Generalized Thin-Plate Spline Warps

Adrien Bartoli, Mathieu Perriollat, Sylvie Chambon

Generalized Thin-Plate Spline Warps

Adrien Bartoli
LASMEA (CNRS / UBP)
Clermont-Ferrand, France
Adrien.Bartoli@gmail.com

Mathieu Perriollat
LASMEA (CNRS / UBP)
Clermont-Ferrand, France
Mathieu.Perriollat@gmail.com

Sylvie Chambon
LTCI (CNRS / ENST) – Paris, France
IRIT (CNRS / UPS) – Toulouse, France
schambon@enst.fr

Abstract

Thin-Plate Spline warps have been shown to be very effective as a parameterized model of the optic flow field between images of various deforming surfaces. Examples include a sheet of paper being manually handled. Recent work has used such warps for images of smooth rigid surfaces. Standard Thin-Plate Spline warps are not rigid, in the sense that they do not satisfy the epipolar geometry constraint, and are intrinsically affine, in the sense of the affine camera model.

We propose three types of warps based on the Thin-Plate Spline. The first one is a flexible rigid warp. It describes the optic flow field induced by a smooth rigid surface, and satisfies the affine epipolar geometry constraint. The second and third ones extend the standard Thin-Plate Spline 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, showing that the proposed warps outperform the standard one in several cases of interest.

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 first image to the corresponding point in the second image. For instance, two images of a rigid planar surface taken by two perspective cameras 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 deforms, the warp is even more involved. Examples of rigid scene models include piecewise or nearly planar structures. Examples of deformable scene models include the flexible low-rank shape and face models.

Representing the warp using a parametric function requires prior assumptions about the observed scene structure. One common, fairly generic assumption is that a smooth surface is observed. This naturally leads to using the Thin-Plate Spline (TPS) as a building block for the warps. TPS are smooth, compact and convenient, $\mathbb{R}^2 \rightarrow \mathbb{R}$ functions. Standard TPS warps, built by 'stacking' a pair of TPS, are very flexible in that they are controlled by centres that may be placed anywhere in the images. They are known to be effective approximations to many types of deformations, see *e.g.* [1]. Standard TPS warps have recently been used as simple parametric warps for images of rigid 3D surfaces by Wills and Belongie [8] and Masson *et al.* [5], for respectively wide-baseline matching and object tracking.

There are, however, two main issues with the use of standard TPS warps in this context, that have not been dealt with in the literature. *(i)* **Standard TPS warps overfit affine images of rigid surfaces.** They do not in general satisfy the rigidity constraint modeled by the affine epipolar geometry. In that sense they are 'too flexible' in affine imaging conditions. *(ii)* **Standard TPS warps do not model perspective projection.** They are intrinsically affine, in the sense of the affine camera model, since their formulation does not include a division. For instance, as mentioned in [8], they are not able to 'reproduce' simple homographic warps with a finite number of centres. Henceforth, we call *DA-Warps* the standard TPS warps (for 'Deformable Affine').

This paper addresses the two above-mentioned issues. First, the DA-Warps are specialized to rigid surfaces in §4. These warps are called *RA-Warps* (for 'Rigid Affine') and are very similar to DA-Warps with an epipolar constraint on the warp coefficients. This solves the first issue. Second, the RA-Warps are extended to the perspective camera model in §5. These warps, dubbed *RP-Warps* (for 'Rigid Perspective'), naturally include FP-Warps ('Flat Perspective') similarly to the RA-Warps including FA-Warps[1]. This solves the second issue for the case of rigid surfaces. Third, we introduce the *DP-Warps* (for 'Deformable Perspective') which are shown to be the perspective analogue of the DA-Warps. This solves the second issue for the case of deformable surfaces. The derivation of these warps is made possible by a feature-driven parameterization of the Thin-Plate Spline

[1]The FP-Warps are 2D homographic warps with 8 parameters. The FA-Warps (for 'Flat Affine') are 2D affine warps with 6 parameters.

we propose in §3. The hierarchy and dependencies between the six types of warps mentioned so far is studied in details in §6. In order to derive warps independent of the intrinsic camera parameters, we consider uncalibrated cameras.

The second line of contributions in this paper, presented in §7, is a set of algorithms for estimating the proposed warps from image point correspondences. Experimental results are reported in §8 and our contributions discussed in §9. Most proofs of our statement will appear in an extended version of the paper.

2. Preliminaries

Previous work. DA-Warps, *i.e.* standard TPS warps, are used in many different contexts. While there is a great body of work on defining alternative warps, such as FFD (Free-Form Deformations) [6] or more recently diffeomorphic warps, DA-Warps are usually used in their original form. Since the seminal paper by Bookstein [1], the literature is mostly focused on estimation methods. Bookstein proposed a method relying on point landmark correspondences, that are chosen as centres for the DA-Warp. The DA-Warp and point matching are simultaneously estimated using the soft-assign in [2]. Algorithms to make faster the computation of DA-Warps from point correspondences are proposed in [3]. Several papers use the integral bending energy for 3D surface reconstruction, for instance, as one of the terms in an energy functional, see *e.g.* [7]. In contrast, we build on the existing DA-Warps to derive new warps by taking into account the possible rigidity of the observed surface and the perspective camera model.

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.* P and $\mathcal{E}$. The elements of a vector are written as in $\mathbf{a}^\mathsf{T} = (a_1 \;\; a_2 \;\; a_3)$ where $^\mathsf{T}$ is vector and matrix transpose. We do not make a difference between coordinate vectors and physical entities. The coordinates of a point in the first image are written with a 2-vector $\mathbf{q}^\mathsf{T} = (x \;\; y)$. $\mathbb{R}^r$ 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 $\check{\mathbf{q}}^\mathsf{T} \sim (\mathbf{q}^\mathsf{T} \;\; 1)$, where $\sim$ means equality up to scale. Scaled homogeneous coordinates are written $\tilde{\mathbf{q}}^\mathsf{T} = (\mathbf{q}^\mathsf{T} \;\; 1)$. The homogeneous to affine coordinate function ψ is defined by $\mathbf{q} = \psi(\check{\mathbf{q}})$. The skew-symmetric (3×3) 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 $\mathsf{X}^\dagger = \left(\mathsf{X}^\mathsf{T}\mathsf{X}\right)^{-1}\mathsf{X}^\mathsf{T}$.

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

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

$$\mathcal{W}_{\text{FA}}(\mathbf{q}; \mathsf{A}) \stackrel{\text{def}}{=} \mathsf{A}\tilde{\mathbf{q}} \qquad \text{and} \qquad \check{\mathcal{W}}_{\text{FP}}(\mathbf{q}; \mathsf{H}) \stackrel{\text{def}}{\sim} \mathsf{H}\tilde{\mathbf{q}}.$$

Thin-Plate Splines. The TPS is an $\mathbb{R}^2 \to \mathbb{R}$ function driven by assigning target values α_k to the l 2D centres $\mathbf{c}_k$ and enforcing several conditions: the TPS is the Radial Basis Function (RBF) that minimizes the integral bending energy. It is usually parameterized by an $l+3$ coefficient vector $\mathbf{h}_{\boldsymbol{\alpha},\lambda}^\mathsf{T} = (\;\mathbf{w}^\mathsf{T} \quad \mathbf{a}^\mathsf{T})$ computed from the target vector $\boldsymbol{\alpha}$ and a regularization parameter $\lambda \in \mathbb{R}^+$. The coefficients in $\mathbf{w}$ must satisfie $\tilde{\mathsf{P}}^\mathsf{T}\mathbf{w} = \mathbf{0}$. These three 'side-conditions' ensure that the TPS has square integrable second derivatives. The TPS is defined by:

$$\omega(\mathbf{q}, \mathbf{h}_{\boldsymbol{\alpha},\lambda}) \stackrel{\text{def}}{=} \boldsymbol{\ell}_\mathbf{q}^\mathsf{T}\mathbf{h}_{\boldsymbol{\alpha},\lambda}, \tag{1}$$

with $\boldsymbol{\ell}_\mathbf{q}^\mathsf{T} \stackrel{\text{def}}{=} (\rho(d^2(\mathbf{q}, \mathbf{c}_1)) \;\; \cdots \;\; \rho(d^2(\mathbf{q}, \mathbf{c}_l)) \;\; \tilde{\mathbf{q}}^\mathsf{T})$ and $\rho(d) \stackrel{\text{def}}{=} d\log(d)$ is the TPS kernel function for the squared distance. Combining the equations obtained for all the l centres $\mathbf{c}_r$ with target values α_r in a single matrix equation gives:

$$\mathsf{K}_\lambda\mathbf{w} + \tilde{\mathsf{P}}\mathbf{a} = \boldsymbol{\alpha}, \;\; K_{r,k} = \begin{cases} \lambda & r = k \\ \rho(d^2(\mathbf{c}_r, \mathbf{c}_k)) & r \neq k. \end{cases} \tag{2}$$

Adding $\lambda\mathsf{I}$ acts as a regularizer. Solving for $\mathbf{h}_{\boldsymbol{\alpha},\lambda}$ using the above equation and the side-conditions is the classical linear method for estimating the TPS coefficients due to Bookstein [1]. The coefficient vector $\mathbf{h}_{\boldsymbol{\alpha},\lambda}$ is a nonlinear function of the regularization parameter λ and a linear function of the target vector $\boldsymbol{\alpha}$.

Rigid surfaces, fundamental and projection matrices. 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. In both cases, the rigidity constraint is $\tilde{\mathbf{q}}'^\mathsf{T}\mathcal{F}\tilde{\mathbf{q}} = 0$. A warp $\mathcal{W}$ is rigid if and only if:

$$\check{\mathcal{W}}(\mathbf{q})^\mathsf{T}\mathcal{F}\tilde{\mathbf{q}} = 0 \qquad \forall \mathbf{q} \in \mathbb{R}^2. \tag{3}$$

Parameterizing the affine fundamental matrix as:

$$\mathcal{A} \stackrel{\text{def}}{\sim} \begin{pmatrix} 0 & 0 & \\ 0 & 0 & \boldsymbol{\imath} \\ & \boldsymbol{\jmath}^\mathsf{T} & \end{pmatrix} \text{ with } \begin{cases} \boldsymbol{\jmath}^\mathsf{T} \stackrel{\text{def}}{=} (c \;\; d \;\; e) \\ \boldsymbol{\imath}^\mathsf{T} \stackrel{\text{def}}{=} (a \;\; b), \end{cases} \tag{4}$$

we rewrite the definition (3) of a rigid affine warp as:

$$\mathcal{W}(\mathbf{q})^{\mathsf{T}}\boldsymbol{\imath} + \tilde{\mathbf{q}}^{\mathsf{T}}\boldsymbol{\jmath} = 0 \qquad \forall \mathbf{q} \in \mathbb{R}^2. \tag{5}$$

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 $\mathbb{P}^4$, see *e.g.* [4, §9.2].

The fundamental matrix is an implicit reconstruction of the two cameras. Canonical cameras are obtained by setting the first (3×4) camera matrix to $(\mathrm{I} \;\; \mathbf{0})\mathcal{M}$ and the second one to $\mathcal{G}_{\mathcal{F}} = ([\breve{\mathbf{e}}']_{\times}\mathcal{F} \;\; \breve{\mathbf{e}}')\mathcal{M}$, where the second epipole $\mathbf{e}'$ is defined by $\mathcal{F}^{\mathsf{T}}\breve{\mathbf{e}}' = \mathbf{0}$. Matrix $\mathcal{M}$ simply swaps the third and fourth coordinates, making affine the first camera, even in the perspective case. Note that $\mathcal{M}\mathcal{M} \sim \mathrm{I}$. In the affine case, we write $\mathcal{S}_{\mathcal{A}}$ the first two rows of $\mathcal{G}_{\mathcal{F}}$, the third row being $(0 \;\; 0 \;\; 0 \;\; 1)$. Within this canonical reconstruction basis, a 3D point with depth δ can be written[2] $\tilde{\mathbf{Q}}^{\mathsf{T}} \sim (\mathbf{q}^{\mathsf{T}} \;\; \delta \;\; 1)$. Reprojecting a 3D point in the second camera $\mathcal{G}_{\mathcal{F}}$ gives the transfer equation $\tilde{\mathbf{q}}' \sim \bar{\mathcal{G}}_{\mathcal{F}}\tilde{\mathbf{q}} + \mathbf{g}_{\mathcal{F}}\delta$ with $\bar{\mathcal{G}}_{\mathcal{F}}$ the first, second and fourth columns of $\mathcal{G}_{\mathcal{F}}$ and $\mathbf{g}_{\mathcal{F}}$ the third one. In the affine case, the second camera matrix $\mathcal{S}_{\mathcal{A}}$ is (2×4). We Define $\bar{\mathcal{S}}_{\mathcal{A}}$ and $\mathbf{s}_{\mathcal{A}}$ similarly to $\bar{\mathcal{G}}_{\mathcal{F}}$ and $\mathbf{g}_{\mathcal{F}}$.

3. Feature-Driven Parameterization of the TPS

We write $\mathbf{h}_{\boldsymbol{\alpha},\lambda} = \mathcal{E}_\lambda\boldsymbol{\alpha}$, *i.e.* as a linear 'back-projection' of the target vector $\boldsymbol{\alpha}$. Matrix $\mathcal{E}_\lambda$ nonlinearly depends on λ. It is given from equation (2) as a function of K_λ and $\tilde{\mathsf{P}}$ by:

$$\mathcal{E}_\lambda \stackrel{\text{def}}{=} \begin{pmatrix} \mathsf{K}_\lambda^{-1}\left(\mathrm{I} - \tilde{\mathsf{P}}\left(\tilde{\mathsf{P}}^{\mathsf{T}}\mathsf{K}_\lambda^{-1}\tilde{\mathsf{P}}\right)^{-1}\tilde{\mathsf{P}}^{\mathsf{T}}\mathsf{K}_\lambda^{-1}\right) \\ \left(\tilde{\mathsf{P}}^{\mathsf{T}}\mathsf{K}_\lambda^{-1}\tilde{\mathsf{P}}\right)^{-1}\tilde{\mathsf{P}}^{\mathsf{T}}\mathsf{K}_\lambda^{-1} \end{pmatrix}.$$

This parameterization has the advantages to separate λ and $\boldsymbol{\alpha}$ and introduces units[3]. The side-conditions are naturally enforced by this parameterization.

Incorporating this parameterization into the TPS (1) we obtain what we call the *feature-driven* parameterization $\tau(\mathbf{q};\boldsymbol{\alpha},\lambda) = \omega(\mathbf{q};\mathbf{h}_{\boldsymbol{\alpha},\lambda})$ for the TPS:

$$\boxed{\tau(\mathbf{q};\boldsymbol{\alpha},\lambda) \stackrel{\text{def}}{=} \boldsymbol{\ell}_{\mathbf{q}}^{\mathsf{T}}\mathcal{E}_\lambda\boldsymbol{\alpha}.} \tag{6}$$

This is a feature-driven parameterization since $\boldsymbol{\alpha}$ contains the coordinates of the target centres. In practice, these are image points. The following important properties hold:

$$\boldsymbol{\ell}_{\mathbf{q}}^{\mathsf{T}}\mathcal{E}_\lambda\mathbf{1} = 1 \text{ and } \tilde{\mathbf{q}}^{\mathsf{T}}\boldsymbol{\theta} = \boldsymbol{\ell}_{\mathbf{q}}^{\mathsf{T}}\mathcal{E}_\lambda\tilde{\mathsf{P}}\boldsymbol{\theta} \quad \forall \mathbf{q} \in \mathbb{R}^2 \quad \forall \boldsymbol{\theta} \in \mathbb{R}^3. \tag{7}$$

[2] δ is actually the inverse of the depth relative to the first 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) and handles points at infinity.

[3] While $\mathbf{h}_{\boldsymbol{\alpha},\lambda}$ has no obvious unit, $\boldsymbol{\alpha}$ in general has (*e.g.* pixels, meters). This stems from $\mathcal{E}_\lambda\tilde{\mathsf{P}} = (0 \;\; \mathrm{I})^{\mathsf{T}}$. The asymptotic regularization behaviour of the TPS is an affine transformation:

$$\lim_{\lambda\to+\infty} \tau(\mathbf{q};\boldsymbol{\alpha},\lambda) = \boldsymbol{\zeta}^{\mathsf{T}}\tilde{\mathbf{q}}, \;\; \boldsymbol{\zeta} \stackrel{\text{def}}{=} \tilde{\mathsf{P}}^{\dagger}\boldsymbol{\alpha}.$$

4. Warps with the Affine Camera Model

4.1. DA-Warps – Standard TPS-Warps

Derivation and properties. Standard $\mathbb{R}^2 \to \mathbb{R}^2$ TPS-Warps are obtained by stacking two $\mathbb{R}^2 \to \mathbb{R}$ TPS sharing their centres and regularization parameter. Using (6), this gives $(\tau(\mathbf{q};\mathbf{x},\lambda) \;\; \tau(\mathbf{q};\mathbf{y},\lambda)) = \boldsymbol{\ell}_{\mathbf{q}}^{\mathsf{T}}\mathcal{E}_\lambda\begin{pmatrix}\mathbf{x} & \mathbf{y}\end{pmatrix}$. DA-Warps are thus defined as:

$$\boxed{\mathcal{W}_{\text{DA}}(\mathbf{q};\mathsf{P}',\lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\text{DA}}\boldsymbol{\ell}_{\mathbf{q}}, \;\; \mathcal{M}_{\text{DA}}^{\mathsf{T}} \stackrel{\text{def}}{=} \mathcal{E}_\lambda\mathsf{P}'.} \tag{8}$$

We call this a Deformable Affine Thin-Plate Spline Warp, or 'DA-Warp' since it models images of deformable surfaces and corresponds to an affine camera model. Thanks to property (7), we write homogeneous DA-Warps as $\tilde{\mathcal{W}}_{\text{DA}}(\mathbf{q};\mathsf{P}',\lambda) \stackrel{\text{def}}{=} \tilde{\mathcal{M}}_{\text{DA}}\boldsymbol{\ell}_{\mathbf{q}}$ with $\tilde{\mathcal{M}}_{\text{DA}}^{\mathsf{T}} \stackrel{\text{def}}{=} \mathcal{E}_\lambda\tilde{\mathsf{P}}'$.

Figure 1. The DA-Warps are interpreted as relating the projection of a deforming 3D surface by two affine cameras.

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

$$\lim_{\lambda\to+\infty} \mathcal{W}_{\text{DA}}(\mathbf{q};\mathsf{P}',\lambda) = \mathcal{L}_{\text{DA}}\tilde{\mathbf{q}}', \;\; \mathcal{L}_{\text{DA}}^{\mathsf{T}} \stackrel{\text{def}}{=} \tilde{\mathsf{P}}^{\dagger}\mathsf{P}'.$$

In other words, a DA-Warp tends to a flat affine warp represented by the (2×3) matrix $\mathcal{L}_{\text{DA}}$. It can be shown that $\mathcal{L}_{\text{DA}}$ minimizes the transfer error[4].

[4] The transfer error is the discrepancy between the data points in the second image and the points transferred by the warp from the first image, see §7 for more details.

A projected deformable surface interpretation. We propose a geometrical interpretation of the DA-Warps as the warps induced by the observation of a deforming surface with two affine cameras, as illustrated in figure 1. In order to model the surface depth, its motion and deformation, we introduce an $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ map, parameterizing the surface in a concise manner by the 3D coordinates $\mathbf{C}'_k$ of the l centres. This map is built by stacking three TPS sharing their centres and regularization parameter. More formally, gathering the '3D centres' $\mathbf{C}'_k$ in a single $(l \times 3)$ matrix $\mathsf{Z}^\mathsf{T} = (\mathbf{C}'_1 \ \cdots \ \mathbf{C}'_l)$, the map is written:

$$\mathcal{R}(\mathbf{q}; \mathsf{Z}, \lambda) \stackrel{\text{def}}{=} \mathcal{M}_{3\text{D}} \boldsymbol{\ell}_\mathbf{q}, \ \ \mathcal{M}_{3\text{D}}^\mathsf{T} \stackrel{\text{def}}{=} \mathcal{E}_\lambda \mathsf{Z}. \tag{9}$$

Reprojecting a 3D surface point gives $\mathbf{q}' = \bar{\mathcal{S}}_\mathcal{A} \mathsf{Z}^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\ell}_\mathbf{q} + \mathbf{s}_\mathcal{A}$. Using property (7), we get $\mathbf{q}' = \mathcal{S}_\mathcal{A} \tilde{\mathsf{Z}}^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\ell}_\mathbf{q}$, that we identify with a DA-Warp (8), giving:

$$\mathbf{q}' = \mathcal{W}_{\text{DA}}(\mathbf{q}; \tilde{\mathsf{Z}} \mathcal{S}_\mathcal{A}^\mathsf{T}, \lambda).$$

This shows that the 3D centres $\mathbf{C}'_k$ can be replaced by the 2D centres $\mathbf{c}'_k$ in the second image since $\mathsf{P}' = \tilde{\mathsf{Z}} \mathcal{S}_\mathcal{A}^\mathsf{T}$.

This geometric interpretation does not only provide a strong intuition on the fact that the DA-Warps are intrinsically affine, but also a setting for naturally deriving the DP-Warps, their perspective projection extension, in §5.2.

4.2. RA-Warps – Rigid Affine Warps

Derivation and properties. Applying the rigid affine warp definition (5) to a DA-Warp (8) gives:

$$\boldsymbol{\ell}_\mathbf{q}^\mathsf{T} \mathcal{E}_\lambda \mathsf{P}' \boldsymbol{\imath} + \tilde{\mathbf{q}}^\mathsf{T} \boldsymbol{\jmath} = 0 \qquad \forall \mathbf{q} \in \mathbb{R}^2.$$

Using property (7) gives $\boldsymbol{\ell}_\mathbf{q}^\mathsf{T} \mathcal{E}_\lambda \left(\mathsf{P}' \boldsymbol{\imath} + \tilde{\mathsf{P}} \boldsymbol{\jmath} \right) = 0, \forall \mathbf{q} \in \mathbb{R}^2$. This implies $\mathsf{P}' \boldsymbol{\imath} + \tilde{\mathsf{P}} \boldsymbol{\jmath} = \mathbf{0}_{(l \times 1)}$, which is the epipolar constraint for all pairs of centres. We call it the rigidity consistency constraint for DA-Warps. This means that if the warp satifies the epipolar geometry, then the centres also have to.

Each pair of centres $\mathbf{c}_k \leftrightarrow \mathbf{c}'_k$ satisfies the epipolar constraint and thus is the projection of a 3D point $\mathbf{C}_k^\mathsf{T} = (\mathbf{c}_k^\mathsf{T} \ \delta_k)$ in the canonical basis. Reprojecting all centres in the second image gives $\mathsf{P}' = (\mathsf{P} \ \boldsymbol{\delta} \ \mathbf{1}) \mathcal{S}_\mathcal{A}^\mathsf{T}$. Substituting into the DA-Warp formulation (8) gives:

$$\mathcal{W}_{\text{DA}}(\mathbf{q}; \mathsf{P}', \lambda) = \mathcal{S}_\mathcal{A} (\mathsf{P} \ \boldsymbol{\delta} \ \mathbf{1})^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\ell}_\mathbf{q},$$

which can be seen as the projection of some 3D point by the second camera, thereby satisfying the rigidity constraint, completing the proof. We thus define:

$$\boxed{\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\text{RA}} \boldsymbol{\ell}_\mathbf{q}, \ \ \mathcal{M}_{\text{RA}}^\mathsf{T} \stackrel{\text{def}}{=} \mathcal{E}_\lambda (\mathsf{P} \ \boldsymbol{\delta} \ \mathbf{1}) \mathcal{S}_\mathcal{A}^\mathsf{T}.} \tag{10}$$

This definition of RA-Warps can be made homogeneous by replacing the (2×4) camera $\mathcal{S}_\mathcal{A}$ by its (3×4) equivalent $\mathcal{G}_\mathcal{A}$ in the above equations, giving $\tilde{\mathcal{W}}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) \stackrel{\text{def}}{=} \tilde{\mathcal{M}}_{\text{RA}} \boldsymbol{\ell}_\mathbf{q}$ with $\tilde{\mathcal{M}}_{\text{RA}}^\mathsf{T} \stackrel{\text{def}}{=} \mathcal{E}_\lambda (\mathsf{P} \ \boldsymbol{\delta} \ \mathbf{1}) \mathcal{G}_\mathcal{A}^\mathsf{T}$.

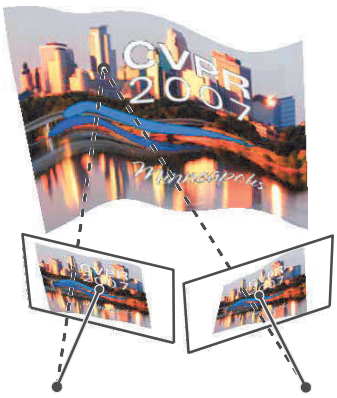

Figure 2. The RA-Warps are interpreted as relating the projection of a rigid smooth 3D surface by two affine cameras.

From the above derivation follows that the set of RA-Warps, denoted $\mathbb{S}_{\text{RA}}$, is a subset of $\mathbb{S}_{\text{DA}}$. In can be shown that the set of flat affine warps $\mathbb{S}_{\text{FA}}$ is included into $\mathbb{S}_{\text{RA}}$. The RA-Warps have $l + 4$ degrees of freedom through $(\boldsymbol{\delta}, \mathcal{A}) \in \mathbb{R}^l \times \mathbb{P}^4$. Parameters $\boldsymbol{\delta}$ are the depth of the centres with respect to the first camera.

The asymptotic regularization behaviour of the RA-Warps is derived directly from the one for the DA-Warps (8):

$$\lim_{\lambda \rightarrow +\infty} \mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \mathcal{L}_{\text{RA}} \tilde{\mathbf{q}}, \ \ \mathcal{L}_{\text{RA}}^\mathsf{T} \stackrel{\text{def}}{=} \tilde{\mathsf{P}}^\dagger (\mathsf{P} \ \boldsymbol{\delta} \ \mathbf{1}) \mathcal{S}_\mathcal{A}^\mathsf{T}.$$

In other words, an RA-Warp tends to a flat affine warp represented by the (2×3) matrix $\mathcal{L}_{\text{RA}}$. It can be shown to be rigid and can be written $\mathcal{L}_{\text{RA}} = \mathcal{S}_\mathcal{A} + \mathbf{s}_\mathcal{A} \boldsymbol{\delta}^\mathsf{T} \tilde{\mathsf{P}}^\dagger$, a plane-induced affine warp minimizing the transfer error, under the assumption that the point correspondences satisfy the rigidity constraint.

A projected rigid surface interpretation. A geometric interpretation of the RA-Warps, illustrated in figure 2, directly stems from their definition (10). The RA-Warps are induced by a surface defined as a Monge patch parameterized by a TPS mapping points from the first image to their depths. This $\mathbb{R}^2 \rightarrow \mathbb{R}$ TPS is of the form (6) and has the same centres as the RA-Warp. This is derived by expanding the formulation (10) of the RA-Warps and property (7):

$$\mathcal{W}_{\text{RA}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{A}, \lambda) = \bar{\mathcal{S}}_\mathcal{A} \tilde{\mathbf{q}} + \mathbf{s}_\mathcal{A} \tau(\mathbf{q}; \boldsymbol{\delta}, \lambda). \tag{11}$$

5. Warps with the Perspective Camera Model

5.1. RP-Warps – Rigid Perspective Warps

We derive the RP-Warps by introducing perspective projection in the RA-Warps. Following §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 Monge patch, and reproject it in the second image, giving from equation (11):

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

Replacing τ by its expression (6), and applying property (7) to each of the three rows of $\bar{\mathcal{G}}_\mathcal{F}$, we get $\check{\mathcal{W}}_{\text{RP}}(\mathbf{q}; \boldsymbol{\delta}, \mathcal{F}, \lambda) \sim$

$(\bar{\mathcal{G}}_{\mathcal{F}}\tilde{\mathsf{P}}^{\mathsf{T}} + \mathbf{g}_{\mathcal{F}}\boldsymbol{\delta}^{\mathsf{T}})\mathcal{E}_{\lambda}^{\mathsf{T}}\boldsymbol{\ell}_{\mathbf{q}}$ and thus:

$$\boxed{\check{\mathcal{W}}_{\text{RP}}(\mathbf{q};\boldsymbol{\delta},\mathcal{F},\lambda) \overset{\text{def}}{\sim} \check{\mathcal{M}}_{\text{RP}}\boldsymbol{\ell}_{\mathbf{q}},\ \ \check{\mathcal{M}}_{\text{RP}}^{\mathsf{T}} \overset{\text{def}}{\sim} \mathcal{E}_{\lambda}(\mathsf{P}\ \ \boldsymbol{\delta}\ \ \mathbf{1})\mathcal{G}_{\mathcal{F}}^{\mathsf{T}}.} \quad (12)$$

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

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

The asymptotic regularization behaviour is:

$$\lim_{\lambda\to+\infty} \check{\mathcal{W}}_{\text{RP}}(\mathbf{q};\boldsymbol{\delta},\mathcal{F},\lambda) \sim \check{\mathcal{L}}_{\text{RP}}\tilde{\mathbf{q}},\ \ \check{\mathcal{L}}_{\text{RP}}^{\mathsf{T}} \overset{\text{def}}{\sim} \tilde{\mathsf{P}}^{\dagger}(\mathsf{P}\ \ \boldsymbol{\delta}\ \ \mathbf{1})\mathcal{G}_{\mathcal{F}}^{\mathsf{T}}.$$

An RP-Warp thus tends to a flat perspective warp represented by the (3×3) homogeneous homography matrix $\check{\mathcal{L}}_{\text{RP}}$. It can be shown to be the plane-induced rigid warp $\check{\mathcal{L}}_{\text{RP}} \sim \mathcal{G}_{\mathcal{F}} + \mathbf{g}_{\mathcal{F}}\boldsymbol{\delta}^{\mathsf{T}}\tilde{\mathsf{P}}^{\dagger}$ minimizing an algebraic transfer error, under the assumption that the point correspondences satisfy the rigidity constraint.

5.2. DP-Warps – Deformable Perspective Warps

The DP-Warps form a superset of all other warps derived so far in this paper, including the standard DA-Warps. The RP-Warps are derived by introducing perspective projection in the RA-Warps. We derive the DP-Warps from the DA-Warps in the same spirit. Consider the deformable surface geometric interpretation shown in figure 1. The surface seen by the second camera is defined by an $\mathbb{R}^2 \to \mathbb{R}^3$ map $\mathcal{R}(\mathbf{q};\mathsf{Z},\lambda)$. Projecting a point on this surface to the second image gives $\check{\mathbf{q}}' \sim \bar{\mathcal{G}}_{\mathcal{F}}\mathcal{R}(\mathbf{q};\mathsf{Z},\lambda) + \mathbf{g}_{\mathcal{F}}$. Substituting the map (9) defining the 3D surface gives:

$$\check{\mathbf{q}}' \sim \bar{\mathcal{G}}_{\mathcal{F}}\mathsf{Z}^{\mathsf{T}}\mathcal{E}_{\lambda}^{\mathsf{T}}\boldsymbol{\ell}_{\mathbf{q}} + \mathbf{g}_{\mathcal{F}}.$$

Using property (7), we get $\check{\mathbf{q}}' \sim \bar{\mathcal{G}}_{\mathcal{F}}\tilde{\mathsf{Z}}^{\mathsf{T}}\mathcal{E}_{\lambda}^{\mathsf{T}}\boldsymbol{\ell}_{\mathbf{q}}$. The centres in the second image are the reprojection of the '3D centres' in matrix Z. The weights of the homogeneous coordinates in $\check{\mathsf{P}}'$ are important: they model the perspective part of the DP-Warps. We thus define the DP-Warps as:

$$\boxed{\check{\mathcal{W}}_{\text{DP}}(\mathbf{q};\check{\mathsf{P}}',\lambda) \overset{\text{def}}{\sim} \check{\mathcal{M}}_{\text{DP}}\boldsymbol{\ell}_{\mathbf{q}},\ \ \check{\mathcal{M}}_{\text{DP}}^{\mathsf{T}} \overset{\text{def}}{\sim} \mathcal{E}_{\lambda}\check{\mathsf{P}}'.} \quad (13)$$

The affine coordinates are obtained as ratios of linear functions through $\mathcal{W}_{\text{DP}}(\mathbf{q};\check{\mathsf{P}}',\lambda) \overset{\text{def}}{=} \psi(\check{\mathcal{W}}_{\text{DP}}(\mathbf{q};\check{\mathsf{P}}',\lambda))$.

The set of DP-Warps, denoted $\mathbb{S}_{\text{DP}}$, forms a superset of $\mathbb{S}_{\text{RP}}$ and a superset of $\mathbb{S}_{\text{DA}}$. The DP-Warps have parameters $\check{\mathsf{P}}'$ and thus $3l-1$ degrees of freedom. Consequently, they can not be estimated by choosing as centres all data points: each point correspondence giving two contraints, we end up with $2l$ contraints, which is less than the $3l-1$ unknowns. Methods for estimating DP-Warps are reported in §7.

The asymptotic 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.* 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-Warps and the number of constraints given by interpolating the l centres:

$$\lim_{\lambda\to+\infty} \check{\mathcal{W}}_{\text{DP}}(\mathbf{q};\check{\mathsf{P}}',\lambda) \sim \check{\mathcal{L}}_{\text{DP}}\tilde{\mathbf{q}},\ \ \check{\mathcal{L}}_{\text{DP}}^{\mathsf{T}} \overset{\text{def}}{\sim} \tilde{\mathsf{P}}^{\dagger}\text{diag}(\mathbf{d})\tilde{\mathsf{P}}'.$$

The $(l\times 1)$ vector $\mathbf{d}$, defined up to scale, represents the $l-1$ free parameters of the limiting flat perspective warp, represented by matrix $\check{\mathcal{L}}_{\text{DP}}$, minimizing an algebraic transfer error, different from the one we use in §7.

6. A Hierarchy of Warps

The aim of this section is to define a hierarchy between the sets of standard DA-Warps $\mathbb{S}_{\text{DA}}$, of flat affine and perspective warps $\mathbb{S}_{\text{FA}}$ and $\mathbb{S}_{\text{FP}}$, and of the three types of warps we introduced, $\mathbb{S}_{\text{RA}}$, $\mathbb{S}_{\text{RP}}$ and $\mathbb{S}_{\text{DP}}$. The whole hierarchy is illustrated in figure 3. So far, we have established $\mathbb{S}_{\text{RA}} \subset \mathbb{S}_{\text{DA}}$ and $\mathbb{S}_{\text{RA}} \subset \mathbb{S}_{\text{RP}}$. Intuitively, the common warps to $\mathbb{S}_{\text{DA}}$ and $\mathbb{S}_{\text{RP}}$ must be rigid and affine. More precisely, we have $\mathbb{S}_{\text{RA}} = \mathbb{S}_{\text{DA}} \cap \mathbb{S}_{\text{RP}}$. We also established that $\mathbb{S}_{\text{DP}}$ is a superset of all the other warps, *i.e.* $\mathbb{S}_{\text{RP}} \subset \mathbb{S}_{\text{DP}}$ and $\mathbb{S}_{\text{DA}} \subset \mathbb{S}_{\text{DP}}$, and thus $\mathbb{S}_{\text{RA}} \subset \mathbb{S}_{\text{DP}}$. The set of DA-Warps $\mathbb{S}_{\text{DA}}$ does not contain any flat perspective warp. More formally, $(\mathbb{S}_{\text{FP}} - \mathbb{S}_{\text{FA}}) \cap \mathbb{S}_{\text{DA}} = \emptyset$, implying $(\mathbb{S}_{\text{FP}} - \mathbb{S}_{\text{FA}}) \cap \mathbb{S}_{\text{RA}} = \emptyset$.

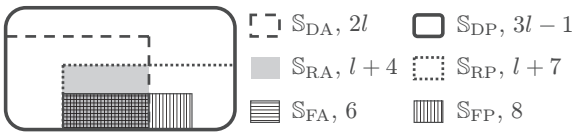

Figure 3. Hierarchical representation for the three proposed types of warps – RA, RP and DP – along with the standard TPS warps, dubbed DA-Warps, and the flat warps FA-Warps and FP-Warps. D stands for Deformable, R for rigid, F for flat, A for Affine and P for Perspective. The number of degrees of freedom for l centres is indicated for each set of warps.

7. Estimation of the Warps

We propose warp estimation methods from m point correspondences $\mathbf{q}_j \leftrightarrow \mathbf{q}'_j$. We examine two cases for the centres in the first image. *(i): 'centre-on-data'* – the centres in the first image coincide with the data points, *i.e.* $m = l$.

(ii): 'arbitrary-centres' – the centres in the first image may not coincide with the data points. They are typically chosen on a regular grid or as interest points. We assume $m \geq l$, *i.e.* we have sufficiently many point correspondences to estimate the warp without having to regularize it.

All warps are estimated by minimizing the transfer error, *i.e.* the discrepancy, measured by the Euclidean distance, between the data points in the second image, and the corresponding points transferred by the sought after warp from the first image. Criteria based on a 3D depth error for the rigid warps are avoided since they are not physically meaningful for uncalibrated cameras. For most warps, the transfer error is nonlinear. 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) of the design matrix, if the system is homogeneous, enforcing unit two-norm on the unknown vector. The initial estimate is refined by iteratively minimizing the transfer error through Nonlinear Least Squares minimization (NLS) with the Levenberg-Marquardt algorithm, see *e.g.* [4, §A]. For a warp $\mathcal{W}$ with parameters $\mathcal{U}$, the minimization problem is generically written:

$$\vartheta(\mathcal{W},\mathcal{U}) \stackrel{\text{def}}{=} \min_{\mathcal{U}} \sum_{j=1}^{n} \|\mathcal{W}(\mathbf{q}_j;\mathcal{U},\lambda) - \mathbf{q}'_j\|^2.$$

DA-Warps. The minimal number of point correspondences is $m \geq 3$. In the centre-on-data case, this is the classical problem solved by Bookstein [1]. With our feature-driven parameterization (8), the transformation is readily expressed in terms of the centre coordinates P'. Note that the data points are interpolated, which nullifies the transfer error. In the arbitrary-centre case, we solve $\vartheta(\mathcal{W}_{\text{DA}}, \mathsf{P}')$. Writing $\boldsymbol{\ell}_{\mathbf{q}_j}$ as $\boldsymbol{\ell}_j$, and replacing $\mathcal{W}_{\text{DA}}$ by its expression (8), we get an LLS problem:

$$\min_{\mathsf{P}'} \sum_{j=1}^{m} \left\| \boldsymbol{\ell}_j^\mathsf{T} \mathcal{E}_\lambda \mathsf{P}' - \mathbf{q}_j'^\mathsf{T} \right\|^2.$$

RA-Warps. A single algorithm solves both the centre-on-data and arbitrary-centre cases for $m \geq 4$ point correspondences. RA-Warps depend on the affine fundamental matrix $\mathcal{A}$ and a depth vector $\boldsymbol{\delta}$. Contrarily to the DA-Warp case, $\vartheta(\mathcal{W}_{\text{RA}}, \{\boldsymbol{\delta}, \mathcal{A}\})$ is an NLS problem due to the coupling between $\boldsymbol{\delta}$ and $\mathcal{A}$ in the expression (10) of the warp. In order to find an initial estimate, we use a two-step procedure. We compute $\mathcal{A}$ using *e.g.* the Gold Standard algorithm in [4, §14.3]. Given $\mathcal{A}$, finding $\boldsymbol{\delta}$ by minimizing the transfer error, *i.e.* solving $\vartheta(\mathcal{W}_{\text{RA}}, \boldsymbol{\delta})$ turns out to be an LLS problem:

$$\min_{\boldsymbol{\delta}} \sum_{j=1}^{m} \left\| \mathbf{s}_\mathcal{A} \boldsymbol{\ell}_j^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\delta} + \boldsymbol{\ell}_j^\mathsf{T} \mathcal{E}_\lambda \tilde{\mathsf{P}} \mathcal{S}_\mathcal{A}^\mathsf{T} - \mathbf{q}_j'^\mathsf{T} \right\|^2.$$

RP-Warps. For both the centre-on-data and the arbitrary-centre cases, the minimal number of point correspondences is $m \geq 7$. RP-Warps depend on the fundamental matrix $\mathcal{F}$ and a depth vector $\boldsymbol{\delta}$. $\vartheta(\mathcal{W}_{\text{RP}}, \{\boldsymbol{\delta}, \mathcal{F}\})$ is an NLS problem, for several reasons: (i) $\boldsymbol{\delta}$ and $\mathcal{F}$ are coupled in the homogeneous expression (12) of the warp, (ii) finding the affine coordinates of the transferred point requires a division and (iii) the fundamental matrix must fulfill a nonlinear rank-deficiency constraint[5]. Similarly to the algorithm for the RA-Warps, we use a two-step initialization procedure. We compute $\mathcal{F}$ using *e.g.* the Gold Standard algorithm in [4, §11.4]. Given $\mathcal{F}$, we estimate $\boldsymbol{\delta}$ by minimizing an algebraic approximation to the transfer error:

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

with $d_a^2(\check{\mathbf{q}}, \check{\mathbf{q}}') = \|\mathsf{S}[\check{\mathbf{q}}]_\times \check{\mathbf{q}}'\|^2$ an algebraic distance between points $\mathbf{q}$ and $\mathbf{q}'$, and $\mathsf{S} = (\mathrm{I}\ \mathbf{0})$ simply selects the first two rows of the cross-product. The algebraic approximation yields an LLS minimization problem since the algebraic distance directly compares homogeneous coordinate vectors, thereby avoiding the need for the perspective division. Normalizing the image coordinates is crucial to make the algebraic distance 'similar to' the Euclidean one [4]. Substituting d_a by its expression, and the RP-Warp by its homogeneous formulation (12), we get $\min_{\boldsymbol{\delta}} \sum_{j=1}^{m} \|\mathsf{S}[\tilde{\mathbf{q}}'_j]_\times \mathcal{G}_\mathcal{F} (\mathsf{P}\ \ \boldsymbol{\delta}\ \ \mathbf{1})^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\ell}_j\|^2$ and as sought, after minor algebraic manipulations, an LLS problem:

$$\min_{\boldsymbol{\delta}} \sum_{j=1}^{m} \left\| \mathsf{S}[\tilde{\mathbf{q}}'_j]_\times \mathbf{g}_\mathcal{F} \boldsymbol{\ell}_j^\mathsf{T} \mathcal{E}_\lambda \boldsymbol{\delta} + \mathsf{S}[\tilde{\mathbf{q}}'_j]_\times \bar{\mathcal{G}}_\mathcal{F} \tilde{\mathsf{P}}^\mathsf{T} \mathcal{E}_\lambda^\mathsf{T} \boldsymbol{\ell}_j \right\|^2.$$

DP-Warps. The minimum number of data points is $m \geq 4$. In the *arbitrary-centre* case, $\vartheta(\mathcal{W}_{\text{DP}}, \check{\mathsf{P}}')$ is an NLS problem, due to the division required for finding the affine coordinates of the transferred point. An initial estimate is found, as for the RP-Warps, by minimizing an algebraic approximation to the transfer error, *i.e.* $\min_{\check{\mathsf{P}}'} \sum_{j=1}^{m} d_a^2(\check{\mathcal{W}}_{\text{DP}}(\mathbf{q}_j; \check{\mathsf{P}}', \lambda), \tilde{\mathbf{q}}'_j)$. The arbitrary scale of $\check{\mathsf{P}}'$ is fixed by enforcing its norm to unity, *i.e.* $\|\check{\mathsf{P}}'\| = 1$. Replacing d_a by its expression, and $\check{\mathcal{W}}_{\text{DP}}$ by its homogeneous expression (13), we obtain, after some minor algebraic manipulations, an LLS problem:

$$\min_{\check{\mathsf{P}}', \|\check{\mathsf{P}}'\|=1} \sum_{j=1}^{m} \left\| \mathsf{S}[\tilde{\mathbf{q}}'_j]_\times \mathrm{diag}_3(\boldsymbol{\ell}_j^\mathsf{T} \mathcal{E}_\lambda) \mathrm{vect}(\check{\mathsf{P}}') \right\|^2,$$

with $\mathrm{diag}_r(\mathbf{x})$ an r block diagonal matrix with $\mathbf{x}$ the repeated block, and with vect the row-wise matrix vectorization.

[5] In practice, we directly optimize over the 12 entries of the second projection matrix to avoid parameterizing the nontrivial variety of fundamental matrices.

The simple centre-on-data setting in not possible for this type of warps. Indeed, as already discussed, the $2l$ constraints provided in general by l centre correspondences are not enough to constrain the $3l-1$ degrees of freedom of the warp. In other words, there is no solution to estimate the warp parameters by taking as centres the whole set of point correspondences: $l-1$ other point correspondences are needed. We thus consider a *weak centre-on-data* case: we pick a subset of l out of the m data points as centres, with $l \leq \lfloor \frac{2m+1}{3} \rfloor$. In case where $3l = 2m+1$ holds, there is a unique solution, which obviously depends on which data points are chosen as centres. The general minimal case is $m = 3k+1$ and $l = 2k+1$ with $k \in \mathbb{R}^{+*}$. Both the minimal and redundent cases are solved by writing an algebraic approximation to the transfer error. Given the centre coordinates P and P' in both images, we are looking for the parameters $\check{\mathsf{P}}'$ of the DP-Warp. A simple way of parameterizing the problem is to use $\check{\mathsf{P}}' = \mathrm{diag}(\mathbf{d})\tilde{\mathsf{P}}'$, and compute the 'scale vector' $\mathbf{d}$ only. This enforces interpolation of the centres and leaves only the remaining $l-1$ unknowns since $\mathbf{d}$ is an l-vector defined up to scale, leading to $\min_{\mathbf{d}} \sum_{j=1}^{m} d_a^2(\check{\mathcal{W}}_{\mathrm{DP}}(\mathbf{q}_j; \mathrm{diag}(\mathbf{d})\tilde{\mathsf{P}}', \lambda), \tilde{\mathbf{q}}'_j)$ such that $\|\mathbf{d}\| = 1$. Substituting d_a by its expression, as well as the DP-Warp from equation (13), gives an LLS problem:

$$\min_{\mathbf{d}, \|\mathbf{d}\|=1} \sum_{j=1}^{m} \left\| \mathsf{S}[\tilde{\mathbf{q}}'_j]_\times \tilde{\mathsf{P}}'^{\mathsf{T}} \mathrm{diag}(\mathcal{E}_\lambda^{\mathsf{T}} \boldsymbol{\ell}_j)\mathbf{d} \right\|^2 .$$

8. Experimental Evaluation

8.1. Simulated Data

We synthetically generated training and test data sets by projecting 3D points into two cameras with focal length f. The 3D points lie on paper-like 3D surfaces, *i.e.* with vanishing Gaussian curvature. We believe that it is representative of the kind of real images one may use the proposed warps on. The 50, Gaussian noise corrupted training data points are used as centres to estimate the warps. The errors reported below are means over 500 trials of the transfer error estimated over the 200 points of the test set, so as to reflect the quality of the estimated warps.

Figure 4 (left) shows the results we obtained. We must not compare the affine and perspective warps with these results, since in order to assess the influence of noise only, we use affine cameras for the affine warps, and perspective cameras with $f = 300$ pixels for the perspective warps. We observe that the quality of the warps linearly degrades with the noise level, and is actually around twice the noise standard deviation. This is a satisfying behaviour, holding true for all warps.

We gradually increase the focal length of the cameras, which has the effect of making the images more affine and keep the scene rigid. In order to preserve the transfer error scale, we keep invariant the size of the imaged object by translating each camera along its optical axis. Figure 4 (right) shows the result, with a 2 pixel noise level. As expected, the transfer error for the RP-Warps appears to be invariant to the focal length (the slight variations are due to the image scale). We make the same observation for the DA-Warps. This is due to the fact that the DA-Warps have extra degree of freedom compared to the rigid warps approximating the perspective projection of the rigid surface. The RA-Warps have the highest error for short focal lengths, which dramatically decreases as the focal length increases. It eventually converges to the same error as for the RP-Warps when the affine camera model becomes numerically equivalent to the perspective one. We see that beyond $f_{\mathrm{DA}} = 2000$ pixels, the RA-Warps do better than the DA-Warps. This is the breakdown focal length for the DA-Warps, above which the lack of rigidity, and the fact that the affine approximation is better fulfilled, makes the RA-Warps better capture the underlying true warp. We note that 2000 pixels is the order of magnitude one may have with real images. We experimentally measured $f_{\mathrm{DA}} \approx \{\infty, 4000, 2000, 500, 300\}$ pixels for $\{0, 1, 2, 5, 10\}$ pixel noise levels. This shows that enforcing the rigidity constraint is very important to capture a warp as close as possible to the true one from limited image measurements. Another conclusion is due to the significant difference between the RA-Warps and the RP-Warps. The latter achieves consistently lower test transfer errors, much lower for short to medium focal lengths. This shows that much more accurate warps are captured by modeling perspective projection. Experimental results with a deforming surface, not shown here, allow us to draw similar conclusions for the DA-Warps and the DP-Warps. It also shows that the DP-Warps overfit the data and usually have large variance. Another experiment shows that gradually merging the test and the training sets decreases the error to zero for the deformable warps and to the epipolar geometry fitting error for the rigid warps, as expected.

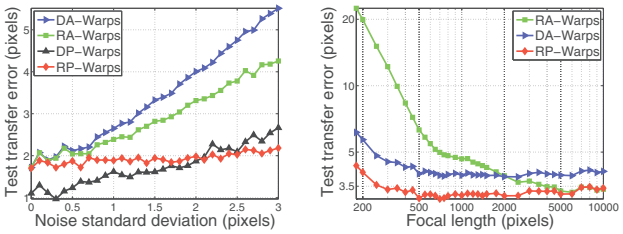

Figure 4. Influence of the noise level (left) and the focal length (right) onto the quality of the warps.

8.2. Real Data

We took a set of images of a manually handled poster with short and long focal lengths and various deformations.

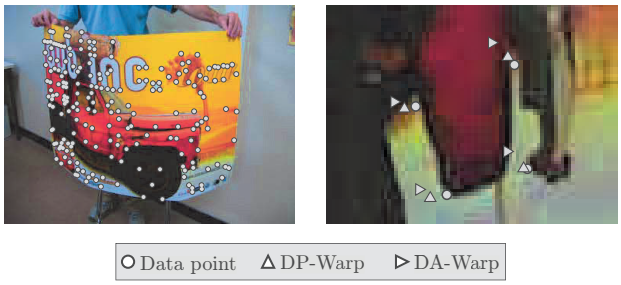

Figure 5. One of the perspective images (left) and closeup on transferred points in another image for the deformable warps (right).

Figure 6. Negative difference image for an RA-Warp (left) and an RP-Warp (right). Bright colors indicate low discrepancy. The RMS errors are respectively 33.66 and 19.01.

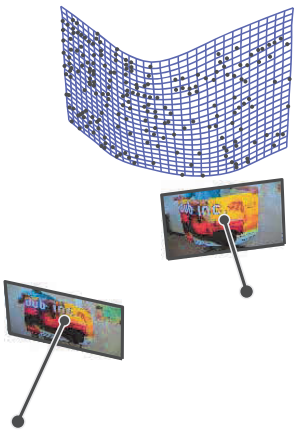

Figure 7. The cameras and 3D surface reconstructed through an RP-Warp.

We then manually clicked 206 point correspondences into all the images and estimated the warps. One of the experiments consisted in comparing the RA-Warps and the RP-Warps in the presence of significant perspective projection effects, see one of the images, overlaid with the point correspondences, on figure 5 (left) but *without surface deformations*. We selected 52 points to serve as centres (approximately 25% of the data points), and minimized the transfer error over all data points to fit the warps. The transfer error we obtained is 19.01 pixels and 12.53 pixels for the RA-Warp and the RP-Warp respectively. We used the two computed warps to warp the second image onto the first one, and computed their difference, shown on figure 6. These ideally are black, zero value images. Non-zero values are to be interpreted as unmodeled physical phenomena, mainly the extent to which the warp models the image deformation. The difference image clearly reflects the quality of the warp. We see that the RP-Warps do much better than the RA-Warps. The mean color alignment error in pixel value units is 33.66 and 19.01 for the RA-Warp and the RP-Warp respectively. We observed in particular that there is no data points in the top-right hand corner of the poster. This is were the difference is the highest for the RA-Warp, showing that the deformation, and thus the surface, is not very well captured by this model. The 3D surface reconstructed by the RP-Warp is shown in figure 7.

Similar comparison results for the DA-Warps and the DP-Warps were obtained by using two images of the poster with different deformations, giving a transfer error of 6.66 and 5.83 pixels respectively. Figure 5 (right) shows a representative closeup on transferred points. We observe that the data points are better predicted by the DP-Warp, meaning that it effectively models perspective projection but has however a high variance, *i.e.* is very dependent on which data points are used.

9. Discussion

Three types of $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ image warps are proposed, using the $\mathbb{R}^2 \rightarrow \mathbb{R}$ Thin-Plate Spline as a building block. They are designed to overcome some limitations of the standard Thin-Plate Spline warps, and derived based on a feature-driven parameterization we introduce. These warps have direct practical impacts since they better model image deformations than the standard DA-Warps in several cases, *e.g.* for rigid smooth surfaces and images with perspective projection effects. The DP-Warps are unstable because they overfit the data and tend to have high variance since they actually depend on the depth of the centres. One remedy may be to regularize their denominator. The warps essentially use two view visual geometry and the Thin-Plate Spline. Since they are given geometric interpretations, they can probably be extended to multiple views.

References

[1] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *PAMI*, 1989. 1, 2, 6

[2] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *CVIU*, 2003. 2

[3] G. Donato and S. Belongie. Approximate thin-plate spline mappings. *ECCV*, 2002. 2

[4] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003. 3, 6

[5] L. Masson, M. Dhome, and F. Jurie. Tracking 3D objects using flexible models. *BMVC*, 2005. 1

[6] T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. SIGGRAPH, 1986. 2

[7] D. Terzopoulos. Multilevel computational processes for visual surface reconstruction. *CVGIP*, 1983. 2

[8] J. Wills and S. Belongie. A feature-based approach for determining dense long range correspondences. In *ECCV*, 2004. 1