

Deformable 3D Reconstruction with an Object Database

Pablo F. Alcantarilla
pablofdezalc@gmail.com

Adrien Bartoli
adrien.bartoli@gmail.com

ISIT-UMR 6284 CNRS
Université d'Auvergne
Clermont-Ferrand, France

Abstract

Deformable 3D reconstruction from 2D images requires prior knowledge on the scene structure. Template-free methods use *generic* prior knowledge such as piecewise smoothness but require multiple images with significant baseline. Template-based methods require only one image but handle only one object for which they need specific prior knowledge, namely a 3D template. We here propose a novel method that alleviates the strong assumptions of both the template-free and template-based methods: our method uses multiple templates to achieve deformable 3D reconstruction from only one image and for multiple objects. It uses object recognition to automatically discover what objects are visible in the input image and to select the appropriate templates for deformable 3D reconstruction. The object database is built offline. Crucially, this database does not only contain appearance descriptors as in existing object recognition frameworks, but also material properties to facilitate deformable 3D reconstruction. We show successful experimental results with objects made of various materials such as paper, cloth and plastic.

1 Introduction

Deformable 3D reconstruction from monocular video data has been an active area of research over the last few years [1, 2, 3, 4, 5]. Techniques for the rigid case are well advanced: it is now possible to build city-scale 3D models from photo collections in a few hours [6, 7]. Deformable 3D reconstruction however still lags behind. Deformable 3D reconstruction requires prior knowledge on the scene structure such as the shape space or the type of deformations in order to constrain the problem and resolve the ambiguities. Deformable 3D reconstruction approaches can be classified in two groups: template-free [8, 9, 10, 11, 12] and template-based methods [13, 14, 15]. Template-free methods use *generic* prior knowledge such as piecewise smoothness and require multiple images with significant baseline to recover the deformable structure. Template-based methods require only one image but currently handle only one object at a time for which they need a 3D template.

In this paper, we propose a novel approach that alleviates the strong assumptions of both the template-free and template-based approaches: we use multiple templates from an object database to achieve deformable 3D reconstruction from only one image and for multiple objects at a time. Our methodology combines different techniques including object recognition by registration and template-based deformable 3D reconstruction into a single framework.

In our framework, an object database is built offline. This database contains not only appearance descriptors as in existing object recognition frameworks [20, 22, 25], but also material properties such as deformation constraints to facilitate deformable 3D reconstruction. At runtime we use object recognition to automatically discover what objects are visible in the current input image and to select the appropriate templates for deformable 3D reconstruction. For this purpose, we perform wide-baseline image matching between the stored templates in the database and the input image that contains the deforming surfaces. We use an outlier rejection method [26] to obtain a set of clean-up matches between each detected template and the objects present in the 2D input image. For those objects that have a number of clean-up matches higher than a defined threshold, we compute an image warp [6, 11]. Each warp encodes the particular deformation of an object in the image. Finally, given the estimated warps we perform deformable 3D reconstruction for the detected objects [9].

Our approach is the first to use an object database to aid deformable 3D reconstruction. In terms of genericity, it lies between existing template-based and template-free methods, as it assumes that strong priors on the world can be modeled but is not object-specific.

We show successful deformable 3D reconstruction results of multiple objects from a single image. The objects in the database are made of different materials such as paper, cloth and plastic. The rest of the paper is organized as follows: we describe the related work in Section 2. Then, we introduce notation and a general system overview in Section 3. In Section 4 we describe how to build an object database that encodes both appearance and material properties for developable and non-developable objects. Our deformable 3D reconstruction with multiple objects framework is introduced in Section 5 and experimental results are reported in Section 6. Finally, conclusions and future work are presented in Section 7.

2 Related Work

Template-free 3D deformable reconstruction methods estimate the 3D shape of a deforming object from a series of 2D projections through a sequence of images. These methods require features tracks with enough baseline between the images and make use of spatio-temporal smoothness priors to constrain the problem. Early approaches follow the seminal work of Bregler *et al.* [9, 12, 23, 31]. Recent work follow the so called *piecewise* approach, where surface patches are reconstructed first and glued together in a second step [15, 27, 30, 33]. Interesting results were also obtained considering the trajectory space approach [3], in which the evolving 3D structure is represented in a trajectory space by means of a linear combination of basis trajectories. Recently, [1] proposed a Bayesian Finite Element Method (FEM) modeling of deformations integrated within an Extended Kalman Filter (EKF) framework.

Template-based deformable 3D reconstruction methods require one single image, since they rely on a prior template where the 3D shape of the object at rest is known. These methods have two main steps: (i) registration of the input image to the template and (ii) 3D deformable reconstruction from reprojection and deformation constraints. Some works [11, 24, 28] proposed convex formulations considering inextensible deformation constraints maximizing the surface’s depth. Recently, analytical solutions for the isometric and conformal deformation cases were proposed [8]. They showed that template-based isometric surface reconstruction from a single view registered to a template, generally has a single solution for both developable and non-developable surfaces.

In the last few years, there has also been an increasing interest on combining object

recognition with Simultaneous Localization and Mapping (SLAM) approaches, adding a *semantic* meaning to the generated maps [13, 14]. In addition, object recognition and augmentation from a large object database are becoming standard techniques in large-scale augmented reality applications such as in [25, 26]. However, these approaches are only valid for performing detection and augmentation of rigid objects. Therefore, our approach can be of interest towards a more complete augmentation and scene interaction with 3D deformable objects.

3 Notation and System Overview

In this section we describe the main notation and provide a general system overview. Figure 1 depicts a general overview of the different steps in our algorithm using figures from our experiments.

The object database DB is composed of M different objects $DB = \{O_1, \dots, O_M\}$, either non-developable or developable¹. Each object in the database O_i has a set of N_i appearance descriptors $D_{O_i} = \{d_1, \dots, d_{N_i}\}$ and a 3D template T_{O_i} . In addition, associated to each 3D template, each object also has a 2D parameterization of the template denoted as P_{O_i} .

Given an input image, we compute a set of feature descriptors D_q and perform object recognition to detect the objects from the database that are present in the image. For this purpose, we first use a hierarchical vocabulary tree [27] to determine a set of M_1 object hypotheses ($M_1 \leq M$). Then, we match the set of features D_q to the associated set of descriptors in the database for each particular object hypothesis to form a set of correspondences that may be corrupted with outliers. We use the feature-based surface detection method described in [28] to obtain a set of clean-up matches between the target image and each template, estimating an image warp W_{O_i} that models the image deformation between the target image and of each possible template.

After feature-based surface detection we have a set of M_2 objects, whose deformation is consistent with the computed warps and the number of clean-up matches is higher than a defined threshold. Our problem is now to solve for the deformable 3D reconstruction of each detected object given their 2D parameterization P_{O_i} and estimated warps W_{O_i} . In more details, we are interested in obtaining for each detected object the unknown surface function φ_{O_i} that maps a point from the 2D parameterization P_{O_i} to the 3D deforming surface S_i . At this point we can perform template-based isometric surface reconstruction, since we can obtain a single solution considering isometric deformations for both the developable and non-developable cases [6].

4 Building the Object Database

We now describe how to build an object database that encodes appearance information and deformation constraints for each object. Basically, an object in our database is an entity that comprises of a set of appearance descriptors D_{O_i} , a 3D template T_{O_i} and its 2D parameterization P_{O_i} . This information is obtained from a set of images I_{O_i} of the object of interest. For developable objects only one image is necessary, whereas for non-developable objects

¹The surface of developable objects has a zero Gaussian curvature and can be flattened onto a plane without distortion such as stretching or compression.

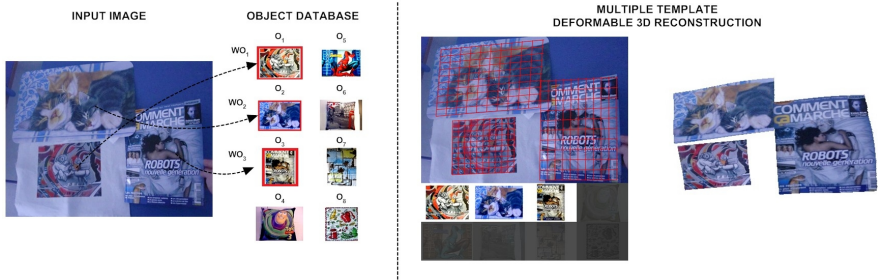


Figure 1: Given an input image, we first perform object recognition to detect database objects in the input image. Then, we compute 2D image warps that model the deformation of each particular object between the input image and a 2D parameterization of the 3D template. Finally, using the estimated warps we perform template-based isometric surface reconstruction. The detected objects from the database are highlighted, whereas non-detected objects are depicted in a darker color.

several images are used. Figure 2 depicts an example of the object database construction for the two different types of surfaces non-developable and developable.

For feature detection and description, we use the Speeded Up Robust Features (SURF) [17] for obtaining a set of keypoints with associated descriptors. We decided to use SURF since it has a good compromise between speed and performance. However, our framework is valid for any particular feature detector-descriptor choice.

Regarding the 3D templates of the objects and their 2D parameterizations, we distinguish between two different cases: developable and non-developable objects. For the first case, obtaining a 3D template of the object T_{O_i} is trivial, since developable objects can be physically flattened to a plane without any distortion. For these objects we use a single image from an object of known dimensions flattened into a plane and extract SURF features to obtain the set of descriptors D_{O_i} . The 3D reconstruction of the planar object with known dimensions can be obtained by means of pose rectification or homography [17]. For developable objects the 2D parameterization P_{O_i} is the same as its 3D template T_{O_i} .

For non-developable objects we use multiple images of the object and perform rigid Structure-from-Motion (SfM) to build a 3D template, matching features between different views. We use the software *Bundler* [29] for the generation of a sparse 3D point cloud for the non-developable objects. This 3D point cloud can be densified by using multi-view stereo algorithms such as [16]. In the resulting 3D point cloud, a 3D point can be visible in several images from the whole sequence, and therefore the point will have different descriptors with a small variance between them. For those 3D points we compute an average descriptor vector by averaging the descriptors from all the different views.

The next step is building a 3D triangle-mesh from the 3D point cloud. For this purpose, well-known surface reconstruction algorithms from point clouds such as [8, 18] can be used. The resulting 3D mesh will be used as the 3D template of the surface with the list of associated descriptors. For non-developable objects we obtain the 2D parameterization P_{O_i} by means of conformal flattening [19]. This technique is very popular in the computer graphics literature and widely used in many different applications such as texture mapping. Conformal flattening maps a 3D mesh onto a plane while approximately preserving angles. To sum

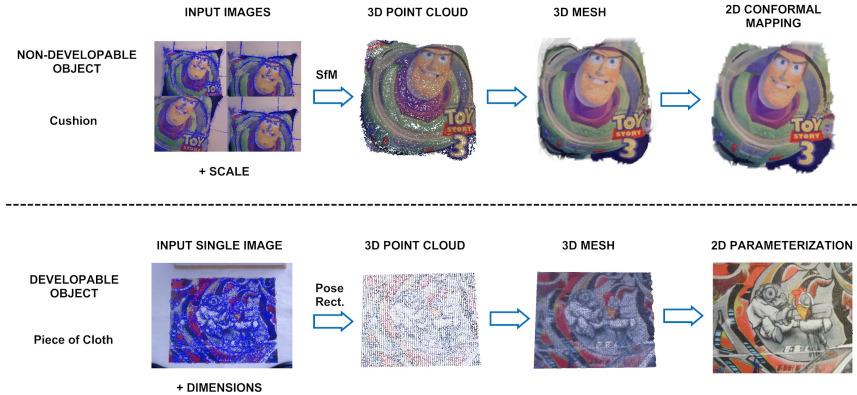


Figure 2: The image shows the process from the input image(s) to the computation of the 3D template and 2D parameterization for non-developable and developable objects. The extracted features are depicted in blue in the input images. First row: non-developable object. Second row: developable object. Best viewed in color.

up, an object in the database O_i is composed of:

$$O_i = \begin{cases} d_{O_i} = \{d_1 \dots d_{N_i}\} & , & \text{Appearance Descriptors} \\ T_{O_i} = \{(x_1, y_1, z_1) \dots (x_{N_i}, y_{N_i}, z_{N_i})\} & , & \text{Geometric 3D Template} \\ P_{O_i} = \{(u_1, v_1) \dots (u_{N_i}, v_{N_i})\} & , & \text{Geometric 2D Parameterization} \end{cases} \quad (1)$$

where the template T_{O_i} is a 3D triangle-mesh with N_i vertices. Similarly, the parameterization P_{O_i} is a 2D triangle-mesh with N_i vertices.

5 Deformable 3D Reconstruction with Multiple Objects

In this section we describe the main steps of our deformable 3D reconstruction with multiple objects approach. Basically, there are three main steps: object recognition, feature-based deformable surface detection and deformable 3D reconstruction.

5.1 Object Recognition

We use a hierarchical vocabulary tree [24] to determine a set of M_1 object hypotheses that can be present in the input image. The vocabulary tree is constructed with the set of descriptors from each object in the database and is built offline. Given a target image, we first extract SURF features and quantize the extracted descriptors D_q according to the hierarchical vocabulary tree. Then, we use a *Term Frequency-Inverse Document Frequency* (TF-IDF) score to find a similarity measure between each of the objects in the database and the visual words from the input image. We choose the set of M_1 best scores as the possible objects present in the image.

Thanks to the use of a hierarchical vocabulary tree we can perform image queries in constant operation time even when considering very large object databases. At the end of the vocabulary tree step, we have a set of M_1 potential objects present in the image that need to be validated with the image deformation model.

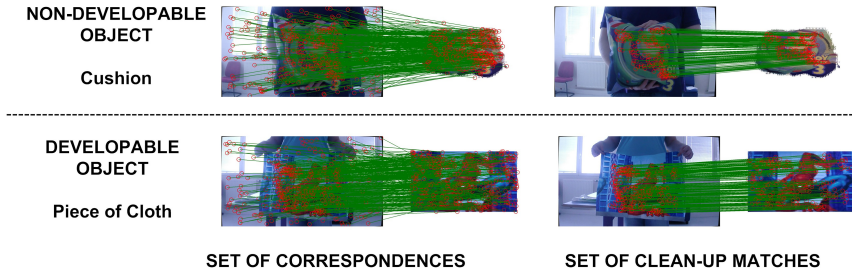


Figure 3: Results of the outlier rejection method for image matching in deformable surfaces. The set of matches is computed between the target image and the 2D parameterization of the 3D template. First row: a piece of cloth (developable object). The algorithm returns 248 clean-up matches from a set of 412 correspondences. Second row: a cushion (non-developable object). The algorithm returns 198 clean-up matches from a set of 386 correspondences.

5.2 Feature-Based Deformable Surface Detection

For performing deformable 3D reconstruction, we need to check also for geometric consistency of the potential detected objects with the deforming surfaces in the image. Given the set of M_1 potential objects, we use the feature-based approach described in [26] for detecting deformable surfaces in the input image. This algorithm performs wide-baseline feature matching between a template image P_{O_i} and an input image I_q estimating a parametric warp that represents the image deformation model. The algorithm returns a set of clean-up matches rejecting outliers between the template and the input image. Figure 3 depicts two examples of the outlier rejection results for two different objects in our database, one piece of cloth and one cushion.

For each of the M_1 potential objects, we match the set of descriptors from each object D_{O_i} with respect to the set of descriptors from the target image D_q . Image matching is performed in a fast way by checking for the Approximated Nearest Neighbors (ANNs) [27] of the template image. The set of correspondences between the query image and descriptors from potential objects may contain outliers. Therefore, we need to find the set of inliers between each template and the query image. In addition, we search for the warp function W_{O_j} that maps a point from the 2D template P_{O_j} to a point in the target image I . We use a parametric warp based on Radial Basis Functions (RBF), concretely the Thin Plate Spline (TPS), that minimizes the integral bending energy. For more details about the TPS warp, see [4, 5].

The outliers are removed from the set of correspondences by assuming that the surface is locally smooth and its local topology must be preserved. From the set of clean-up matches, the final warp W_{O_j} is obtained.

5.3 Deformable 3D Reconstruction

In order to constrain the 3D reconstruction problem, we assume that the surfaces deform isometrically. This is a reasonable assumption that has been made in most of the works from the literature [24, 28]. These works also assume the surfaces to be developable, however in our formulation we can also deal with non-developable surfaces.

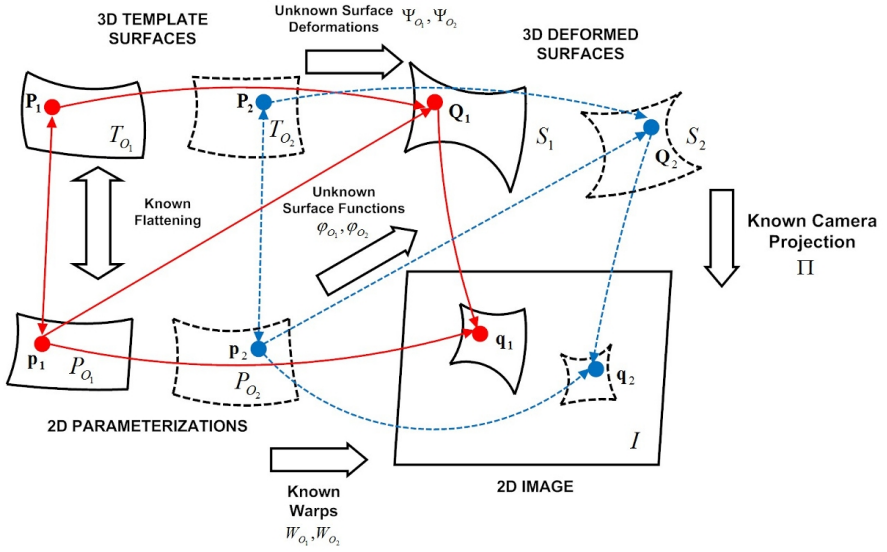


Figure 4: Geometric modeling of monocular multiple-template based reconstruction. For clarity reasons the figure just show the formulation for two objects.

Template-based isometric surface reconstruction from a single view registered to a template generally has a single solution as shown in [6]. Their approach is based on a system of partial differential equations (PDEs) that take into account reprojection and surface deformation constraints. In this paper, we extend that framework to allow the reconstruction of multiple templates. Figure 4 depicts the geometric modeling of monocular multiple template-based reconstruction, assuming that the surfaces have disk topology and that the object templates have been correctly detected. The geometric modeling depicted in Figure 4 takes into account reprojection and deformation constraints for each detected object. The projection operator Π maps a 3D point $Q_i = (X Y Z)^t$ from a 3D deformed surface S_i to a 2D point in the image plane $q_i = \Pi(Q_i)$, with $q_i = (u, v)^t$. Assuming a pin-hole camera model, the projection operator Π contains the known camera intrinsics parameters in matrix K . If the intrinsics effects are *undone* in the image (by multiplying by K^{-1} , the projection operator is simply $\Pi(Q_i) = \frac{1}{Z} (X Y)^t$. Deformation constraints are encoded in the warps W_{O_i} that map a 2D point from an object’s parameterization p_i to a 2D point in the target image q_i , i.e. $q_i = W_{O_i}(p_i)$.

Considering isometric deformations, we can obtain an analytical solution for the unknown surface functions φ_{O_i} , that maps a 2D point p_i from the 2D parameterization P_{O_i} to a 3D point $Q_i = \varphi_{O_i}(p_i)$ in the deformed surface S_i . The analytical solution is obtained by solving a system of PDEs, yielding a solution that depends only on the warp W_{O_i} and the set of first and second order warp derivatives. Every solution of the system of PDEs is of the form:

$$\varphi_{O_i} = fA^{-\frac{1}{2}} \begin{pmatrix} W_{O_i} \\ 1 \end{pmatrix}, \quad (2)$$

where f and A are functions that depend on the warp W_{O_i} and the set of first and second order warp derivatives. It is important to notice here, that even though the geometric modeling takes into account the camera intrinsics parameters, the camera used for building the database



Figure 5: Reconstruction results for a single object in the image. First row: *Graffiti* experiment. Second row: *T-Shirt* experiment. Left Image: Reconstructed mesh projected in the 2D image. Middle: Reconstructed mesh in 3D. Right: Sets of clean-up matches between the target image and the object template.

can be different from the one used for the deformable reconstruction.

6 Experimental Results

The database used in our experiments comprises of 8 different objects from different materials and deformation properties. We used typical objects made of paper, cloth and plastic. In particular, we have 6 developable objects and 2 non-developable objects. The images for building the database have a resolution of 2048×1536 pixels and were obtained by using a standard consumer digital camera. Regarding the images for the deformable 3D reconstruction, we used two different digital cameras and up to three different image resolutions 2048×1536 , 1280×720 and 640×480 pixels.

Figure 5 depicts the deformable reconstruction results considering a single object in the image for two different objects. For the first example, we used a low image resolution of 640×480 pixels and for the second experiment we used an image resolution of 1280×720 pixels. The algorithm is able to return 130 clean-up matches from a set of 167 correspondences, capturing the deformation of the surface accurately for the first experiment. On the second experiment the algorithm returns 147 clean-up matches from a set of 255 correspondences.

Figure 6 depicts reconstruction results for experiments considering two objects in the image. We show three different experiments considering an image resolution of 1280×720 pixels. As can be observed, our algorithm is able to detect multiple objects in the image and obtain a good reconstruction even in the presence of significant deformations. We also performed the reconstruction of three different objects simultaneously. Figure 7 depicts one example in which three different deforming objects are reconstructed from a single image with an image resolution of 2048×1536 pixels.

We did not observe any false positive regarding object detection during our experiments.

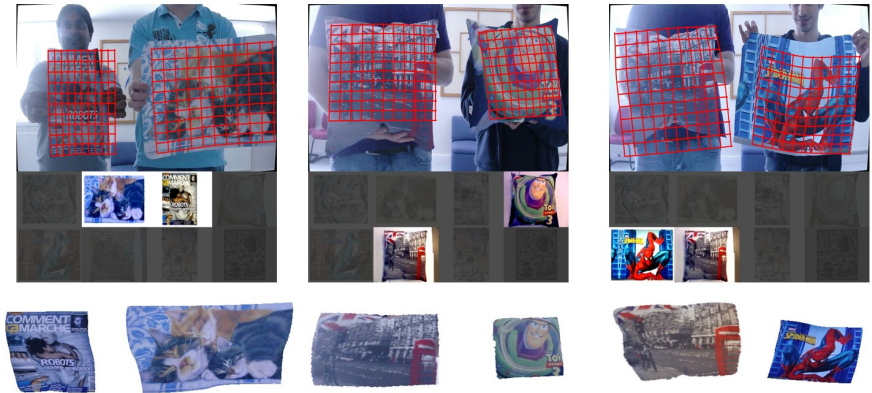


Figure 6: Reconstruction results for two objects in the image. First row: Reconstructed meshes projected in 2D. Second row: Reconstructed meshes in 3D.



Figure 7: Reconstruction results for three objects in the image. Left: Reconstructed meshes projected in 2D. Middle and right: two different views of the reconstructed meshes in 3D.

This is because, once a set of object hypotheses has been computed from the hierarchical vocabulary tree, we check each of the hypotheses for geometric consistency with the deforming surfaces in the image. However, it is possible that sometimes the object can be detected properly by the vocabulary tree but the small number of correspondences is too few to obtain an accurate deformable 3D reconstruction. This may happen when many objects are present in the image or the objects are far from the camera.

We implemented the whole system in C++ and obtained a processing time of approximately 2 seconds per frame when dealing with two objects in the image using a Core 2 Duo 2.4GHz laptop computer and images of 1280×720 pixels resolution. However, we expect to have real-time performance in the near future, by using multiple cores and GPGPU processing for speeding-up the computation of some steps of our algorithm. For example, by using faster feature detection and description methods, the computation time can be reduced considerably, since feature detection is one of the most time consuming steps in our framework (approximately 1 second per frame).

7 Conclusions and Future Work

In this paper we have presented the first deformable 3D reconstruction approach using an object database. Our method uses multiple templates to achieve deformable 3D reconstruction from only one image and for multiple objects. It uses object recognition to automatically discover which objects are visible in the input image and to select the appropriate templates for deformable 3D reconstruction. Our work opens a whole new area of approaches that can benefit from using strong priors encoded in a versatile object database.

Possible extensions of our work include the study of other possible deformations such as conformal or quadratic local models. In addition, the object database framework allows one to incorporate more information as for example material properties of each particular object such as stiffness, texture or local deformation models. Another possible extension of our work is to use specific feature detectors (e.g. face detector) to deal with other classes of objects and deformations.

References

- [1] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski. Building Rome in a Day. In *Intl. Conf. on Computer Vision (ICCV)*, 2009.
- [2] A. Agudo, B. Calvo, and J.M.M. Montiel. Finite Element based Sequential Bayesian Non-Rigid Structure from Motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Providence, Rhode Island, USA, 2012.
- [3] I. Akhter, Y. Sheikh, S. Khan, and T. Kanade. Trajectory space: A dual representation for nonrigid structure from motion. *IEEE Trans. Pattern Anal. Machine Intell.*, 2011.
- [4] A. Bartoli. Maximizing the predictivity of smooth deformable image warps through cross-validation. *Journal of Mathematical Imaging and Vision*, 31(2):133–145, 2008.
- [5] A. Bartoli, M. Perriollat, and S. Chambon. Generalized thin-plate spline warps. *Intl. J. of Computer Vision*, 88(1):85–110, May 2010.
- [6] A. Bartoli, Y. Gérard, F. Chadebecq, and T. Collins. On template-based reconstruction from a single view: Analytical solutions and proofs of well-posedness for developable, isometric and conformal surfaces. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Providence, Rhode Island, USA, 2012.
- [7] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. SURF: Speeded up robust features. *Computer Vision and Image Understanding*, 110(3):346–359, 2008.
- [8] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin. The ball-pivoting algorithm for surface reconstruction. *IEEE Trans. on Visualization and Computer Graphics*, 5(4):349–359, Oct. 1999.
- [9] C. Bregler, A. Hertzmann, and H. Biermann. Recovering non-rigid 3D shape from image streams. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2000.

- [10] F. Brunet, R. Hartley, A. Bartoli, N. Navab, and R. Malgouyres. Monocular template-based reconstruction of smooth and inextensible surfaces. In *Asian Conf. on Computer Vision (ACCV)*, 2010.
- [11] A. Del Bue and A. Bartoli. Multiview 3D warps. In *Intl. Conf. on Computer Vision (ICCV)*, Barcelona, Spain, 2011.
- [12] A. Del Bue, X. Lladó, and L. Agapito. Non-rigid metric shape and motion recovery from uncalibrated images using priors. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, New York, NY, USA, 2006.
- [13] R.O. Castle and D.W. Murray. Keyframe-based recognition and localization during video-rate parallel tracking and mapping. *Image and Vision Computing*, 29(8):524–532, 2011.
- [14] J. Civera, D. Gálvez-López, L. Riazuelo, J.D. Tardós, and J.M.M. Montiel. Towards semantic SLAM using a monocular camera. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, San Francisco, California, USA, 2011.
- [15] T. Collins and A. Bartoli. Locally affine and planar deformable surface reconstruction from video. In *Proceedings of the International Workshop on Vision, Modeling and Visualization (VMV)*, Siegen, Germany, 2010.
- [16] Y. Furukawa, B. Curless, S.M. Seitz, and R. Szeliski. Towards Internet-scale multi-view stereo. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [17] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [18] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, 2006.
- [19] B. Lévy, S. Petitjea, N. Ray, and J. Maillot. Least squares conformal maps for automatic texture atlas generation. In *Conf. on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 362–371, 2002.
- [20] D.G. Lowe. Object recognition from local scale-invariant features. In *Intl. Conf. on Computer Vision (ICCV)*, pages 1150–1157, Corfu, Greece, 1999.
- [21] M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *Intl. Conf. on Computer Vision Theory and Applications (VISAPP)*, 2009.
- [22] D. Nistér and H. Stewénus. Scalable recognition with a vocabulary tree. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2006.
- [23] M. Paladini, A. Bartoli, and L. Agapito. Sequential non-rigid structure from motion with the 3D-implicit low rank shape model. In *Eur. Conf. on Computer Vision (ECCV)*, Crete, Greece, 2010.
- [24] M. Perriollat, R. Hartley, and A. Bartoli. Monocular template-based reconstruction of inextensible surfaces. *Intl. J. of Computer Vision*, 95(2):124–137, 2011.

- [25] J. Pilet and H. Saito. Virtually augmenting hundreds of real pictures: An approach based on learning, retrieval, and tracking. In *IEEE Virtual Reality (VR)*, Waltham, MA, USA, 2010.
- [26] D. Pizarro and A. Bartoli. Feature-based deformable surface detection with self-occlusion reasoning. *Intl. J. of Computer Vision*, 97(1):54–70, March 2012.
- [27] C. Russell, J. Fayad, and L. Agapito. Energy based multiple model fitting for non-rigid structure from motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [28] M. Salzmann and P. Fua. Reconstructing sharply folding surfaces: A convex formulation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Miami, USA, 2009.
- [29] N. Snavely, S.S. Seitz, and R. Szeliski. Modeling the world from Internet photo collections. *Intl. J. of Computer Vision*, 80(2):189–210, 2008.
- [30] J. Taylor, A. D. Jepson, and K.N. Kutulakos. Non-rigid structure from locally-rigid motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [31] L. Torresani, A. Hertzmann, and C. Bregler. Non rigid structure-from-motion: Estimating shape and motion with hierarchical priors. *IEEE Trans. Pattern Anal. Machine Intell.*, 30(5):878–892, May 2008.
- [32] H. Uchiyama and E. Marchand. Toward augmenting everything: Detecting and tracking geometrical features on planar objects. In *IEEE and ACM Intl. Sym. on Mixed and Augmented Reality (ISMAR)*, pages 17–25, Basel, Switzerland, 2011.
- [33] A. Varol, M. Salzmann, E. Tola, and P. Fua. Template-free monocular reconstruction of deformable surfaces. In *Intl. Conf. on Computer Vision (ICCV)*, 2009.