Template-Based Conformal Shape-from-Motion from Registered Laparoscopic Images

Abed Malti
http://isit.u-clermont1.fr/~abed
Adrien Bartoli
http://isit.u-clermont1.fr/~ab
Toby Collins
http://isit.u-clermont1.fr/content/Toby-Collins

ALCoV-ISIT
Université d'Auvergne
Clermont-Ferrand
France

Abstract

One of the current limits of laparosurgery is the absence of a 3D vision facility for standard laparoscopes. While there has been significant progress made in visual SLAM (Simultaneous Localization And Mapping) with a single camera, most of the current approaches relies on the assumption that the tissues are rigid or undergo a cyclic deformation. However, in laparoscopic surgery none of these assumptions commonly apply, due to unpredictable and non-isometric deformation of the tissues.

Our main contribution in this paper is a new sequential 3D reconstruction method well-adapted to the peritoneal environment. We draw on recent computer vision results exploiting a template of the environment. The state-of-the-art methods assume that the environment can be modeled as an isometric developable surface, *i.e.* which deforms isometrically to a plane. While this assumption holds for paper-like surfaces, it certainly does not fit for peritoneal surfaces. Our new method tackle these limits: it uses a full 3D template and copes with non-isometric 3D deformations, thanks to two phases. First the 3D template is reconstructed using rigid SfM (Shape-from-Motion) while the surgeon is exploring – but not deforming – the peritoneal environment. Second the 3D template is used during surgery to infer a quasi-conformal deformation to the current 3D shape from a single laparoscopic image only. This makes both sequential processing and effective self-recovery from tracking errors possible.

The proposed approach has been validated on in-vivo laparoscopic videos of the abdominal wall and a uterus. Experimental results illustrate the ability of our method to deal with extensible deformations of the tissues.

1 Introduction

Over the last few years significant efforts have been made toward developing systems for computer aided laparosurgery. The idea is to assist the practitioner during the intervention in order to improve their perception of the intra-operative environment [9]. 3D visualization is one of the major possible improvements. However, due to the unpredictable, complex and non-isometric behaviour of living peritoneal tissues, 3D shape recovery from laparoscopic images is a difficult and open problem. Visual SLAM was proposed for stereo-laparoscopes [8] and more recently for standard monocular laparascopes [3]. The accuracy of 3D reconstruction for these methods depends on a temporal window and on the ability of the state model to account for the complex phenomena occuring in the environment. Errors may accumulate through navigation and produce artefacts in the reconstructed 3D map. Recently [5] proposed a method for 3D reconstruction of the beating heart by assuming cyclic deformation modeled as a combination of basis shapes; this method cannot be used in lapascopy where cyclicity assumption does not hold.

We propose a novel approach to SfM in the complex laparoscopic setting. We extend recent 2D-template-based deformable SfM methods proposed in computer vision for developable surfaces (paper-like) surfaces [10, 11]. The problem of monocular 3D shape recovery is under-constrained because there is an infinite number of 3D surfaces that can project to the same image data. It is then of critical importance to constrain the problem to have a unique consistent solution or at least a set of plausible solutions. Over the years, different types of constraints have been proposed which can be categorized in statistical and physical constraints. Statistical constraints often model the deformation as a linear combination of basis vectors which can be learned online [6]. A non-linear learning method was proposed in [12] where local deformation models are first learned offline and then combined online. Physical constraints include spatial and temporal priors on the surface. In [1] physical constraints are used as priors within a coarse-to-fine shape basis statistical model. An important physical prior is the isometry constraint [2, 10], which imposes that any surface geodesic distance is preserved after deformation. This constraint is unfortunately too strong to reconstruct human tissues which are typically non-isometric. We propose to relax this constraint to quasi-conformal deformation. We further decompose it into shearing and anisotropic scaling which will receive different weights to account for possible surface stretching or shrinking. While this models quite well the environment, a direct consequence is that the template cannot be taken as flat anymore, as was assumed by most previous methods. Our method thus reconstructs a 3D template shape using classical RSfM (Rigid SfM) by taking advantage of the exploratory phase at which the surgeon navigates with the laparoscope inside the peritoneum. The reconstructed model is deformed afterwards at the surgery phase to fit the different shapes taken by the tissues, thereby providing 3D shape at run-time from a single image. Our algorithm is here dubbed DSfM (Deformable SfM). The technical part consists of three major improvements over the state-of-the-art: *(i)* dealing with non-isometric (conformal) instead of isometric surfaces, *(ii)* using a 3D instead of a flat 2D template and *(iii)* creating a custom 3D template using RSfM.

Paper organization. We first describe our DSfM system in §2 and then give details on the deformable shape recovery phase in §3. We finally report experimental results in §4 and conclude. We assume that the intrinsic laparoscope's parameters have been calibrated. Our notation will be introduced throughout the paper.

2 The DSfM system description

2.1 Architecture overview

As depicted in figure 1, our DSfM system can be splitted into two main components:

1. **3D template reconstruction.** Using an RSfM algorithm [4], a set of 3D point clouds of the organ shape is reconstructed. This reconstruction takes advantage of the early intra-operative exploration phase.

2. **Deformable 3D shape reconstruction.** After the inspection step, the surgeon starts to deform the target organ using the surgery tools. This part of the system offers a newer ability to 3D passive reconstruction of laparoscope images. Given the constructed template, it seeks for the 3D deformation.

2.2 Template reconstruction with RSfM

Given M views, RSfM finds camera parameters as well as a set of 3D points $(x_j \; y_j \; z_j)$, $j = 1, \ldots, N_v$ which will be the template points. There are several ways to proceed for RSfM [4]. We chose the classical sequential approach where two different views with enough

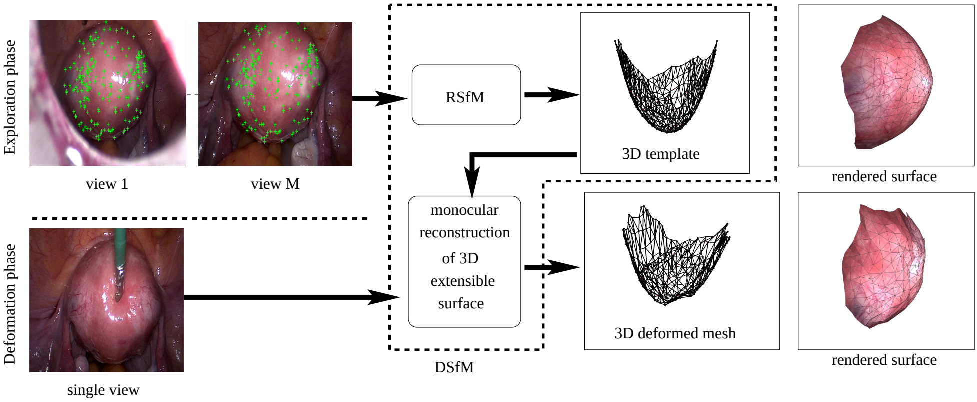

Figure 1: Principle of our DSfM (Deformable Structure-from-Motion) approach. In the first phase the surgeon explores the peritoneal cavity without deforming it; Classical RSfM (Rigid Shape-from-Motion) is used to find the 3D shape called the 3D template. In the second phase, the 3D template is used to infer the 3D shape observed from only a single laparoscopic view. This makes the approach resistant to temporary registration or tracking errors and well-adapted to potentially live sequential processing.

baseline are used to compute the essential matrix, from which the relative camera position can be extracted and used to triangulate a first set of 3D points. The camera position for each other view is then computed on turn using camera resection, and new 3D points are triangulated. Finally, bundle adjustment is launched to minimize the reprojection error, and the 3D points are connected to form a mesh with N_F faces $\mathscr{F}$ and N_v vertices $\mathscr{V}$ given by the set of triangulated 3D points. Image consistent meshing can be used [7].

3 Monocular reconstruction with DSfM

We assume given N_c point correspondences between the deformed shape in an image and the 3D template. In the template, the correspondences are given by their barycentric coordinates $\{(f_i\ b_{1i}\ b_{2i}\ b_{3i})^\top\}$, $i = 1, \ldots, N_c$, relatively to the triangle they lie on. In the image, the correspondences are given in pixel coordinates $\{(u_i w_i)^\top\}$, $i = 1, \ldots, N$. Extensible 3D reconstruction can be formulated as

$$\begin{cases} \min\limits_{\mathscr{V}'} \left\{ \sum_{i=1}^{N_F} \parallel \mathsf{S}_i - \mathsf{S}_i^0 \parallel^2 + \lambda_1 \sum_{i=1}^{N_F} \parallel \mathsf{A}_i - \mathsf{A}_i^0 \parallel^2 + \lambda_2 \parallel \Delta \mathscr{V}' \parallel^2 \right\} \\ \text{s.t. } \sum_{i=1}^{N_c} \parallel \mathsf{K}\, \mathsf{v}'(f_i) \begin{pmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{pmatrix} - \begin{pmatrix} u_i \\ w_i \end{pmatrix} \parallel = 0 \end{cases} \tag{1}$$

where $\mathsf{v}'(f_i)$ is the 3×3 matrix whose columns are the 3D coordinates of the vertices of face i. S_i and A_i are the 2D shearing and anisotropy scaling transforms from the template to the deformed i^{th} face. λ_1 and λ_2 are two real positive weights that tune the importance of the shearing, the anisotropy scaling and the smoothing energy term. The inextensible formulation enforces the edges of the triangles to remain constant when fitting the data correspondence constraint. In contrast, this weighted combination of non-isometric transforms relax the inextensible condition and allows to deal with local extensible deformations. S^0 and A^0 are average amounts of shearing and anisotropy scaling for each face of the 3D template mesh. They can be either learned from training data or experimentally set. Practically, normalized shearing and anisotropy scaling transforms are experimentally set and then scaled by the triangle area of each face f_i to obtain the transforms S_i^0 and A_i^0. The additional weighted energy term smoothes the deformed shape. It is expressed through the linear Laplace-Beltrami discrete linear operator Δ of dimension $N_v \times N_v$ [13]. $\mathscr{V}$ is an $N_v \times 3$ matrix which concatenates the 3D mesh vertices.

4 Experimental results

4.1 Implementation details

In our implementation, the correspondences are assumed known. The constrained optimization problem can be minimized after elimination to exactly satisfy the correspondence constraints. Another solution is to include the constraints as a penalty in the minimization. As is shown by experimental results, this solution allows one to fit the correspondence constraints if the minimization is well initialized. In our implementation, this initialization is done by computing a rigid transform of the initial 3D template to the camera frame of the target image. This rigid transform is estimated as the average rigid transform which maps the image correspondences to the 3D template mesh. Even though this rigid transform is a rough estimate, it gives good results regarding convergence of the subsequent minimization.

4.2 Results and discussions

To validate the proposed approach, the experiment we propose is the 3D reconstruction of both inside abdominal wall and uterus from in-vivo sequences. The 3D template of the abdominal wall and the uterus are generated during the laparoscopy exploration step of the inside body. A set of extensible deformation upon the abdomen are made by the surgeon to introduce the trocar so that the surgery tool can go through it. Another set of deformation may occur on the uterus when the surgeon starts to examine it. In either the first or the second deformation, two different frames of monocular 3D reconstructions are shown in figure 2. It can be seen that the extensible uterus deformation by the tool was recovered. It can be seen also that the abdominal wall bends under the pressure of the external hand of the surgeon when he is inserting the trocar.

5 Conclusion

In this paper, we have presented a new method to reconstruct a quasi-conformal deforming living tissue in 3D using a single laparoscopic image and a 3D template that is previously reconstructed using standard RSfM. We believe that our method provides novel technical contributions and also a new way of tackling the 3D vision problem in laparoscopy. The in-vivo experimental results are very encouraging. We are currently working on improving the step of points matches computation between the 3D template and the current laparoscopic view.

References

[1] A. Bartoli, V. Gay-Bellile, U. Castellani, J. Peyras, S. Olsen, and P. Sayd. Coarse-to-fine low-rank structure-from-motion. *CVPR*, 2008.

[2] F. Brunet, R. Hartley, A. Bartoli, N. Navab, and R. Malgouyres. Monocular template-based reconstruction of smooth and inextensible surfaces. *ACCV*, 2010.

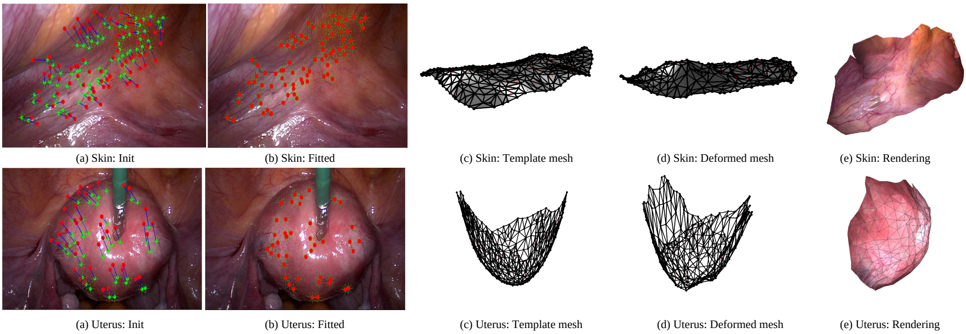

Figure 2: First row: the in-vivo abdominal wall (referred to as skin) experiment. Second row: The in-vivo uterus experiment. From left to right: (a) The initial distance between back projected template and image correspondences is shown with blue segments. Red points are the target image matches and green stars refers to the back projected correspondences from the 3D template. (b) After optimization, the template correspondences are fitted to the target image and the equality constraint is fulfilled. (c) The initial templated 3D mesh and (d) the deformed mesh. The deformed area clearly appears in the image.

[3] O. G. Grasa, J. Civera, and J. Montiel. EKF monocular slam with relocalization for laparoscopic sequences. *ICRA*, 2011.

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

[5] M. Hu, P. Penney, D. Rueckert, P. Edwards, F. Bello, R. Casula, M. Figl, and D. Hawkes. Non-rigid reconstruction of the beating heart surface for minimally invasive cardiac surgery. *MICCAI*, 2009.

[6] F. Moreno-Noguer, M. Salzmann, V. Lepetit, and P. Fua. Capturing 3D stretchable surfaces from single images in closed form. *CVPR*, 2009.

[7] D. D. Morris and T. Kanade. Image-consistent surface triangulation. *CVPR*, 2000.

[8] P. Mountney, D. Stoyanov, A. Davison, and G. Yang. Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery. *MICCAI*, 2006.

[9] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. A complete augmented reality guidance system for liver punctures: First clinical evaluation. *MICCAI*, 2005.

[10] M. Perriollat, R. Hartley, and A. Bartoli. Monocular template-based reconstruction of inextensible surfaces. *BMVC*, 2008.

[11] M. Salzmann, R. Hartley, and P. Fua. Convex optimization for deformable surface 3-D tracking. *ICCV*, 2007.

[12] M. Salzmann, R. Urtasun, and P. Fua. Local deformation models for monocular 3D shape recovery. *CVPR*, 2008.

[13] D. Yoo. Three-dimensional morphing of similar shapes using a template mesh. *International Journal of Precision Engineering and Manufacturing*, 2009.