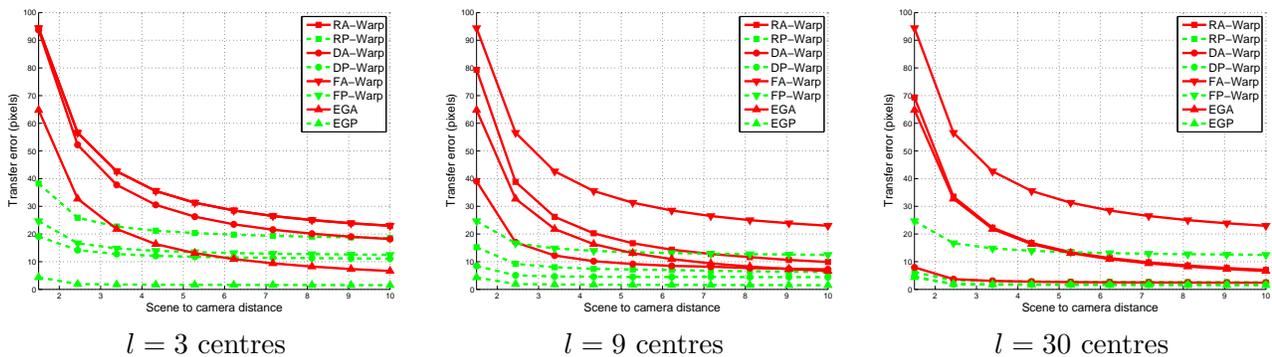


Figure 16: The influence on the estimated warps of the amount of perspective of the camera with respect to the observed surface.

As expected, the transfer error for the affine warps decreases with the amount of perspective. The perspective warps are much less influenced. As already observed for the real poster data in §4.1.4, the DA-Warp models the data only when the number of centres grows large enough. It thus becomes quite independent of the amount of perspective for  $l = 30$  centres. These results confirm that the DA-Warp and RA-Warp are intrinsically affine.

**Influence of the amount of surface bending.** We evaluate the influence of the amount of bending on the estimated warps. Figure 16 shows the results we obtained.

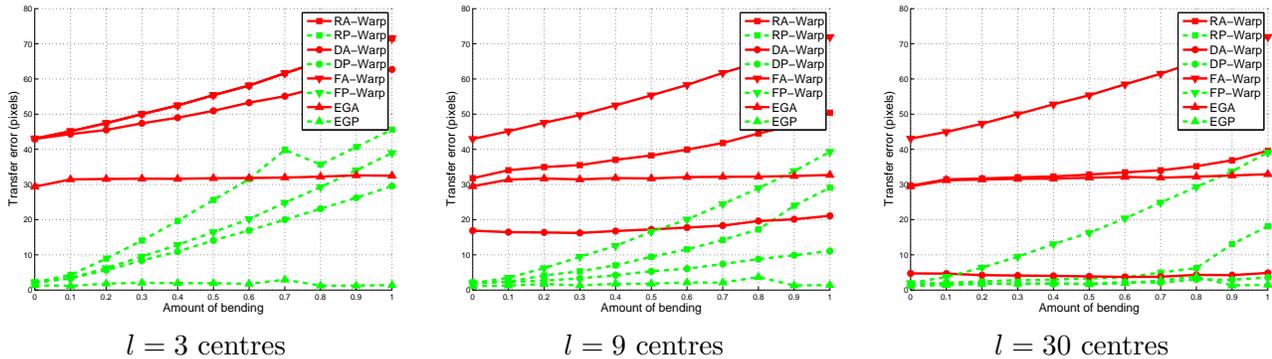


Figure 17: The influence of the amount of bending of the observed surface on the estimated warps.

The epipolar geometry, whether affine or perspective, is not influenced by the amount of bending. All the other warps, whether rigid or deformable, degrade as the bending increases. While this is obvious for the flat warps, it shows that increasing the amount of bending, thereby making more complex the structure of the image deformation relating the two images, requires more flexibility for the warp to model the data. This is confirmed by the fact that when the number of centres grows to  $l = 30$ , the deformable warps are not influenced by the amount of bending, while the rigid RA-Warp and RP-Warp are much less influenced than for lower numbers of centres.

## 8.2 Real Data

We report experimental results complementary to those already shown to illustrate the different warps through the paper. Three pairs of images were used ; the same are considered below, as well as a fourth one to compare the DA-Warp and the DP-Warp in the presence of perspective and deformation.

**Deformable warps and perspective.** For the pair of images shown in figure 18 (a) and (b), a low fitting transfer error RMSR of 1.39 pixels shows that the relationship is well modeled by an FP-Warp. Figure 18 (c) confirms our previous observations: the DP-Warp models the images very well, even with the minimum

$l = 3$  centres. The DA-Warp however requires much more centres to model the perspective effect as a deformation.

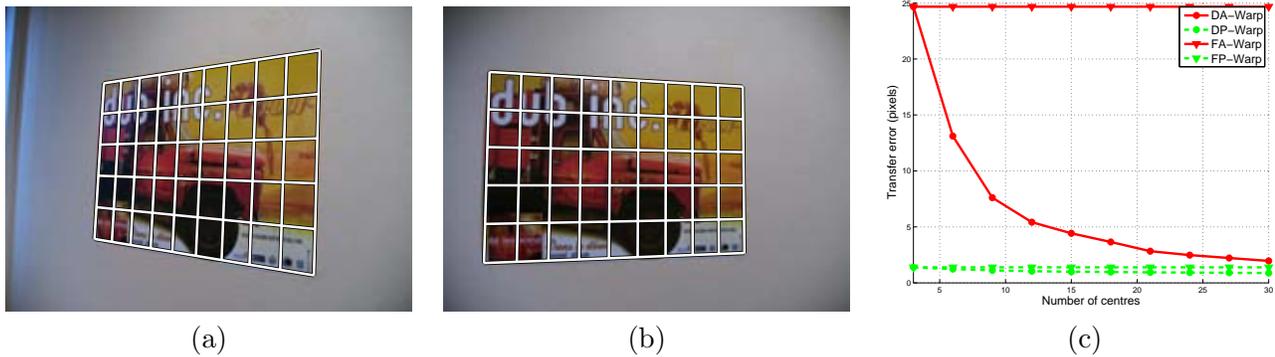


Figure 18: Deformable warps (the DA-Warp and the DP-Warp) and perspective. (a) and (b) show the source and target images, overlaid with the warp visualization grid, transferred from the source to the target with an FP-Warp. (c) shows the fitting transfer error RMSR as a function of the number of centres. This figure matches figures 5 and 12.

Figure 19 shows a comparison between the DA-Warp and the DP-Warp for the pair of images DP11 and DP22. In those images, the perspective effect is quite important, and the surface deforms between the source and the target image. The results we obtained confirmed what we observed in the case of simulated data: with many centres, the flexibility of the DA-Warp models the effect of perspective. The DA-Warp thus fits the images quite well: the RMSR for both the DA-Warp and the DP-Warp reaches 4.17 pixels for  $l = 30$  centres. For fewer centres however, there is a significant difference between the DA-Warp and the DP-Warp. For  $l = 9$  centres, the former has an RMSR of 14.28 pixels, while the latter has an RMSR of 6.10 pixels. As we observed with simulated data, the difference between the DA-Warp and the DP-Warp, even though very significant, is less important for a bent surface than for a flat surface.

**Affine warps and rigidity.** Figure 20 (a) and (b) shows a pair of images for which we get a fitting transfer error RMSR of 3.71 pixels for the affine epipolar geometry. Figure 20 (c) confirms that the error RMSR for the RA-Warp converges to the affine epipolar geometry when the number of centres grows. The DA-Warp remains quite close to the RA-Warp but has its extra degrees of freedom used to model the noise, and thus eventually reaches a lower fitting transfer error.

**Rigid warps and perspective.** Figure 21 (a) and (b) shows a pair of images for which we get a fitting transfer error RMSR of 13.77 for the affine epipolar geometry and 0.57 pixels for the perspective epipolar geometry. Figure 21 (c) confirms that the error RMSR for the RP-Warp converges to the perspective epipolar geometry when the number of centres grows, and really uses the centres to model the deformation. The RA-Warp is ‘stuck’ with a higher error since it cannot outperform the affine epipolar geometry.

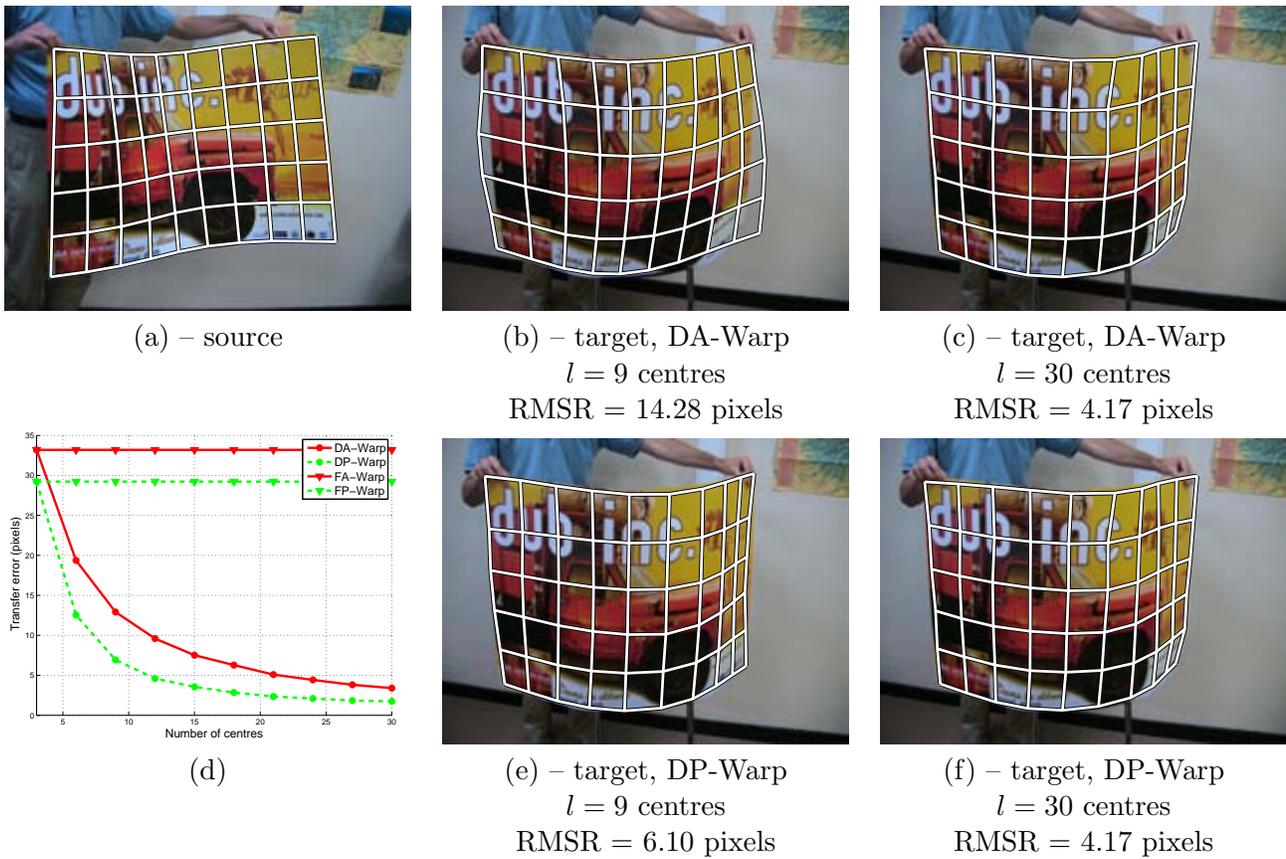


Figure 19: Deformable warps (the DA-Warp and the DP-Warp) and perspective. (a) and (b,c,e,f) show the source and target images, overlaid with the warp visualization grid, transferred from the source to the target with various warps and numbers of centres. (d) shows the fitting transfer error RMSR as a function of the number of centres.

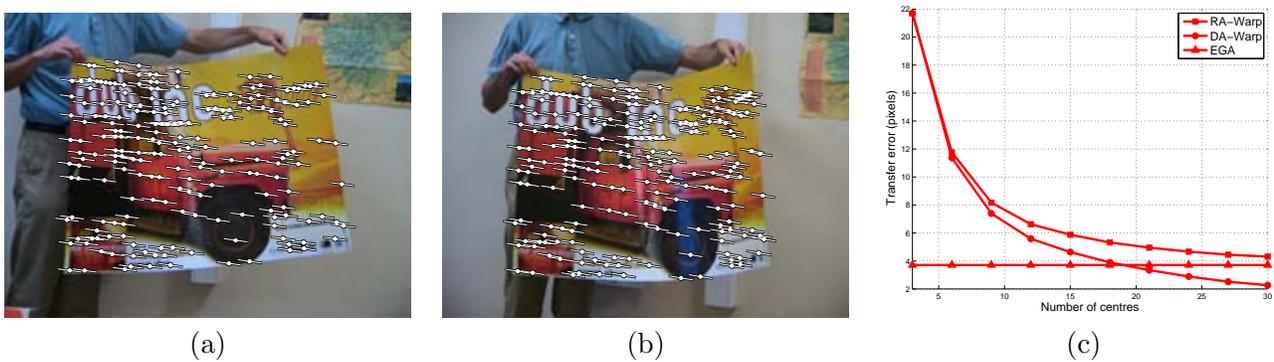


Figure 20: Affine warps (the DA-Warp and the RA-Warp) and rigidity. (a) and (b) show the source and target images, overlaid with epipolar lines for the affine epipolar geometry. (c) shows the fitting transfer error RMSR as a function of the number of centres. This figure matches figures 6 and 8.

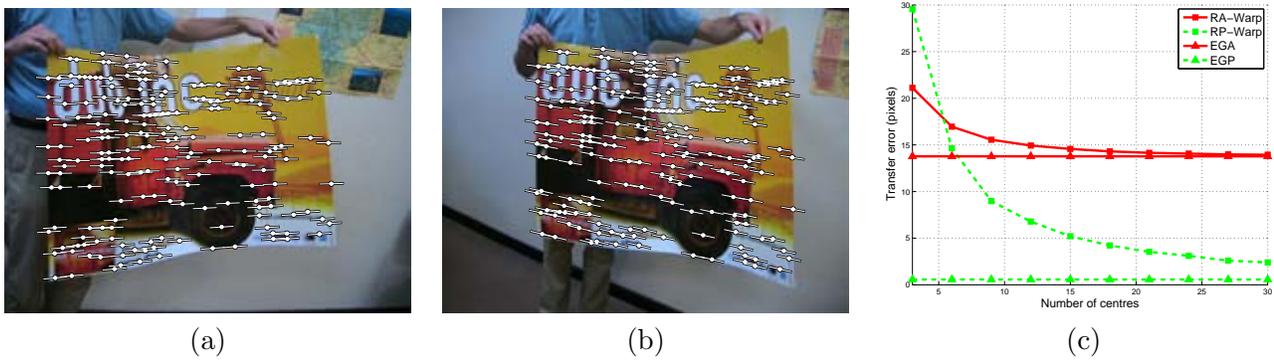


Figure 21: Rigid warps (the RA-Warp and the RP-Warp) and perspective. (a) and (b) show the source and target images, overlaid with epipolar lines for the perspective epipolar geometry. (c) shows the fitting transfer error RMSR as a function of the number of centres. This figure matches figures 9 and 10.

**Rigid warps and 3D surfaces.** The rigid warps, namely the RA-Warp and the RP-Warp, have been shown to define a 3D surface. Figure 22 shows an example of such a surface recovered from an RP-Warp. A pair of perspective images, DP21 and DP22, were used to estimate an RP-Warp. Given the epipolar geometry, and knowing the internal camera calibration parameters from the header of the image files,<sup>12</sup> we formed the calibrated projection matrices. This allowed us to plot the cameras in 3D with the images textured-mapped on them. We then used the TPS function from the RP-Warp (22) to find the depth of the data points and of the points forming the visualization grid. This 3D visualization confirms the well-founded derivation of our rigid warps based on a single TPS function giving the depth.

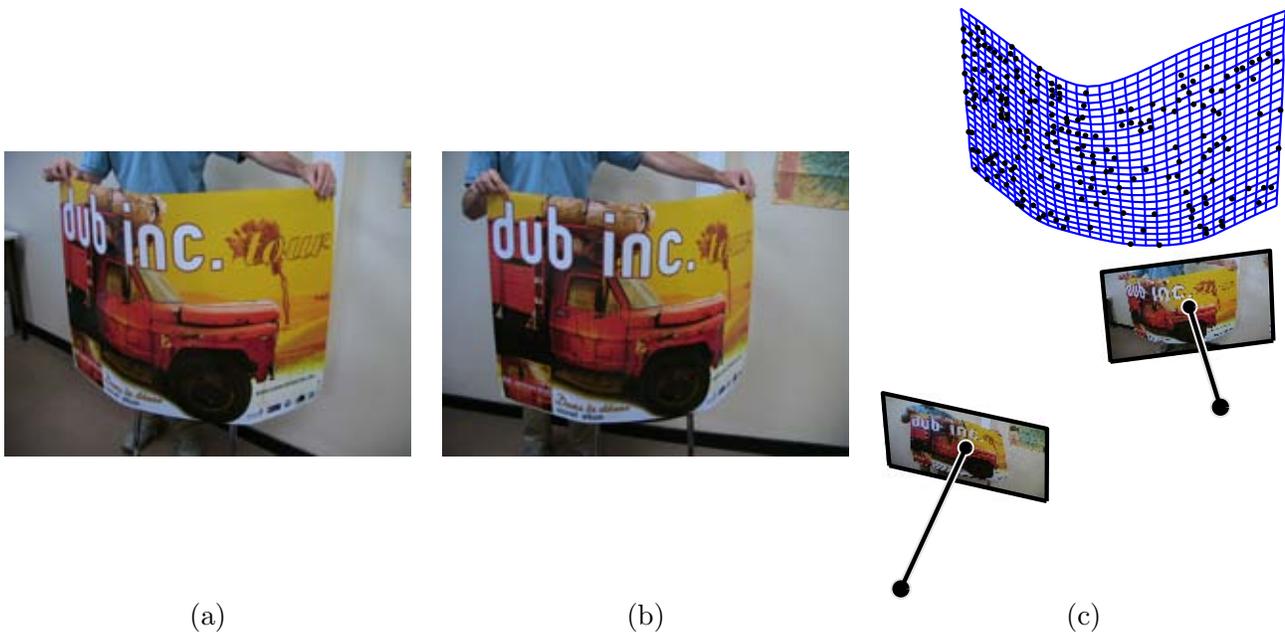


Figure 22: Extracting 3D entities from the RP-Warp. (c) shows the cameras and 3D surface extracted from an RP-Warp estimated between two perspective images (a) and (b) of a rigid surface.

<sup>12</sup>The Exif format gives information on the state of the camera which took the picture, such as its focal length.

## 9 Discussion

Three types of  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ , image warps were proposed. They use the  $\mathbb{R}^2 \rightarrow \mathbb{R}$  Thin-Plate Spline as a building block. They were designed to overcome some limitations of the standard Thin-Plate Spline warp, herein dubbed ‘DA-Warp’ for ‘Deformable Affine’. Their derivation is based on a feature-driven parameterization that we introduced. We formulated a rigid DA-Warp that we called the ‘RA-Warp’ for ‘Rigid Affine’. Several arguments were given to show that both the DA-Warp and the RA-Warp are intrinsically affine, in the sense of the affine camera model. They make linear combinations of the centres in affine coordinates. We derived the perspective projection model analogue of these two warps, respectively called the ‘DP-Warp’ for ‘Deformable Perspective’ and the ‘RP-Warp’ for ‘Rigid Perspective’. They were shown to make linear combinations of the centres in homogeneous coordinates. A geometric interpretation of the warps in terms of projected surfaces was reported. The differences between them lie in the choice of a rigid or of a deformable surface and in the affine or the perspective projection model. These warps have a direct practical impact since instead of using the standard DA-Warp, a better adapted warp can be chosen from the proposed ones for rigid smooth surfaces and images with perspective projection effects. The ‘generalization’ principle that we proposed can also be applied to other types of warps. For instance, a follow-up of our work is the NURBS-Warp (Brunet et al., 2009), the perspective version of the usual tensor-product cubic B-Spline warp.

This paper opens several research possibilities, including:

- **Estimation of the warps.** Various estimation methods were proposed for deformable warps in the literature. In this paper, we showed how (asymmetric) Maximum Likelihood Estimation (MLE) from point correspondences could be performed. The first extension that can be thought of is the one of performing (asymmetric) Maximum A Posteriori (MAP) to include a smoothness prior. This raises the problem of selecting the smoothing weight properly (Bartoli, 2008). Alternatively, the number of centres can be adjusted, as in (Bartoli, 2009). A second possible extension is to make the estimation procedure robust *i.e.*, resistant to erroneous point matches, as is done in (Pilet et al., 2008) for piecewise affine warps. Finally, it is also possible to estimate warps using a direct method, with the pixel color discrepancy as an error function, as proposed in (Bartoli and Zisserman, 2004) for the DA-Warp. Automatic centre placement might also be considered as an open problem.
- **Combined rigid and deformable warps.** Estimating deformable warps with parts constrained to move rigidly is important in medical image registration, for instance in the case of CT and PET images of thoracic and abdominal regions (Moreno et al., 2006). Combined rigid and deformable warps can

be obtained by merging a DA-Warp and several RA-Warps, one for each rigid part.

- **Multiple view warps.** The warps we derived are image to image. For the rigid surface setting at least, it is possible to introduce other images, with warps sharing strong properties with the initial one, namely the centre depth vector  $\delta$ . This opens the issues of occlusion reasoning and the estimation of a warp consistently over multiple views.
- **Automatic surface reconstruction.** For all the warps derived in this paper, there exist an underlying 3D surface. While it is obvious that it matches some ‘true’ surface in the rigid case, it generally does not in the deformable case. Indeed, it is straightforward from the geometric interpretation of the DA-Warp given in §4.1.3 to see that the depth of the ‘3D centres’ that govern the observed 3D surface cannot be recovered from the target image centres (we have only two constraints in the equation  $P' = \tilde{Z}MS_A^T$  for the three unknowns of each ‘3D centres’: their depth is left unconstrained). Enforcing surface smoothness does not raise the ambiguity since it always leads to a flat surface. In the perspective case, however, the homogeneous ‘scale’ of the target image centres is related to their depth, as shown in §5.2.1. It is however expected to give quite noisy and unstable depth estimates since it is based on the perspective effect only and ‘degenerates’ in the affine case. One possible improvement would thus be to include strong 3D surface priors, such as pre-trained linear shape models (Salzmann et al., 2007).

## References

- A. Bartoli. Maximizing the predictivity of smooth deformable image warps through cross-validation. *Journal of Mathematical Imaging and Vision*, 31(2-3):133–145, July 2008.
- A. Bartoli. On computing the prediction sum of squares statistic in linear least squares problems with multiple parameter or measurement sets. *International Journal of Computer Vision*, 85(2):133–142, November 2009.
- A. Bartoli and A. Zisserman. Direct estimation of non-rigid registrations. In *British Machine Vision Conference*, 2004.
- V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), September 2003.
- F. L. Bookstein. Principal warps: Thin-Plate Splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.

- C. Bregler, A. Hertzmann, and H. Biermann. Recovering non-rigid 3D shape from image streams. In *International Conference on Computer Vision and Pattern Recognition*, 2000.
- F. Brunet, A. Bartoli, R. Malgouyres, and N. Navab. NURBS warps. In *British Machine Vision Conference*, 2009.
- W. Chojnacki, M. J. Brooks, A. van den Hengel, and D. Gawley. Revisiting Hartley’s normalised eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1172–1177, 2003.
- H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141, February 2003.
- T. F. Cootes, S. Marsland, C. J. Twining, K. Smith, and C. J. Taylor. Groupwise diffeomorphic non-rigid registration for automatic model building. In *European Conference on Computer Vision*, 2004.
- G. Donato and S. Belongie. Approximate thin-plate spline mappings. In *European Conference on Computer Vision*, 2002.
- J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *RAIRO Analyse Numérique*, 10:5–12, 1976.
- V. Gay-Bellile, A. Bartoli, and P. Sayd. Direct estimation of non-rigid registrations with image-based self-occlusion reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009. To appear.
- R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003. Second Edition.
- M. Irani, P. Anandan, and M. Cohen. Direct recovery of planar-parallax from multiple frames. In *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, 1999.
- S. Y. Lee, K. Y. Chwa, J. Hahn, and S. Y. Shin. Image morphing using deformation techniques. *Journal of Visualization and Computer Animation*, 7:3–26, 1996.
- J. Lim and M.-H. Yang. A direct method for non-rigid motion with thin-plate spline. In *International Conference on Computer Vision and Pattern Recognition*, 2005.
- Q.-T. Luong and T. Vieville. Canonic representations for the geometries of multiple projective views. *Computer Vision and Image Understanding*, 64(2):193–229, 1996.
- L. Masson, M. Dhome, and F. Jurie. Tracking 3D objects using flexible models. In *British Machine Vision Conference*, 2005.

- A. Moreno, G. Delso, O. Camara, and I. Bloch. Non-linear registration between 3D images including rigid objects: Application to CT and PET lung images with tumors. In *Workshop on Image Registration in Deformable Environments (DEFORM) at BMVC*, 2006.
- M. Perriollat and A. Bartoli. A single directrix quasi-minimal model for paper-like surfaces. In *Workshop on Image Registration in Deformable Environments (DEFORM) at BMVC*, 2006.
- J. Pilet, V. Lepetit, and P. Fua. Fast non-rigid surface detection, registration and realistic augmentation. *International Journal of Computer Vision*, 76(2):109–122, February 2008.
- M. Prasad, A. Zisserman, and A. Fitzgibbon. Single view reconstruction of curved surfaces. In *International Conference on Computer Vision and Pattern Recognition*, 2006.
- M. Salzmann, J. Pilet, S. Ilic, and P. Fua. Surface deformation models for nonrigid 3D shape recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1–7, August 2007.
- S. Schaefer, T. McPhail, and J. Warren. Image deformation using moving least squares. In *SIGGRAPH*, 2006.
- T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. In *SIGGRAPH*, 1986.
- D. Terzopoulos. Multilevel computational processes for visual surface reconstruction. *Computer Vision, Graphics and Image Processing*, 24:52–96, 1983.
- J. Wills and S. Belongie. A feature-based approach for determining dense long range correspondences. In *European Conference on Computer Vision*, 2004.
- A. Y. Yang, S. Rao, A. Wagner, and Y. Ma. Segmentation of a piece-wise planar scene from perspective images. In *International Conference on Computer Vision and Pattern Recognition*, 2005.

## A Asymptotic Internal Regularization Behaviour

We derive the form that the proposed warps take when the internal regularization parameter tends to infinity.

### A.1 The DA-Warp

Considering the definition (15) of the DA-Warp, and the asymptotic TPS (9), it is straightforward to see that the asymptotic DA-Warp is an FA-Warp represented by the  $(2 \times 3)$  matrix  $\mathcal{L}_{\text{DA}} = \left(\tilde{\mathbf{P}}^\dagger \mathbf{P}'\right)^\top$ . We can easily see that this FA-Warp minimizes the transfer error, since  $\arg \min_{\mathcal{L}} \left\| \tilde{\mathbf{P}} \mathcal{L}^\top - \mathbf{P}' \right\|^2 = \mathcal{L}_{\text{DA}}$ .

## A.2 The RA-Warp and the RP-Warp

This proof has two parts. First, we show that the FA-Warp and the FP-Warp, resulting from the asymptotically regularized RA-Warp and RP-Warp are plane-induced rigid warps, and that the inducing plane has a reduced equation  $\boldsymbol{\pi} = \tilde{\mathbf{P}}^\dagger \boldsymbol{\delta} \in \mathbb{R}^3$ . Second, we show that  $\boldsymbol{\pi}$  is the reduced equation of the plane minimizing the depth error with respect to the source camera. An interpretation of  $\boldsymbol{\pi}$  as a minimizer of some transfer error is also given.

- **Part 1: Plane-induced flat warps and plane equation.** (Recall that an FA-Warp induced by a plane  $\boldsymbol{\pi}$  is represented by the  $(2 \times 3)$  matrix  $\mathcal{L}_{\text{RA}} = \mathcal{S}_{\mathcal{A}} + \mathbf{s}_{\mathcal{A}} \boldsymbol{\pi}^\top$  and an FP-Warp by the  $(3 \times 3)$  matrix  $\check{\mathcal{L}}_{\text{RP}} \sim \check{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^\top$ .) We give a proof for the perspective case and then specialize the result to the affine case. The asymptotic FP-Warp matrix (24) is rewritten as:

$$\check{\mathcal{L}}_{\text{RP}} \sim \left( \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^\top + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^\top \right) \tilde{\mathbf{P}}^\dagger{}^\top,$$

from which, distributing and using the property<sup>13</sup>  $\mathbf{X}^\top \mathbf{X}^\dagger{}^\top = \mathbf{I}$ , we get:

$$\check{\mathcal{L}}_{\text{RP}} \sim \bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^\top \quad \text{with} \quad \boldsymbol{\pi} = \tilde{\mathbf{P}}^\dagger \boldsymbol{\delta}.$$

Replacing  $\mathcal{F}$  by the affine fundamental matrix  $\mathcal{A}$ , we get:

$$\check{\mathcal{L}}_{\text{RA}} \sim \bar{\mathcal{G}}_{\mathcal{A}} + \check{\mathbf{g}}_{\mathcal{A}} \boldsymbol{\pi}^\top.$$

Observing that the last row of  $\bar{\mathcal{G}}_{\mathcal{A}}$  is  $(0 \ 0 \ 1)$  and that the last element of  $\check{\mathbf{g}}_{\mathcal{A}}$  vanishes, we get:

$$\mathcal{L}_{\text{RA}} = \bar{\mathcal{S}}_{\mathcal{A}} + \mathbf{s}_{\mathcal{A}} \boldsymbol{\pi}^\top.$$

- **Part 2:  $\boldsymbol{\pi}$  is the reduced equation of the depth error minimizing plane.** For a rigid warp, the depth error  $\epsilon_{3\text{D}}$  is the discrepancy between the depth  $\delta_j$  of the data points and the depth  $\tau(\mathbf{q}_j; \boldsymbol{\delta}, \lambda)$  predicted by the warp:

$$\epsilon_{3\text{D}}^2 \stackrel{\text{def}}{=} \sum_{j=1}^m (\tau(\mathbf{q}_j; \boldsymbol{\delta}, \lambda) - \delta_j)^2.$$

The depth of a 3D point reconstructed on the plane  $\boldsymbol{\zeta}$  is  $\boldsymbol{\zeta}^\top \tilde{\mathbf{q}}_j$ . The depth error is thus  $\epsilon_{3\text{D}}^2(\boldsymbol{\zeta}) = \sum_{j=1}^m \left\| \boldsymbol{\zeta}^\top \tilde{\mathbf{q}}_j - \hat{\delta}_j \right\|^2$ . It is straightforward to see that  $\arg \min_{\boldsymbol{\zeta}} \epsilon_{3\text{D}}^2(\boldsymbol{\zeta}) = \boldsymbol{\pi} = \tilde{\mathbf{P}}^\dagger \boldsymbol{\delta}$ . Therefore, the

<sup>13</sup>This property holds on a full column rank portrait matrix such as  $\tilde{\mathbf{P}}$ . It is straightforward to show that  $\mathbf{X}^\top \mathbf{X}^\dagger{}^\top = \mathbf{X}^\top \left( (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \right)^\top = \mathbf{X}^\top \mathbf{X} (\mathbf{X}^\top \mathbf{X})^{-\top} = (\mathbf{X}^\top \mathbf{X})^\top (\mathbf{X}^\top \mathbf{X})^{-\top} = \mathbf{I}$ .

asymptotic FA-Warp and FP-Warp are depth error minimizing, plane induced affinity and homography, respectively. Using the results in Appendix B, we conclude that  $\pi$  has an interpretation as the minimizer of some transfer error.

### A.3 The DP-Warp

Consider the following error function for an FP-Warp represented by the  $(3 \times 3)$  matrix  $\check{H}$  :

$$\tilde{\epsilon}_a(\check{H}, \mathbf{d}) = \sum_{j=1}^m \|d_j \tilde{\mathbf{q}}'_j - \check{H} \tilde{\mathbf{q}}_j\|^2.$$

This is an algebraic transfer error function, different from the one we use in §§7.4 and 7.5 to initialize the RP-Warp and the DP-Warp, based on the distance  $d_a$  derived from the norm of the vector product of homogeneous point coordinates. This error function can be rewritten in matrix form as:

$$\tilde{\epsilon}_a(\check{H}, \mathbf{d}) = \left\| \text{diag}(\mathbf{d}) \tilde{P}' - \tilde{P} \check{H}^\top \right\|^2,$$

from which it is straightforward to obtain the minimizing  $\check{H}$  as a function of  $\mathbf{d}$  by setting  $\frac{\partial \tilde{\epsilon}_a^2}{\partial \check{H}} = \mathbf{0}$ :

$$\check{H}^\top = \tilde{P}^\dagger \text{diag}(\mathbf{d}) \tilde{P}'.$$

which is the matrix representing the asymptotic FP-Warp for the DP-Warp.

Intuitively,  $\tilde{\epsilon}_a^2$  and the above discussed algebraic error based on  $d_a$  are different since the former compares all homogeneous coordinates while the one compares only the leading two ones, corresponding to the  $x$  and  $y$  coordinates, weighted by the homogeneous coordinate of the transferred point. More formally, setting  $d_j = (H_{31} \ H_{32} \ H_{33}) \tilde{\mathbf{q}}_j$  equalizes the two errors. Computing  $\mathbf{d}$  by solving:

$$\min_{\mathbf{d}} \tilde{\epsilon}_a^2(\check{H}, \mathbf{d}),$$

yields:

$$d_j = \frac{\tilde{\mathbf{q}}_j'^\top \check{H} \tilde{\mathbf{q}}_j}{\|\tilde{\mathbf{q}}_j'\|^2}.$$

This shows that the two algebraic error functions do not in general share the same minimizer. Interpreting the homogeneous point coordinates as 3D vectors, we note that  $\tilde{\epsilon}_a^2$  is actually the sum of orthogonal distances between the ‘direction vector’ of  $\tilde{\mathbf{q}}_j'$  and the transferred point  $\check{H} \tilde{\mathbf{q}}_j$ .

## B Relationship Between the Depth and the Transfer Errors

The rigid warps can have their  $\delta$  parameters estimated by minimizing the discrepancy  $\epsilon_{3D}$  between the depth of the triangulated data points and the predicted depth. A tight relationship is revealed between the 3D depth error and the transfer error. The reasoning is based on the assumption that the image points are ‘corrected’, *i.e.*, that they exactly satisfy the rigidity constraint. In the affine case, the two errors are equal up to a global scale, while in the perspective case, each term has a different weight.

### B.1 Affine Imaging Model

The transfer error for an RA-Warp is:

$$\epsilon_e^2 = \sum_{j=1}^m \left\| \mathcal{W}_{\text{RA}}(\mathbf{q}_j; \boldsymbol{\delta}, \mathcal{A}, \lambda) - \mathbf{q}'_j \right\|^2.$$

Substituting the expression (19) of the RA-Warp, and using the rigidity assumption through the transfer equation (14), we get:

$$\epsilon_e^2 = \sum_{j=1}^m \left\| \bar{\mathcal{S}}_{\mathcal{A}} \left( \tilde{\mathbf{P}}^\top \mathcal{E}_\lambda^\top \ell_j - \tilde{\mathbf{q}}_j \right) + \mathbf{s}_{\mathcal{A}} \left( \boldsymbol{\delta}^\top \mathcal{E}_\lambda^\top \ell_j - \hat{\delta}_j \right) \right\|^2,$$

that simplifies, using property (6), to:

$$\epsilon_e^2 = \sum_{j=1}^m \left\| \mathbf{s}_{\mathcal{A}} \boldsymbol{\zeta}^\top \boldsymbol{\delta}^\top \mathcal{E}_\lambda^\top \ell_j - \mathbf{s}_{\mathcal{A}} \hat{\delta}_j \right\|^2,$$

and factors as:

$$\epsilon_e^2 = \|\mathbf{s}_{\mathcal{A}}\|^2 \sum_{j=1}^m \left( \tau(\mathbf{q}_j; \boldsymbol{\delta}, \lambda) - \hat{\delta}_j \right)^2.$$

We thus have  $\epsilon_e = \|\mathbf{s}_{\mathcal{A}}\| \epsilon_{3D} \propto \epsilon_{3D}$ , the 3D depth error.

### B.2 Perspective Imaging Model

The algebraic transfer error for an RP-Warp is:

$$\epsilon_a^2 = \sum_{j=1}^m \left\| \mathbf{S} [\tilde{\mathbf{q}}'_j]_{\times} \mathcal{W}_{\text{RP}}(\mathbf{q}_j; \boldsymbol{\delta}, \mathcal{F}, \lambda) \right\|^2.$$

Substituting the expression (23) of the RP-Warp, and using the rigidity assumption through the transfer equation (13), we get:

$$\epsilon_a^2 = \sum_{j=1}^m \left\| S \left[ \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}}_j + \check{\mathbf{g}}_{\mathcal{F}} \hat{\delta}_j \right]_{\times} \left( \bar{\mathcal{G}}_{\mathcal{F}} \bar{\mathbf{P}}^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j \right) \right\|^2.$$

Using property (6), expanding and noting that two cross-products vanish gives:

$$\epsilon_a^2 = \sum_{j=1}^m \left\| S \left[ \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}}_j \right]_{\times} \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j + S \left[ \check{\mathbf{g}}_{\mathcal{F}} \right]_{\times} \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}}_j \hat{\delta}_j \right\|^2.$$

Reordering the two factors in the leading cross-product, we factor the error as:

$$\epsilon_a^2 = \sum_{j=1}^m \left\| S \left[ \check{\mathbf{g}}_{\mathcal{F}} \right]_{\times} \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}}_j \right\|^2 \left( \boldsymbol{\delta}^{\top} \mathcal{E}_{\lambda}^{\top} \ell_j - \hat{\delta}_j \right)^2.$$

Noting that  $[\check{\mathbf{g}}_{\mathcal{F}}]_{\times} \bar{\mathcal{G}}_{\mathcal{F}} \sim \mathcal{F}$ , the fundamental matrix, and identifying a TPS (4), we get:

$$\epsilon_a^2 = \sum_{j=1}^m \left\| S \mathcal{F} \tilde{\mathbf{q}}_j \right\|^2 \left( \tau(\mathbf{q}_j; \boldsymbol{\delta}, \lambda) - \hat{\delta}_j \right)^2.$$

The algebraic transfer error is thus similar to the 3D depth error with each term reweighted by the norm of the direction vector of the corresponding epipolar line.

## C Some Relationships Between the Sets of Warps

### C.1 $\mathbb{S}_{\text{FA}} \subset \mathbb{S}_{\text{RA}}$ and $\mathbb{S}_{\text{FP}} \subset \mathbb{S}_{\text{RP}}$

We prove the perspective case. Specializing the result to the affine case is straightforward. Let  $\check{\mathbf{H}}$  be a  $(3 \times 3)$  homography matrix representing an element of  $\mathbb{S}_{\text{FP}}$  *i.e.*, an FP-Warp. We have to find the parameters  $\boldsymbol{\delta}$  of an RP-Warp (23) such that:

$$\check{\mathbf{H}} \tilde{\mathbf{q}} \sim \left( \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^{\top} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\delta}^{\top} \right) \mathcal{E}_{\lambda}^{\top} \ell_{\mathbf{q}}.$$

Using property (6), the first term reduces to  $\bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{q}}$ . Choosing  $\boldsymbol{\delta}$  such that  $\boldsymbol{\delta}^{\top} \mathcal{E}_{\lambda}^{\top} = (\mathbf{0} \ \boldsymbol{\pi})$ , the second term reduces to  $\check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^{\top} \tilde{\mathbf{q}}$ , giving:

$$\check{\mathbf{H}} \tilde{\mathbf{q}} \sim \left( \bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^{\top} \right) \tilde{\mathbf{q}},$$

which is the general form of a plane-induced FP-Warp. Note that choosing  $\delta = \mathbf{0}$  also does the trick. Replacing the (perspective) fundamental matrix  $\mathcal{F}$  by an affine one  $\mathcal{A}$  in the above, and keeping only the first two rows of the equation gives:

$$\mathbf{B}\tilde{\mathbf{q}} \sim \left( \bar{\mathcal{S}}_{\mathcal{A}} + \mathbf{s}_{\mathcal{A}}\boldsymbol{\pi}^{\top} \right) \tilde{\mathbf{q}},$$

where the  $(2 \times 3)$  matrix  $\mathbf{B}$  represents an FA-Warp, and the right hand side of the equation is the general form of such warps.

### C.2 $(\mathbb{S}_{\text{FP}} \setminus \mathbb{S}_{\text{FA}}) \cap \mathbb{S}_{\text{DA}} = \emptyset$

We prove that the set of DA-Warps does not contain any flat, purely perspective warp *i.e.*, an FP-Warp which would not also be an FA-Warp. Let  $\check{\mathbf{H}}$  be a  $(3 \times 3)$  homography matrix representing an element of  $\mathbb{S}_{\text{FP}}$ . Equating with a DA-Warp in homogeneous coordinates (16) gives:

$$\check{\mathbf{H}}\tilde{\mathbf{q}} \sim \tilde{\mathbf{P}}'^{\top} \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}} \quad \forall \mathbf{q} \in \mathbb{R}^2,$$

that we rewrite, thanks to property (7) as:

$$\check{\mathbf{H}}\tilde{\mathbf{q}} \sim \begin{pmatrix} \mathbf{P}'^{\top} \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}} \\ 1 \end{pmatrix} \quad \forall \mathbf{q} \in \mathbb{R}^2,$$

from which, defining  $\bar{\mathbf{H}}$  as the first two rows and  $\mathbf{r}^{\top}$  as the last row of  $\check{\mathbf{H}}$ :

$$\bar{\mathbf{H}}\tilde{\mathbf{q}} = \left( \mathbf{r}^{\top} \tilde{\mathbf{q}} \right) \mathbf{P}'^{\top} \mathcal{E}_{\lambda}^{\top} \boldsymbol{\ell}_{\mathbf{q}} \quad \forall \mathbf{q} \in \mathbb{R}^2,$$

which holds if and only if  $\mathbf{r}^{\top} \sim (0 \ 0 \ 1)$ , implying that  $\check{\mathbf{H}}$  is an FA-Warp represented by the  $(2 \times 3)$  matrix  $\mathbf{B} = \frac{1}{H_{33}} \bar{\mathbf{H}}$ .

### C.3 $\mathbb{S}_{\text{RA}} = \mathbb{S}_{\text{DA}} \cap \mathbb{S}_{\text{RP}}$

We prove that the warps which lie in both the set of DA-Warps and the set of RP-Warps lie in the set of RA-Warps. The converse is straightforward by definition. We begin by equating a DA-Warp and an RP-Warp in homogeneous coordinates, respectively (16) and (23):

$$\tilde{\mathcal{M}}_{\text{DA}} \boldsymbol{\ell}_{\mathbf{q}} \sim \tilde{\mathcal{M}}_{\text{RP}} \boldsymbol{\ell}_{\mathbf{q}} \quad \forall \mathbf{q} \in \mathbb{R}^2.$$

This implies  $\tilde{\mathcal{M}}_{\text{DA}} \sim \tilde{\mathcal{M}}_{\text{RP}}$  and thus:

$$\mathcal{E}_\lambda \tilde{\mathbf{P}}' \sim \mathcal{E}_\lambda \begin{pmatrix} \tilde{\mathbf{P}} & \boldsymbol{\delta} \end{pmatrix} \check{\mathcal{G}}_{\mathcal{F}}^\top.$$

Since  $\mathcal{E}_\lambda$  is a full column rank, portait matrix, this implies:

$$\tilde{\mathbf{P}}' \sim \begin{pmatrix} \tilde{\mathbf{P}} & \boldsymbol{\delta} \end{pmatrix} \check{\mathcal{G}}_{\mathcal{F}}^\top.$$

This equation holds if and only if  $\check{\mathcal{G}}_{\mathcal{F}}$  is an affine camera matrix *i.e.*, if and only if  $\mathcal{F} = \mathcal{A}$ , meaning that the warp is actually an RA-Warp, as is seen by inspecting its homogeneous definition (20).

## D The 3D Coordinate Frame for Rigid Warps

We show that the 3D coordinate frame fixed by choosing a particular canonical projection matrix in §3.3 has no effect on the RA-Warp and the RP-Warp. We demonstrate it for the RP-Warp. Specializing the result to the case of the RA-Warp is straightforward. We assume that a different projection matrix  $(\bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^\top \quad \nu \check{\mathbf{g}}_{\mathcal{F}})$  was used instead of  $\mathcal{G}_{\mathcal{F}}$ . The RP-Warp matrix (23) with parameters  $\boldsymbol{\delta}'$  is rewritten as:

$$\begin{pmatrix} \bar{\mathcal{G}}_{\mathcal{F}} + \check{\mathbf{g}}_{\mathcal{F}} \boldsymbol{\pi}^\top & \nu \check{\mathbf{g}}_{\mathcal{F}} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{P}}^\top \\ \boldsymbol{\delta}'^\top \end{pmatrix} \mathcal{E}_\lambda^\top.$$

Expanding and factorizing this equation gives:

$$\left( \bar{\mathcal{G}}_{\mathcal{F}} \tilde{\mathbf{P}}^\top + \check{\mathbf{g}}_{\mathcal{F}} \left( \boldsymbol{\pi}^\top \tilde{\mathbf{P}}^\top + \nu \boldsymbol{\delta}'^\top \right) \right) \mathcal{E}_\lambda^\top,$$

that we identify with our RP-Warp as:

$$\mathcal{W}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \quad \text{with} \quad \boldsymbol{\delta} = \tilde{\mathbf{P}} \boldsymbol{\pi} + \nu \boldsymbol{\delta}'.$$

Consequently, whatever the chosen target canonical projection matrix and warp parameters  $\boldsymbol{\delta}'$  in the coordinate frame it induces, we can find a strictly equivalent RP-Warp in another canonical coordinate frame.