

A Closed-Form Solution to Local Non-Rigid Structure-from-Motion

Shaifali Parashar, Yuxuan Long, Mathieu Salzmann and Pascal Fua
 CVLAB, EPFL, Switzerland

shaifali.parashar,yuxuan.long,mathieu.salzmann,pascal.fua@epfl.ch

Abstract

A recent trend in Non-Rigid Structure-from-Motion (NRSfM) is to express local, differential constraints between pairs of images, from which the surface normal at any point can be obtained by solving a system of polynomial equations. The systems of equations derived in previous work, however, are of high degree, having up to five real solutions, thus requiring a computationally expensive strategy to select a unique solution. Furthermore, they suffer from degeneracies that make the resulting estimates unreliable, without any mechanism to identify this situation.

In this paper, we show that, under widely applicable assumptions, we can derive a new system of equation in terms of the surface normals whose two solutions can be obtained in closed-form and can easily be disambiguated locally. Our formalism further allows us to assess how reliable the estimated local normals are and, hence, to discard them if they are not. Our experiments show that our reconstructions, obtained from two or more views, are significantly more accurate than those of state-of-the-art methods, while also being faster.

1. Introduction

Reconstructing the 3D shape of deformable objects from monocular image sequences is known as Non-Rigid Structure-from-Motion (NRSfM) and has applications in domains ranging from entertainment [19] to medicine [14]. Early methods relied on low-rank representations of the surfaces [4, 5, 1, 28, 8, 16, 10] while more recent ones exploit local surface properties to derive local constraints and can handle larger deformations [29, 30, 27, 6, 7, 11]. Unfortunately, these constraints have to be enforced jointly over the entire point set and the whole sequence. Hence, the computational cost increases non-linearly with the number of images and quickly becomes prohibitive. Furthermore, these *global* methods cannot handle missing data.

The *local* methods of [20, 21, 22] offer an interesting alternative. Expressing isometry, conformality or equiareality constraints in terms of differential properties makes

it possible to derive systems of polynomial equations that can be solved to infer depth derivatives. Hence, surface normals can be obtained in any frame by pairing it with all others and the complexity only grows linearly with the number of images because the number of local variables remains fixed. Unfortunately, the systems of equations that arise in this computation are bivariate of high degree. They can have up to five real solutions. In theory, a unique solution can be obtained from 3 images but this requires either a complex sum-of-squares formulation or complex reduction methodologies that add phantom solutions. Hence, in practice, these formulations then require more than 3 images to obtain reliable estimates. Furthermore, when the motion between frames is too small, the system becomes ill-posed and the estimates unreliable, without any mechanism to flag such situations as problematic.

In this paper, we use the same philosophy as in [20, 21] but, instead of expressing the constraints on depth derivatives, we express them in terms of surface normals. Specifically, given a 2D warp (registration) between two images, we consider the two tangent planes at corresponding points in these images. We use the homography relating these planes to compute the normals at these points by solving local differential constraints in closed form. This has two solutions, instead of five in the earlier approaches. For each plane and we pick the right one by enforcing an easy-to-compute measure of local smoothness. Furthermore, our formalism lets us assess how well-conditioned the problem was and, hence, how usable the resulting normals are. In other words, we can derive from an image pair a set of reliable normals and discard the others.

We will demonstrate on both synthetic data that we outperform state-of-the-art local and global methods at a fraction of the computational cost. Our contribution is therefore an approach to Non-Rigid Structure-from-Motion that relies on solving in closed form a set of equations relating surface normals at corresponding points. Being entirely local, the computation is both fast and reliable.

2. Related Work

NRSfM was introduced in [4] and the ill-posedness of the problem was handled by constraining the deformations to lie on a low-dimensional manifold. Later variants introduced additional constraints [5, 1, 10, 8, 16], or relied on machine learning to tune the dimensionality of the deformation space [13, 24]. However, the low-dimension assumption these methods all rely on precludes them from modeling complex deformations. Furthermore, they do not naturally handle missing data and occlusions, and complex formulations are required to overcome this limitation. As a result, these methods have been limited to objects that deform in a relatively predictable way, such as human faces. Recently, these fundamental limitations have been addressed by imposing constraints between corresponding points across images in one of the two following ways.

Modeling Global Deformations. Several methods seek to enforce physical surface properties, such as inextensibility [7, 11], piece-wise inextensibility [27, 30] or piece-wise rigidity [29]. A globally optimal solution is then found by solving for constraints over all corresponding points jointly. These methods usually require a computationally expensive optimization, which makes them impractical for handling large numbers of images.

Modeling Local Deformations. Other methods rely on local constraints. This makes it possible to solve for each point independently, reduces the computation cost, and enables the handling of missing data. The constraints can stem from assuming that distances are preserved [6, 20, 21] or even only from assuming that the deformations are locally diffeomorphic [22]. These methods tend to perform significantly better than their global counterparts but suffer from one key drawback: The local constraints are not always well-posed, leading to manyfold ambiguities or even degenerate solutions, without any mechanism for telling when this happens. This is the problem we address in this paper.

3. Formalism and Assumptions

At the heart of our approach is the fact that the normals at two different instants at a point on a deforming 3D surface can be computed given its projections in two images and a 2D warp between these images under the sole assumptions of local surface planarity and deformation local linearity. In this section, we first introduce the NRSfM setup we will use in the rest of this paper, which is similar to the one of [21]. We then explain what our assumptions mean and why they are widely applicable.

3.1. Setup

Fig. 1 depicts our setup when using only two images \mathcal{I} and $\bar{\mathcal{I}}$ acquired by a calibrated camera. In each one, we

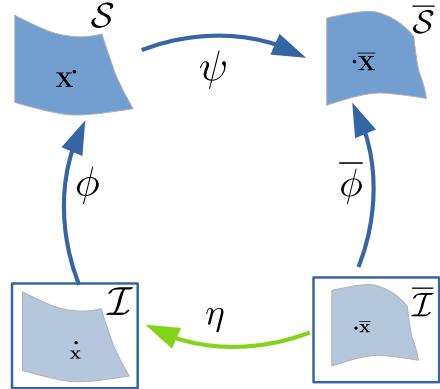


Figure 1: A 2-view model for Diff-NRSfM. Assuming ψ to be locally isometric, our goal is to find $\phi, \bar{\phi}$ given that η is known.

denote the deforming surface as \mathcal{S} and $\bar{\mathcal{S}}$, respectively, and model it in terms of functions $\phi, \bar{\phi} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ that associate a surface point to an image point. Let us assume that we are given an image registration function $\eta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ that associates points in the first image to points in the second. In practice, it can be computed using standard image matching techniques such as optical flow [26, 25] or SIFT [17]. These functions can be composed to create a mapping $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ between 3D surface points seen in the two images. We use a parametric representation of η and ϕ with B-splines [3], which allows us to accurately obtain first- and second-order derivatives of these functions. A finite-difference approach could also be used.

Given a point $\mathbf{x} = (u, v)$ on \mathcal{I} and its corresponding 3D point $\mathbf{X} = \phi(\mathbf{x})$ on \mathcal{S} , we write $\phi(\mathbf{x}) = \frac{1}{\beta(u,v)} (u \ v \ 1)^\top$, where β represents the inverse of the depth. The Jacobian of ϕ is given by

$$\mathbf{J}_\phi = \frac{1}{\beta(u,v)} \begin{pmatrix} 1 - uk_1 & -uk_2 \\ -vk_1 & 1 - vk_2 \\ -k_1 & -k_2 \end{pmatrix}, \quad (1)$$

where $k_1 = \frac{\partial_u \beta}{\beta}$ and $k_2 = \frac{\partial_v \beta}{\beta}$. $\bar{u}, \bar{v}, \bar{\phi}, \bar{k}_1$, and \bar{k}_2 are defined similarly in $\bar{\mathcal{I}}$.

3.2. Local Planarity and Linearity

In this work, we assume local planarity of the 3D surfaces and local linearity of the deformations as described in [15, 12]. We now describe these two assumptions and argue that they are weak ones that are generally applicable.

Surface Local Planarity. Let \mathbf{x}_0 be an image point with surface normal \mathbf{n} at $\phi(\mathbf{x}_0)$. All points $\mathbf{x} = (u, v)$ sufficiently close to \mathbf{x}_0 can be accurately described as lying on the tangent plane. Hence, they satisfy $\mathbf{n}^\top \phi(\mathbf{x}) + d = 0$, where d is a scalar, which we can rewrite as $\beta =$

$-\frac{\mathbf{n}^\top}{d} (u \ v \ 1)^\top$. Therefore the inverse depth β that appears in Eq. 1 is a linear function of \mathbf{x} even though ϕ is not. Nevertheless, all higher-order derivatives of ϕ can be expressed in terms of β and its first order derivatives. This is widely viewed as a weak assumption that applies to most smooth manifolds [15]. For example, our planet is a sphere that can be treated as locally planar.

Deformation Local Linearity. We assume the deformation ψ that relates locally two planes to be smooth enough to be well described locally by its first order approximation, so that we can ignore its second derivatives. In other words, we use a first order approximation for the local deformations but a second order one for the surface depth to allow for globally non-planar shapes. This is a looser set of assumptions than what is normally used in NRSfM. For example, [16, 8] and other low-rank methods assume the deformation space to be of low rank. Physics-based methods that use inextensibility [7, 30] or piecewise-rigidity [29, 27] make a much stronger assumption.

3.3. Differential Constraints across Images

To express constraints between values in \mathcal{I} and $\bar{\mathcal{I}}$, we define *metric tensors* and *connections* as described in [15].

Metric Tensors. The metric tensors \mathbf{g} in \mathcal{I} and $\bar{\mathbf{g}}$ in $\bar{\mathcal{I}}$ are first-order differential quantities that capture local distances and angles. They can be written as

$$\mathbf{g} = \mathbf{J}_\phi^\top \mathbf{J}_\phi \text{ and } \bar{\mathbf{g}} = \mathbf{J}_{\bar{\phi}}^\top \mathbf{J}_{\bar{\phi}}. \quad (2)$$

These tensors can be used to impose isometric, conformal, and equiareal constraints by forcing the scalars k_1 and k_2 of Eq. 1 to satisfy one of the three conditions below:

$$\begin{cases} \bar{\mathbf{g}} = \mathbf{J}_\eta^\top \mathbf{g} \mathbf{J}_\eta, & \text{Isometry} \\ \bar{\mathbf{g}} = \lambda^2 \mathbf{J}_\eta^\top \mathbf{g} \mathbf{J}_\eta, \lambda^2 \in \mathbb{R}^+ - \{1\}, & \text{Conformality} \\ \sqrt{\det(\bar{\mathbf{g}})} = \sqrt{\det(\mathbf{J}_\eta^\top \mathbf{g} \mathbf{J}_\eta)}. & \text{Equiareality} \end{cases} \quad (3)$$

Connections. Connections are second-order differential quantities that express the rate of change of the metric tensors. Under the assumption of local linearity, it can be shown [21] that

$$\left(\begin{array}{c} \bar{k}_1 \\ \bar{k}_2 \end{array} \right) = \mathbf{J}_\eta^\top \left(\begin{array}{c} k_1 \\ k_2 \end{array} \right) - \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \mathbf{J}_\eta^{-1} \frac{\partial^2 \eta}{\partial \bar{u} \partial \bar{v}}. \quad (4)$$

Solutions to isometric, conformal and equiareal NRSfM can be obtained by solving the metric tensor preservation equations in Eq. 3 under the constraint of Eq. 4.

4. Computing Normals from Two Images

In earlier approaches [21], the NRSfM problem was addressed by solving the system of Eq. 3 under the isometric, conformal, and equiareal constraints of Eq. 3 with respect to the variables k_1 and k_2 of Eq. 1. Here, we solve this system of equations directly in terms of the surface normals. We will show that, not only can this be done in closed form, but it also allows us to identify degenerate situations that result in unreliable estimates.

Differentiating the Warp. Let us consider a point $\bar{\mathbf{x}} = (\bar{u}, \bar{v})^\top$ in $\bar{\mathcal{I}}$ and its corresponding point $(u, v)^\top = \eta(\bar{u}, \bar{v})$ in \mathcal{I} . Assuming the surfaces to be locally planar means that there is a 3×3 homography matrix $\mathbf{H} = [h_{ij}]_{1 \leq i,j \leq 3}$ relating $\bar{\mathbf{x}}$ and \mathbf{x} . We can write

$$\left(\begin{array}{c} u \\ v \\ 1 \end{array} \right) = \frac{1}{s} \left(\begin{array}{ccc} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{array} \right) \left(\begin{array}{c} \bar{u} \\ \bar{v} \\ 1 \end{array} \right), \quad (5)$$

where $\bar{s} = h_{31}\bar{u} + h_{32}\bar{v} + h_{33}$. Thus the first and second-order derivatives of η can be computed as

$$\begin{aligned} \mathbf{J}_\eta &= \left(\begin{array}{cc} \frac{\partial \eta}{\partial \bar{u}} & \frac{\partial \eta}{\partial \bar{v}} \end{array} \right) = \frac{1}{s} \left(\begin{array}{cc} h_{11} - h_{31}u & h_{12} - h_{32}u \\ h_{21} - h_{31}v & h_{22} - h_{32}v \end{array} \right), \\ \left(\begin{array}{ccc} \frac{\partial^2 \eta}{\partial \bar{u}^2} & \frac{\partial^2 \eta}{\partial \bar{u} \partial \bar{v}} & \frac{\partial^2 \eta}{\partial \bar{v}^2} \end{array} \right) &= -\frac{1}{s} \mathbf{J}_\eta \left(\begin{array}{ccc} 2h_{31} & h_{32} & 0 \\ 0 & h_{31} & 2h_{32} \end{array} \right). \end{aligned} \quad (6)$$

Image Embedding and Local Normal. The unit normal \mathbf{n} at \mathbf{x} is the cross product of the columns of the matrix \mathbf{J}_ϕ from Eq. 1. This lets us write

$$\mathbf{n} = \frac{1}{\beta^2 \sqrt{\det \mathbf{g}}} \begin{pmatrix} k_1 \\ k_2 \\ 1 - uk_1 - vk_2 \end{pmatrix} \quad (7)$$

$$\begin{aligned} &= \frac{1}{\beta^2 \sqrt{\det \mathbf{g}}} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ -\mathbf{x}^\top & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ 1 \end{pmatrix}. \\ &\Rightarrow \begin{pmatrix} k_1 \\ k_2 \\ 1 \end{pmatrix} = \beta^2 \sqrt{\det \mathbf{g}} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ \mathbf{x}^\top & 1 \end{pmatrix} \mathbf{n}. \end{aligned} \quad (8)$$

Given the normal \mathbf{n} as written in Eq. 7, we rewrite the matrix \mathbf{J}_ϕ of Eq. 1 as

$$\begin{aligned} \mathbf{J}_\phi &= \frac{1}{\beta} \begin{pmatrix} 0 & uk_1 + vk_2 - 1 & k_2 \\ 1 - uk_1 - vk_2 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ v & -u \end{pmatrix} \\ &= \beta \sqrt{\det \mathbf{g}} [\mathbf{n}]_\times \mathbf{E}. \end{aligned} \quad (9)$$

We can now rewrite the differential constraints across images introduced in Section 3.3 in terms of the normals.

Connections. Given the η derivatives from Eq. 6, the connections transfer relation of Eq. 4 becomes

$$\begin{aligned} \begin{pmatrix} \bar{k}_1 \\ \bar{k}_2 \end{pmatrix} &= \mathbf{J}_\eta^\top \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} + \frac{1}{\bar{s}} \begin{pmatrix} h_{31} \\ h_{32} \end{pmatrix}, \\ \Rightarrow \begin{pmatrix} \bar{k}_1 \\ \bar{k}_2 \\ 1 \end{pmatrix} &= \begin{pmatrix} \mathbf{J}_\eta^\top & \mathbf{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ 1 \end{pmatrix}. \end{aligned} \quad (10)$$

Using Eq. 8, we rewrite the above expression as

$$\begin{aligned} \bar{\mathbf{n}} &= \frac{\beta^2}{\bar{\beta}^2} \sqrt{\frac{\det \mathbf{g}}{\det \bar{\mathbf{g}}}} \mathbf{T} \mathbf{n} \\ &= \frac{\beta^2}{\bar{\beta}^2} \sqrt{\frac{\det \mathbf{g}}{\det \bar{\mathbf{g}}}} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ -\bar{\mathbf{x}}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{J}_\eta^\top & \mathbf{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ \mathbf{x}^\top & 1 \end{pmatrix} \mathbf{n} \\ &= \frac{\beta^2}{\bar{s}\bar{\beta}^2} \sqrt{\frac{\det \mathbf{g}}{\det \bar{\mathbf{g}}}} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ -\bar{\mathbf{x}}^\top & 1 \end{pmatrix} \begin{pmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{32} \\ \bar{s}u & \bar{s}v & \bar{s} \end{pmatrix} \mathbf{n} \\ &= \frac{\beta^2}{\bar{s}\bar{\beta}^2} \sqrt{\frac{\det \mathbf{g}}{\det \bar{\mathbf{g}}}} \mathbf{H}^\top \mathbf{n}, \end{aligned} \quad (11)$$

which directly relates the two normals.

Metric Tensor. As shown in Fig. 1, we can write $\bar{\phi} = \psi \circ \phi \circ \eta$. Differentiating this expression and multiplying it by its transpose yields

$$\bar{\mathbf{g}} = \mathbf{J}_\phi^\top \mathbf{J}_{\bar{\phi}} = \mathbf{J}_\eta^\top \mathbf{J}_\phi^\top \mathbf{J}_\psi^\top \mathbf{J}_\psi \mathbf{J}_\phi \mathbf{J}_\eta. \quad (12)$$

Using Eq. 9, we write $\mathbf{J}_\phi \mathbf{J}_\eta = \beta \sqrt{\det g} [\mathbf{n}]_\times \mathbf{E} \mathbf{J}_\eta$. Given the η derivatives of Eq. 6, we simplify $\mathbf{E} \mathbf{J}_\eta$ to $\frac{1}{\bar{s}} (\mathbf{h}_1 \times \hat{\mathbf{x}} \quad \mathbf{h}_2 \times \hat{\mathbf{x}})$, where $\mathbf{h}_1, \mathbf{h}_2$ are the first two columns of the homography matrix \mathbf{H} , and $\hat{\mathbf{x}} = (u \ v \ 1)^\top$. By writing $\mathbf{z}_1 = \mathbf{n} \times (\mathbf{h}_1 \times \hat{\mathbf{x}})$ and $\mathbf{z}_2 = \mathbf{n} \times (\mathbf{h}_2 \times \hat{\mathbf{x}})$, Eq. 3 reduces to

$$\begin{cases} \bar{\mathbf{g}} = \frac{\lambda^2 \beta^2 \det(\mathbf{g})}{\bar{s}^2} \begin{pmatrix} \mathbf{z}_1^\top \mathbf{z}_1 & \mathbf{z}_1^\top \mathbf{z}_2 \\ \mathbf{z}_1^\top \mathbf{z}_2 & \mathbf{z}_2^\top \mathbf{z}_2 \end{pmatrix}, \\ \sqrt{\det(\bar{\mathbf{g}})} = \|[\mathbf{J}_{\bar{\phi}}]_1 \times [\mathbf{J}_{\bar{\phi}}]_2\| = \sqrt{\det(\mathbf{J}_\eta^\top \mathbf{g} \mathbf{J}_\eta)}, \end{cases} \quad (13)$$

Here, $[\mathbf{J}_{\bar{\phi}}]_1$ and $[\mathbf{J}_{\bar{\phi}}]_2$ are the columns of $\mathbf{J}_{\bar{\phi}}$. We will use this definition of area to draw equiareal constraints.

NRSfM from isometric/conformal constraints. So far, we have expressed the metric preservation conditions in terms of the normals of the two surfaces under consideration. The only unknown left in the system is therefore \mathbf{n} . We now show that this unknown can in fact be computed in closed form.

Given the multiplicative nature of the cross product, the constraints on the normals of Eq. 11 imply that

$$[\bar{\mathbf{n}}]_\times = \frac{\beta^2}{\bar{s}\bar{\beta}^2} \sqrt{\frac{\det \mathbf{g}}{\det \bar{\mathbf{g}}}} \det(\mathbf{H}^\top) \mathbf{H}^{-1} [\mathbf{n}]_\times \mathbf{H}^{-\top}. \quad (14)$$

This lets us rewrite the matrix $\mathbf{J}_{\bar{\phi}}$ of Eq. 9 as

$$\begin{aligned} \mathbf{J}_{\bar{\phi}} &= \frac{\beta^2}{\bar{\beta}} \sqrt{\det \mathbf{g}} \mathbf{H}^{-1} [\mathbf{n}]_\times \left(\frac{\det \mathbf{H}^\top}{\bar{s}} \mathbf{H}^{-\top} \bar{\mathbf{E}} \right) \\ &= \frac{\beta^2 \sqrt{\det \mathbf{g}}}{\bar{\beta}} \mathbf{H}^{-1} [\mathbf{n}]_\times (\mathbf{h}_1 \times \hat{\mathbf{x}} \quad \mathbf{h}_2 \times \hat{\mathbf{x}}) \\ &= \frac{\beta^2 \sqrt{\det \mathbf{g}}}{\bar{\beta}} \mathbf{H}^{-1} (\mathbf{z}_1 \quad \mathbf{z}_2). \end{aligned} \quad (15)$$

Injecting this expression into the isometric/conformal metric tensor preservation relation of Eq. 13 yields

$$\begin{aligned} \begin{pmatrix} \mathbf{z}_1^\top \mathbf{H}^{-\top} \mathbf{H}^{-1} \mathbf{z}_1 & \mathbf{z}_1^\top \mathbf{H}^{-\top} \mathbf{H}^{-1} \mathbf{z}_2 \\ \mathbf{z}_1^\top \mathbf{H}^{-\top} \mathbf{H}^{-1} \mathbf{z}_2 & \mathbf{z}_2^\top \mathbf{H}^{-\top} \mathbf{H}^{-1} \mathbf{z}_2 \end{pmatrix} &= \frac{\lambda^2 \bar{\beta}^2}{\bar{s}^2 \beta^2} \begin{pmatrix} \mathbf{z}_1^\top \mathbf{z}_1 & \mathbf{z}_1^\top \mathbf{z}_2 \\ \mathbf{z}_1^\top \mathbf{z}_2 & \mathbf{z}_2^\top \mathbf{z}_2 \end{pmatrix}, \\ \Rightarrow \mathbf{z}_i^\top \left(\bar{\mathbf{H}}^\top \bar{\mathbf{H}} - \frac{\lambda^2 \bar{\beta}^2}{\bar{s}^2 \beta^2} \mathbf{I}_{3 \times 3} \right) \mathbf{z}_j &= 0, \forall i, j \in \{1, 2\}, \end{aligned} \quad (16)$$

where $\bar{\mathbf{H}} = \mathbf{H}^{-1}$. Assuming \mathbf{H} to be normalized, that is, its second singular value to be 1, the relation between a 3D point observed in the two input images is given by $\phi(\mathbf{x}) = \mathbf{H} \phi(\bar{\mathbf{x}})$. Using Eq. 5 yields $\bar{\beta} = \beta \bar{s}$. By writing $\mathbf{z}_i = [\mathbf{n}]_\times [\mathbf{h}_i]_\times \hat{\mathbf{x}}$, the above constraints further simplify to

$$[\mathbf{n}]_\times^\top (\bar{\mathbf{H}}^\top \bar{\mathbf{H}} - \lambda^2 \mathbf{I}_{3 \times 3}) [\mathbf{n}]_\times = 0. \quad (17)$$

Since $\mathbf{H} \sim \alpha \mathbf{H}$, we divide the above expression by λ^2 and, with a slight abuse of notation, write $\frac{1}{\lambda} \bar{\mathbf{H}}$ as $\bar{\mathbf{H}}$. This simplifies the above expressions to

$$[\mathbf{n}]_\times^\top (\bar{\mathbf{H}}^\top \bar{\mathbf{H}} - \mathbf{I}_{3 \times 3}) [\mathbf{n}]_\times = [\mathbf{n}]_\times^\top \mathbf{S} [\mathbf{n}]_\times = 0. \quad (18)$$

Degenerate Cases. The system of Eq. 18 holds as long as \mathbf{S} is a non-null matrix, which means $\bar{\mathbf{H}}^\top \bar{\mathbf{H}} \neq \mathbf{I}_{3 \times 3}$. Therefore, $\bar{\mathbf{H}}$ should not be an orthogonal matrix which makes the pure rotations and reflections to cause degeneracies.

Affine Stability. Under affine imaging conditions, $h_{31} = h_{32} = 0$, and $h_{33} = 1$. In this case, \mathbf{z}_i and \mathbf{S} remain non-null, and thus the system in Eq. 18 does not become degenerate, and we can still compute the normal.

Solution. The solution to the system in Eq. 18 can be obtained by homography decomposition [18]. We give an overview of the solution here but recommend reading [18] for more detail.

$\mathbf{S} = s_{ij}$ is a symmetric matrix expressed in terms of $\bar{\mathbf{H}}$. It can be numerically computed using η and image observations $(\mathbf{x}, \bar{\mathbf{x}})$. Eq. 11 shows a closed-form definition of $\mathbf{H}^\top = \mathbf{T} = \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ -\bar{\mathbf{x}}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{J}_\eta^\top & \mathbf{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ \mathbf{x}^\top & 1 \end{pmatrix}$. Let us write $\mathbf{n} = (n_1 \ n_2 \ n_3)^\top$. Since $n_3 \neq 0$, we define $y_1 = \frac{n_1}{n_3}$ and $y_2 = \frac{n_2}{n_3}$ and expand the system in Eq. 18

accordingly. This yields 6 constraints, out of which only 3 are unique. They are given by

$$\begin{aligned} s_{33}y_2^2 - 2s_{23}y_2 + s_{22} &= 0, \\ s_{33}y_1^2 - 2s_{13}y_1 + s_{11} &= 0, \\ s_{22}y_1^2 - 2s_{12}y_1y_2 + s_{11}y_2^2 &= 0. \end{aligned} \quad (19)$$

By solving the first two, we obtain $y_1 = \frac{s_{13} \pm \sqrt{s_{13}^2 - s_{33}s_{11}}}{s_{33}}$ and $y_2 = \frac{s_{23} \pm \sqrt{s_{23}^2 - s_{33}s_{22}}}{s_{33}}$.

We use the third expression to disambiguate the solutions. Ultimately, this gives us a closed-form expression for the pair of normals, written as

$$\begin{aligned} \mathbf{n}_a &= (s_{13} + s\sqrt{s_{13}^2 - s_{33}s_{11}} \quad s_{23} + \sqrt{s_{23}^2 - s_{33}s_{22}} \quad s_{33})^\top, \\ \mathbf{n}_b &= (s_{13} - s\sqrt{s_{13}^2 - s_{33}s_{11}} \quad s_{23} - \sqrt{s_{23}^2 - s_{33}s_{22}} \quad s_{33})^\top, \end{aligned}$$

where $s = \text{sign}(s_{23}s_{13} - s_{12}s_{33})$. (20)

Normal Selection. Using Eq. 8, the local depth derivatives (k_1, k_2) at \mathbf{X} are given by $k_i = \frac{n_i}{un_1 + vn_2 + n_3}$. From the solution in Eq. 20, we thus obtain two possible solutions for the local depth derivatives (k_{1a}, k_{2a}) and (k_{1b}, k_{2b}) . We pick the normal that minimizes the corresponding sum of squares of depth derivatives. That is, we compute the normal \mathbf{n} as

$$\mathbf{n} = \begin{cases} \mathbf{n}_a & \text{if } k_{1a}^2 + k_{2a}^2 \leq k_{1b}^2 + k_{2b}^2 \\ \mathbf{n}_b & \text{otherwise} \end{cases} \quad (21)$$

Following Eq. 5, $\bar{\mathbf{n}}$ is then obtained as $\mathbf{H}^\top \mathbf{n}$.

Measure of Degeneracy. In degenerate situations, the singular values $(\sigma_1, \sigma_2, \sigma_3)$ of \mathbf{H} are all one. We use the ratio $\frac{\sigma_1}{\sigma_3}$ to quantify the degeneracy. Thus, we only reconstruct from \mathbf{S} if $\frac{\sigma_1}{\sigma_3} > \tau$, and we set $\tau = 1.2$. We also put an upper limit $\frac{\sigma_1}{\sigma_3} = 10$ to ensure that the data is well-conditioned, as we observed that a high value of τ leads to inaccuracies caused by measurement errors on η .

Equiareality. We write the area-preservation constraint of Eq. 13 using the \mathbf{J}_ϕ and $\mathbf{J}_{-\phi}$ given in Eqs. 9 and 15. The constraint is given by

$$\begin{aligned} (\mathbf{z}_1 \times \mathbf{z}_2)^\top \frac{\beta^4 \det \mathbf{g}}{\bar{\beta}^2 (\det \mathbf{H})^2} \mathbf{H} \mathbf{H}^\top (\mathbf{z}_1 \times \mathbf{z}_2) \\ = (\mathbf{z}_1 \times \mathbf{z}_2)^\top \frac{\beta^2 \det \mathbf{g}}{\bar{s}^2} (\mathbf{z}_1 \times \mathbf{z}_2) \end{aligned} \quad (22)$$

Assuming that the homography is normalized, we write $\bar{\beta} = \bar{s}\beta$ and $\mathbf{H} \sim \frac{\mathbf{H}}{\det \mathbf{H}}$. The above equation is simplified to

$$(\mathbf{z}_1 \times \mathbf{z}_2)^\top \mathbf{H} \mathbf{H}^\top (\mathbf{z}_1 \times \mathbf{z}_2) = (\mathbf{z}_1 \times \mathbf{z}_2)^\top (\mathbf{z}_1 \times \mathbf{z}_2). \quad (23)$$

Here, $\mathbf{z}_i = \mathbf{n} \times (\mathbf{h}_i \times \hat{\mathbf{x}})$. We write $\hat{\mathbf{h}}_i = \mathbf{h} \times \hat{\mathbf{x}}$, and the cross product is given by $\mathbf{z}_1 \times \mathbf{z}_2 = (\mathbf{n} \cdot (\hat{\mathbf{h}}_1 \times \hat{\mathbf{h}}_2))\mathbf{n}$. Since

$\mathbf{h}_1, \mathbf{h}_2$ are linearly independent, $\mathbf{n} \cdot (\hat{\mathbf{h}}_1 \times \hat{\mathbf{h}}_2)$ is a non-zero scalar. Therefore, the area-preservation constraint is given by

$$\mathbf{n}^\top (\mathbf{H} \mathbf{H}^\top - \mathbf{I}_{3 \times 3}) \mathbf{n} = 0. \quad (24)$$

This constraint is a consequence to isometric/conformal deformations in Eq. 18. Given that \mathbf{H} is non-orthogonal both $\mathbf{n}_a, \mathbf{n}_b$ in Eq. 20 satisfy this relation.

Diffeomorphism. In a generic case of deformation where no metric quantity is preserved, we can assume that the areas are related by a scalar α . Therefore, the area-preservation constraint in Eq. 24 can be re-written as

$$\mathbf{n}^\top (\mathbf{H} \mathbf{H}^\top - \alpha \mathbf{I}_{3 \times 3}) \mathbf{n} = 0. \quad (25)$$

Since $\mathbf{H} \sim \alpha \mathbf{H}$, this equation is also a consequence to our isometric/conformal NRSfM solution.

To summarise, a local closed-form solution to surface normals can be obtained by solving Eq. 18. There may be more solutions to generic deformation constraints that satisfy Eq. 25, the surface normals obtained by solving Eq. 18 is one of them.

5. Computing Normals from Multiple Images

Methods such as those of [21, 22] pick a reference image and formulate reconstruction constraints between it and the other images, which are then solved by solving a least-squares problem over the entire set of images. We use the same strategy except that we reconstruct from image pairs, one of them being the reference image. Therefore, for N images, we obtain $N-1$ estimates for the reference image and 1 estimate for each of the non-reference images.

More formally, let $\{\mathbf{x}_j^i\}$, $i \in [1, M]$, $j \in [1, N]$, be a set of N point correspondences between M images. Our goal is to find the 3D point \mathbf{X}_j^i and the normal \mathbf{n}_j^i corresponding to each \mathbf{x}_j^i . Using Eq. 11, we write the local homography for each point correspondence \mathbf{H}_j^{ik} between image pairs $(i, k) \in [1, M]$, $i \neq k$, using the warp η . Each local homography \mathbf{H}_j^{ik} is normalized by dividing it by its second singular value. We compute \mathbf{Hc}_j^{ik} given by the ratio of the first and third singular value, and the normals for each local homography \mathbf{Hc}_j^{ik} using Eq. 20. We then pick a unique solution using Eq. 21. The solution on the reference and non-reference image is given by \mathbf{n}_j^{kk} and \mathbf{n}_j^{ki} , respectively. For non-degenerate cases, where $\frac{\sigma_1}{\sigma_3} \in [1.1, 10]$, we compute the normal \mathbf{n}_j^i by taking the mean of the \mathbf{n}_j^{ik} 's computed over k reference images. We obtain the 3D point \mathbf{X}_j^i by integrating the normals on each surface using a variant of the method described in [21]. We divide the surface into regions with similar normals and integrate them independently. This is faster and more accurate than performing a global computation. We can do this because our normals are much less noisy than those of [21]. We summarize our complete pipeline in Algorithm 1.

Algorithm 1: Our NRSfM Algorithm

Data: $\mathbf{x}_j^i, \mathbf{H}_j^{ik}$ and \mathbf{Hc}_j^{ik}

Result: \mathbf{n}_j^i

$$\frac{\sigma_1}{\sigma_3 \min} = 1.1, \frac{\sigma_1}{\sigma_3 \max} = 10;$$

for each reference image $k = [1, M]$ **do**

- for** each point $j = [1, N]$ **do**

 - for** images $i = [1, M], i \neq k$ **do**

 - if** $\mathbf{Hc}_j^{ik} > \frac{\sigma_1}{\sigma_3 \min}$ and $\mathbf{Hc}_j^{ik} < \frac{\sigma_1}{\sigma_3 \max}$
 - then**

 - Compute normals using (20);
 - Pick a solution \mathbf{n}_j^{kk} using (21);
 - Write $\mathbf{n}_j^{ik} = (\mathbf{Hc}_j^{ik})^\top \mathbf{n}_j^{kk}$;

 - else**

 - Set $\mathbf{n}_j^{ik}, \mathbf{n}_j^{kk}$ to zero;

 - end**

 - end**

- end**

for each point $j = [1, N]$ **do**

- for** images $i = [1, M]$ **do**

 - Obtain \mathbf{n}_j^i by averaging the non-zero normals in \mathbf{n}_j^{ik} ;

- end**

- end**

6. Experiments

We experiment on synthetic and real data to validate our proposed method, which we denote as **Ours**.

Baselines and Metrics. We compare our method to local linearity-based diffeomorphic NRSfM **Pa20** [22], jointly solving isometric/conformal NRSfM **Pa19** [21], local and piecewise homography decomposition, **Ch14** [6] and **Va09** [29], respectively. These are methods that, like ours, reconstruct local/piecewise surface normals and integrate them to obtain depth. The key idea behind integration is to obtain a smooth reconstruction by enforcing a local smoothness on the normals. As a consequence, this step improves the quality of reconstructed normals. The improvement is huge for **Ch14** and **Va09**, substantial for **Pa19** and **Pa20** and minor for our method. In this paper, to truly compare the NRSfM techniques themselves, we report the accuracy of the computed normals rather than the smoothed ones. We report errors in terms of accuracy of the normals En and 3D points Ed . En is computed as the average dot product between ground-truth and computed normals. Ed is the RMSE between the ground-truth and computed 3D points.

We also compare against two of the best global meth-

Table 1: **Synthetic experiments results.** 'X' indicates that the method does not return a result because we are not using enough images.

Method	Surfaces 1, 2		Surfaces 1,3		Surfaces 1,2,3	
	En	Ed	En	Ed	En	Ed
Lee16	X	X	X	X	X	X
An17	X	X	X	X	X	X
Va09	16.4	10.2	24.5	12.1	17.3	9.8
Ch14	X	X	X	X	28.1	20.2
Ch17	X	X	X	X	X	14.5
Pa19	X	X	X	X	17.4	4.3
Pa20	X	X	X	X	24.7	9.3
Ours	4.0	2.1	8.3	4.3	9.3	3.2

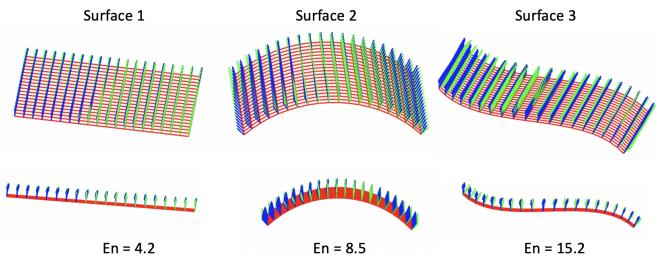


Figure 2: **Reconstructed normals.** A synthetic deforming surface reconstructed in three different frames. The predicted normals are shown in blue and the ground-truth ones in green.

ods, **Ch17** [7] and **Lee16** [16], along with a dense method, **An17** [2]. They directly return 3D points. Hence, we only report Ed for these methods.

Results on Synthetic Data. We created 3 smooth surfaces: a plane, a cylindrical surface and a stretched surface, as shown in Figure 2. Our results are the average of 10 trials with added gaussian noise of variance 3 pixels on the correspondences. As **Ours** can reconstruct from two images only, we perform both pairwise reconstructions and joint reconstruction from the three pairs in the triplet of images available for each surface. We report the results in Table 1. We obtain a very accurate reconstruction from 2 images only. Beside **Ours**, **Va09** is the only baseline that can reconstruct from 2 images. However, it does not perform well on this data. **Lee16** and **An17** are designed for video sequences, and thus need more than 3 images to perform effectively. The remaining methods can operate on three images, but their accuracy is lower than ours, especially in terms of normal accuracy. Since we can discard the normals that have a low reliability, we obtain the accuracy of reconstruction is strengthened using multiple images. Fig. 2 further confirms the quality of our reconstructions by depicting the normals we obtain *without* any smoothing.

Results on Real Data. We tested our approach on the **Paper** [23] and **Rug** [20] datasets. The first comprises 190

images of a deforming sheet of paper with 1500 point correspondences, and the second 160 images of a deforming rug with 4000 point correspondences. The correspondences in the **Paper** dataset were obtained using SIFT with a manual supervision of accuracy and are thus highly accurate. By contrast, those in the **Rug** dataset were computed using the dense optical flow method of [9] and contain errors due to the lack of texture.

Because the computational complexity of the global baselines grows rapidly with the number of correspondences, we evaluated all methods on the full set of correspondences and on a subset of 350 correspondences. For example, **Ch17** has a cubic complexity and hence yields a very high computation time when there are many correspondences. Its Matlab implementation crashes when using all correspondences and using only 1000 correspondences still takes hours on a modern CPU. Similarly **Ch14** and **Va09** take almost 1 hour to reconstruct 20 images and we therefore did not evaluate them on these datasets.

We report our quantitative results in Table 2, and Figure 3 depict qualitative ones. We outperform all baselines except **Ch17** in terms of *Ed* on the partial **Paper** dataset. Crucially, this is achieved at a much reduced computational cost by solving a set of equations in closed-form as opposed to invoking a complex solver. As a result, our approach is about 150 times faster than **Ch17** on 350 correspondences and can handle thousands whereas **Ch17** cannot. Furthermore, our approach is also 50 times faster than **Pa19**, the best performing competitor that relies on a local approach, because we do not have to make a complex formulation to obtain a unique solution for each correspondence.

Table 3 provides a detailed analysis of the run-times of all the methods on 350 and 1500 points. We assume that the input point correspondences and their derivatives are pre-computed. Therefore, the timings only encode the computation of the normals or 3D points. Our approach yields the fastest run-times, seconded only by **An17**. Note, however, that **An17** has a parallel implementation and is computationally optimized. By contrast, our approach, as all the other ones, is implemented in Matlab and not optimized for speed. The relative slowness of the other local method arises from the local normal estimators of **Pa19** and **Pa20** having to minimize the sum of squares of polynomials, which is expensive even if it has linear complexity. **Pa20** is further slowed down by having to transform polynomials into univariate expressions. By contrast, our local normal estimator is computationally cheap as it has a closed-form solution.

Results on dense data. We recorded a video sequence featuring a textureless blue sheet deforming isometrically using a Kinect. It comprises 60 images and 7K point correspondences are tracked using the dense optical flow [9]. Optical flow on textureless surfaces is prone to large error

Table 2: **RMSE results on Paper and Rug datasets.** 'X' indicates that the method does not evaluate normals. '—' indicates that method failed to return a result due to its high computational complexity.

Method	Paper (partial)		Rug (partial)		Paper (full)		Rug (full)	
	En	Ed	En	Ed	En	Ed	En	Ed
Lee16	X	21.6	X	89.8	X	21.9	X	90.7
An17	X	14.7	X	60.6	X	14.7	X	63.7
Ch17	X	5.4	X	63.5	-----	-----	-----	-----
Pa19	17.3	8.3	34.5	52.4	18.4	10.2	35.8	54.3
Pa20	20.7	10.2	28.1	46.1	24.8	11.3	29.4	47.1
Ours	10.2	6.1	14.0	42.0	10.4	5.9	14.4	42.4

Table 3: **Computation times** as a function of the number of images and points used.

Images	350 points							
	Lee16	An17	Va09	Ch14	Ch17	Pa19	Pa20	Ours
10	17.8	10	69.4	75.3	31.3	9.7	24.5	0.2
30	23.4	12	3103	3407	129	12.5	32.7	0.6
60	45.6	19	-----	-----	-----	14.8	45.3	1.9
Images	1500 points							
	Lee16	An17	Va09	Ch14	Ch17	Pa19	Pa20	Ours
10	256	12	1435	1256	995	103	745	0.5
30	987	14	-----	-----	3400	118	1205	2.0
60	1705	22	-----	-----	-----	124	1807	5.8

and the flow we obtain confirms this. We compare the performance of our method against that of **Pa19** and **An17**. Fig. 4 shows grid views of the reconstructions for two specific images. **Ours** performs better as it captures the geometry of the object well. The average reconstruction error for **Ours** is 27 mm whereas for **Pa19** and **An17** it is 37.2 and 48.5 mm, respectively.

7. Conclusion

We have proposed an approach to Non-Rigid Structure-from-Motion (NRSfM) that can estimate normals in closed form from images pairs given a 2D warp and point correspondences between the two images. It does so in closed form from individual correspondences and is therefore fast. Furthermore, it can estimate if these normals are reliable given the motion from one image to the next. And when they are found to be, our experiments show that they are indeed very accurate. As a result, our method performs well with both semi-dense and dense data and can reconstruct large and small deformations at a low computational cost.

Our next step will be to remove the dependency on expensive methods to compute warps and integrate normals so that a truly real-time application can be developed.

Acknowledgements. This work was supported in part by the Swiss National Science Foundation.

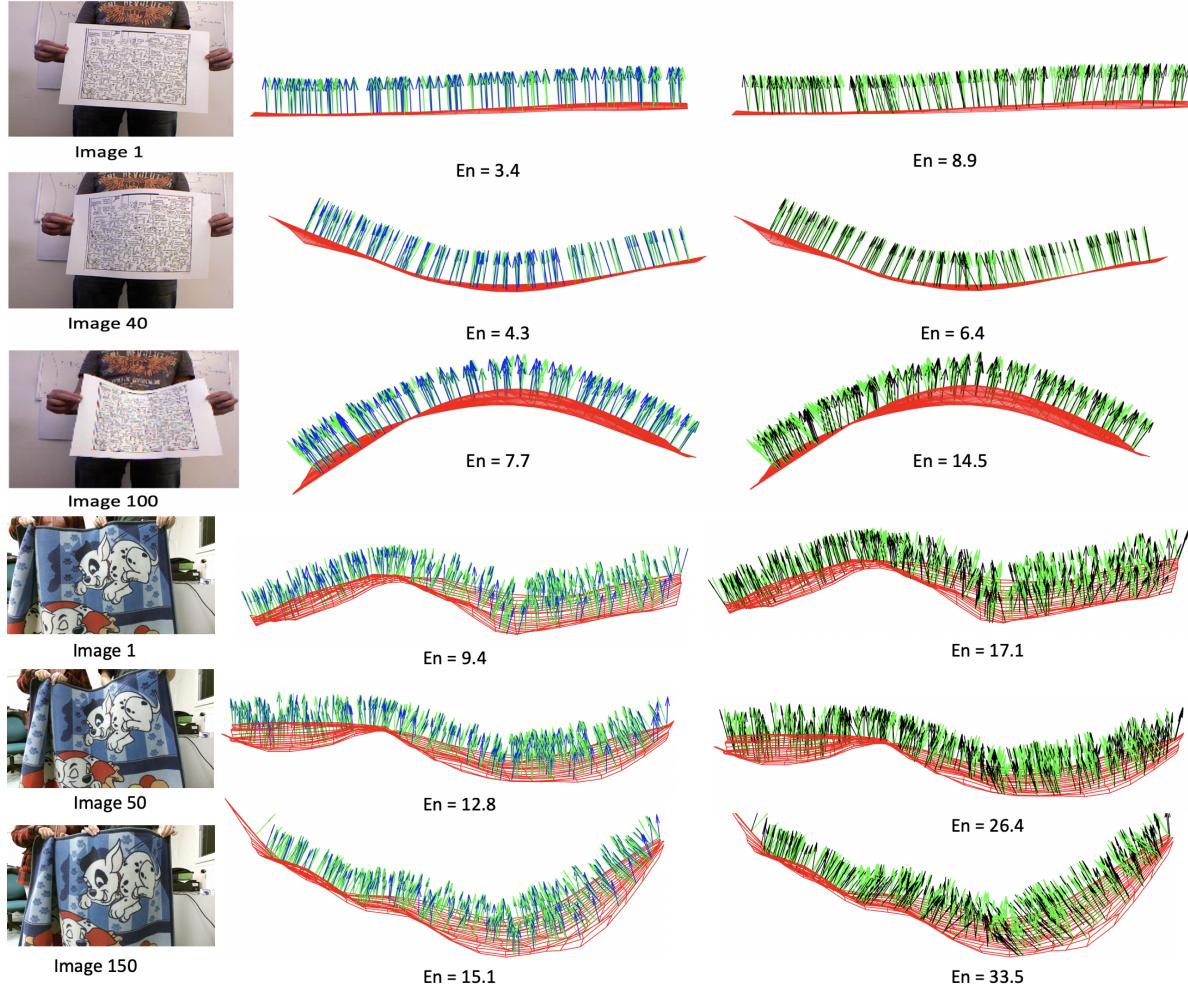


Figure 3: **Paper** and **Rug** datasets. Reconstructed normals on three images. The ground-truth normals are shown in green, the ones predicted by **Ours** in blue, and those by **Ch19** in black. Note that our normals are far less noisy.

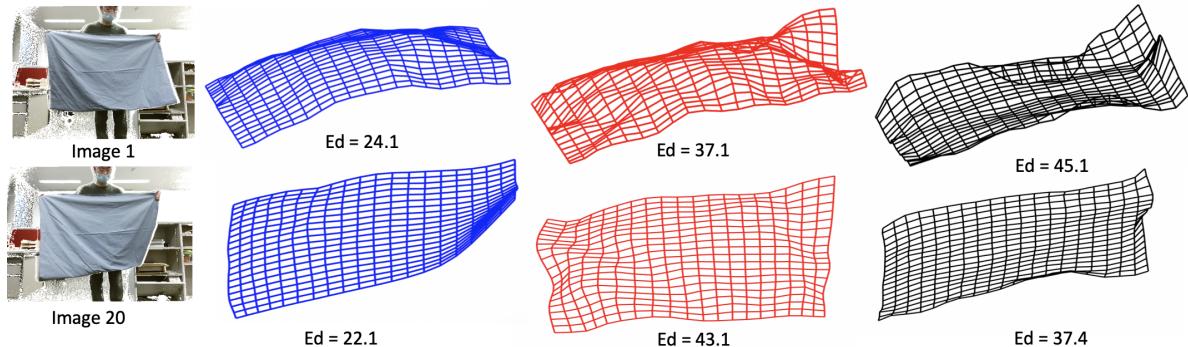


Figure 4: **Sheet dataset**. Reconstructed surfaces for two images. The predictions of **Ours** are shown in blue, of **Pa19** in red, and of **An17** in black. Note that our reconstruction is less noisy and matches the surface 3D shape much better.

References

- [1] I. Akhter, Y. Sheikh, S. Khan, and T. Kanade. Nonrigid Structure from Motion in Trajectory Space. In *Advances in Neural Information Processing Systems*, December 2008. [1](#), [2](#)
- [2] Mohammad Dawud Ansari, Vladislav Golyanik, and Didier

- Stricker. Scalable Dense Monocular Surface Reconstruction. In *3DV*, 2017. 6
- [3] 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, 1989. 2
- [4] C. Bregler, A. Hertzmann, and H. Biermann. Recovering Non-Rigid 3D Shape from Image Streams. In *Conference on Computer Vision and Pattern Recognition*, 2000. 1, 2
- [5] A. Del Bue, X. Llad, and L. Agapito. Non-Rigid Metric Shape and Motion Recovery from Uncalibrated Images Using Priors. In *Conference on Computer Vision and Pattern Recognition*, 2006. 1, 2
- [6] A. Chhatkuli, D. Pizarro, and A. Bartoli. Non-Rigid Shape-From-Motion for Isometric Surfaces Using Infinitesimal Planarity. In *British Machine Vision Conference*, 2014. 1, 2, 6
- [7] A. Chhatkuli, D. Pizarro, T. Collins, and A. Bartoli. Inextensible Non-Rigid Structure-From-Motion by Second-Order Cone Programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2017. 1, 2, 3, 6
- [8] Y. Dai, H. Li, and M. He. A Simple Prior-Free Method for Non-Rigid Structure-From-Motion Factorization. *International Journal of Computer Vision*, 107(2):101–122, 2014. 1, 2, 3
- [9] R. Garg, A. Roussos, and L. Agapito. Dense Variational Reconstruction of Non-Rigid Surfaces from Monocular Video. In *Conference on Computer Vision and Pattern Recognition*, 2013. 7
- [10] P. F. Gotardo and A. M. Martinez. Computing Smooth Time Trajectories for Camera and Deformable Shape in Structure from Motion with Occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10):2051–2065, 2011. 1, 2
- [11] P. Ji, H. Li, Y. Dai, and I. Reid. Maximizing Rigidity Revisited: A Convex Programming Approach for Generic 3D Shape Reconstruction from Multiple Perspective Views. In *International Conference on Computer Vision*, 2017. 1, 2
- [12] A. Kock. *Synthetic Geometry of Manifolds*. Cambridge University Press, 2010. 2
- [13] C. Kong and S. Lucey. Deep Non-Rigid Structure from Motion. In *International Conference on Computer Vision*, 2019. 2
- [14] J. Lamarca, S. Parashar, A. Bartoli, and J. M. M. Montiel. DefSLAM: Tracking and Mapping of Deforming Scenes from Monocular Sequences, 2019. 1
- [15] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer, 1997. 2, 3
- [16] M. Lee and S. Cho, J. and Oh. Consensus of Non-Rigid Reconstructions. In *Conference on Computer Vision and Pattern Recognition*, 2016. 1, 2, 3, 6
- [17] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 20(2):91–110, November 2004. 2
- [18] E. Malis and M. Vargas. Deeper Understanding of the Homography Decomposition for Vision-Based Control. Technical report, INRIA, 2007. 4
- [19] S. Parashar and A. Bartoli. 3DVFX: 3D Video Editing Using Non-Rigid Structure-From-Motion. In *Eurographics*, 2019. 1
- [20] S. Parashar, D. Pizarro, and A. Bartoli. Isometric Non-Rigid Shape-From-Motion with Riemannian Geometry Solved in Linear Time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. 1, 2, 6
- [21] S. Parashar, D. Pizarro, and A. Bartoli. Local Deformable 3D Reconstruction with Cartan’s Connections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019. 1, 2, 3, 5, 6
- [22] S. Parashar, M. Salzmann, and P. Fua. Local Non-Rigid Structure-From-Motion from Diffeomorphic Mappings. In *Conference on Computer Vision and Pattern Recognition*, 2020. 1, 2, 5, 6
- [23] M. Salzmann, R. Hartley, and P. Fua. Convex Optimization for Deformable Surface 3D Tracking. In *International Conference on Computer Vision*, October 2007. 6
- [24] V. Sidhu, E. Tretschk, V. Golyanik, A. Agudo, and C. Theobalt. Neural Dense Non-Rigid Structure from Motion with Latent Space Constraints. In *European Conference on Computer Vision*, 2020. 2
- [25] D. Sun, X. Yang, M. Liu, and J. Kautz. Pwc-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. In *Conference on Computer Vision and Pattern Recognition*, 2018. 2
- [26] N. Sundaram, T. Brox, and K. Keutzer. Dense Point Trajectories by Gpu-Accelerated Large Displacement Optical Flow. In *European Conference on Computer Vision*, 2010. 2
- [27] J. Taylor, A. D. Jepson, and K. N. Kutulakos. Non-Rigid Structure from Locally-Rigid Motion. In *Conference on Computer Vision and Pattern Recognition*, June 2010. 1, 2, 3
- [28] L. Torresani, D. B. Yang, E.J. Alexander, and C. Bregler. Tracking and Modeling Non-Rigid Objects with Rank Constraints. In *Conference on Computer Vision and Pattern Recognition*, pages 493–500, 2001. 1
- [29] A. Varol, M. Salzmann, E. Tola, and P. Fua. Template-Free Monocular Reconstruction of Deformable Surfaces. In *International Conference on Computer Vision*, September 2009. 1, 2, 3, 6
- [30] S. Vicente and L. Agapito. Soft Inextensibility Constraints for Template-Free Non-Rigid Reconstruction. In *European Conference on Computer Vision*, 2012. 1, 2, 3