31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Soft Tissue Deformation Tracking for Robotic Assisted Minimally Invasive Surgery

Danail Stoyanov and Guang-Zhong Yang

Abstract— **This paper presents a new framework for tracking soft tissue deformation in robotic assisted minimally invasive surgery. The method combines optical feature tracking based on stereo-laparoscope images and a constrained geometrical surface model that deforms with feature motion. This has the advantage of relying on reliable salient feature tracking while embedding underlying constraints on the tissue surface for deriving consistent temporal deformation. The proposed framework is resilient to occlusions and specular highlights. The accuracy and robustness of the proposed method are validated using a phantom heart model with known ground truth. To demonstrate the practical value of the method, example *in vivo* results are also provided.**

I. Introduction

Robotic assisted minimally invasive surgery (MIS) is increasingly being used to perform procedures on deformable anatomical structures. In cardiac surgery, totally endoscopic coronary artery bypass graft (TECAB) procedures can deliver many of the patient benefits commonly associated with MIS. However, respiratory and cardiac induced epicardial motion can make delicate anastomosis tasks difficult. With recent developments in port access techniques, the effect of physiological tissue motion can be observed in other surgical procedures within the peritoneal cavity [1]. Tissue motion can complicate instrument control during interventional tasks. It also makes pre- and intra-operative image-guidance difficult. One potential solution is the use of motion stabilization via robotic instruments. Reliable tracking the soft-tissue deformation during the course of intervention is an important pre-requisite [2].

Recently, the computation of structure and motion from stereo-laparoscope images as shown in Fig 1 has made marked progress [3,4,5]. Most approaches for dense stereo tracking are based on an underlying surface model using a geometric representation of the soft-tissue such as thin-plate splines [4] or a bilinear map [3]. The model is then deformed subject to a image similarity measure between the new image and a reprojection of the model template. While this approach can provide accurate information about soft tissue deformation, its robustness is dictated by the presence of prominent textural information or dense salient features. In practice, the method is susceptible to view dependent specular highlights. Another problem associated with the existing technique is that occlusion due to the surgical instruments is not properly handled during reprojection. Typically, the final solution is governed by the presence of salient image features but these are not explicitly considered and may be erroneous and scarcely distributed.

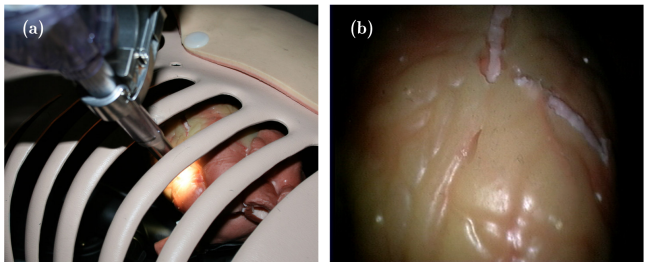

Fig. 1. (a) Stereo-laparoscope of the daVinci surgical system observing a silicone heart phantom model used for validation in this study; (b) image of the phantom from one camera of the scope.

More recently, techniques for the recognition and tracking of deformable surface using a small number of correspondences have been investigated [6,7,8]. The advantage of these methods is that any sparse matching strategy can be used, thus allowing for the use of invariant feature descriptors, which are important for robust performance in the presence of large deformations [9]. Occlusions can be handled naturally by avoiding correspondences that are potentially erroneous during the tracking process. In addition, due to the optimization methods used, additional constraints on the surface orientation can be incorporated. This greatly facilitates the handling of specular highlights.

In this study, a method for tracking tissue deformation using a monocular endoscope is proposed. It is robust to dynamic effects and performs well even for surfaces with homogenous appearance. The method is based on a geometric surface representation but it does not optimize the full reprojection of the model in the image space. Instead, the surface deformation is inferred from a reliable set of tracked salient feature points, which may be obtained using any reliable feature-based approach. This allows the integration of reliable image measurements with a surface model using a robust selection scheme. Experiments with both phantom data and *in vivo* video sequence demonstrate the potential clinical value of the method.

Manuscript received April 7, 2009 and revised June 20, 2009. This work was supported in part by the DTI Robotic Visualisations Project and the EC FP7 Project ARAKNES (224565).

D. Stoyanov is with the Institute of Biomedical Engineering and the Department of Biosurgery and Surgical Technology, Imperial College London, SW7 2AZ, UK (phone: +44-207-5940805; fax: +44-207-5940805; e-mail: danail.stoyanov@ imperial.ac.uk).

G.-Z. Yang is with the Institute of Biomedical Engineering and the Department of Computing, Imperial College London, SW7 2AZ, UK

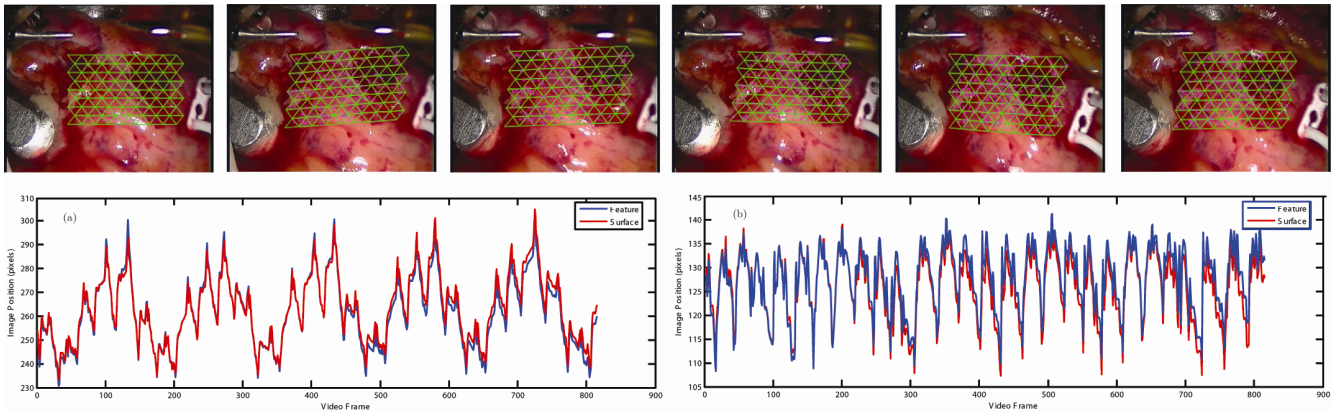

Fig. 2. Example images from a TECAB sequence using a daVinci surgical system. The triangulated mesh tracked by using the method proposed in this study is superimposed and features tracked using the method in [12] and used to propagate the surface mesh are also shown in purple; (a) shows the recovered motion for a feature along the X axis and (b) shows the motion of the feature along the Y axis.

II. Geometric Surface Model and Tracking

A. Soft-Tissue Surface Model

In this study, we represent the soft-tissue surface model M as a triangulated mesh consisting of hexagonal elements as illustrated in Fig 2. Each vertex v_i of the mesh is specified by its image coordinates (x_i, y_i) and the collection of all vertices defines a state vector S which controls the shape and position of M. Any image point p_i within the mesh can be described by its barycentric coordinates $B_i(p)$ and the vertices of the facet that it occupies (v_1, v_2, v_3). When the model undergoes motion or deformation, the new position of the point p_i can be computed using the mapping $T_S(p)$ defined as:

$$T_S(p) = \sum_{i=1}^{3} B_i(p) \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (1)$$

B. Surface Tracking

The goal of the proposed tracking framework is to determine the deformation and movement of M such that an observed set of feature matches within the surface mesh align correctly according to (1). This is performed as proposed in [6] by minimizing the following objective function $\varepsilon(S)$ subject to the parameter of S:

$$\varepsilon(S) = \lambda \varepsilon_D(S) + \varepsilon_C(S) \quad (2)$$

where $\varepsilon_D(S)$ defines deformation energy which regularizes the position of vertices and $\varepsilon_C(S)$ defines the correspondence error between the sparse set of features matching between the images. The weight term λ is used to determine the influence of regularization. In this study, we set it to be 100/length of edges at start. Both $\varepsilon_D(S)$ and $\varepsilon_C(S)$ can be formulated as Eucledian distance measures. Therefore the minimization of $\varepsilon(S)$ can be performed via iterative non-linear least squares techniques such as the Levenberg-Marquardt algorithm.

The effect of tracking the deformation of M is illustrated in Fig 2 where the initial surface model is tracked in subsequent frames to recover the cardiac and respiratory induced epicardial surface motion. Features used to drive the framework are shown within the triangulation. In the following sections, the terms of (2) will be discussed in more detail.

III. Tracking Constraints

A. Deformation Constraints

In general, the surface mesh described in the previous section is regular. Therefore, the deformation energy $\varepsilon_D(S)$ can be formulated as a geometric regularization constraint. Given all vertices within the mesh, we may consider subsets of vertices that define a regular hexagon. The regularization term can be formulated to keep the distance between collinear vertices equidistant [10]. This has the effect of pulling the overall energy in the tracking process to preserve the original regular starting point of the surface. Non-linear deformations are penalized by using $\varepsilon_D(S)$, however, this only affects the immediate locality of the vertices. The overall deformation can therefore accommodate a wide range of changes in surface shape.

Formulating $\varepsilon_D(S)$ in this manner is effective in practice even though it does not directly consider biomechanical tissue properties. It enables fast computation of the regularization process, thus real-time implementation. For a collinear set of vertices v_i, v_j, v_k, we can write the deformation energy as:

$$\varepsilon_D(S) = \frac{1}{2} \sum_{i,j,k \in E} \begin{matrix} (-x_i + 2x_j - x_k)^2 \\ +(-y_i + 2y_j - y_k)^2 \end{matrix} \quad (3)$$

where E represents all vertex indices i, j, k that are collinear. This constraint can be succinctly written in terms of a matrix multiplication for faster computation [8,10].

B. Visual Constraints on Soft-Tissue Dynamics

The feature correspondence energy $\varepsilon_C(S)$ in (2) is used to ensure that the surface mesh deforms to optimally fit the matched feature between image frames. All feature correspondences found in the images C contain matches in the form of $c = [c_t, c_{t-1}]$, where the indices denote feature positions at times t and $t-1$, respectively. Following [7], the correspondence energy can be defined as:

$$\varepsilon_C(S) = -\sum_{c \in C} w_c \rho\left(\left\| c_1 - T_S(c_0) \right\|, r\right) \tag{4}$$

where function ρ is a robust estimator with radius r and w_c represents the weight given to each correspondence (in this study, we set all weight coefficients to 1). The use of a robust estimator is crucial to the handling of outliers due to mismatches and occlusions. In essence, this function limits the influence of potentially unreliable information during computation. Many robust estimators are available in the literature including Tunkey and Huber functions [11]. The estimator used in this study is defined as:

$$\rho\left(\delta, r\right) = \begin{cases} \dfrac{3(r^2 - \delta^2)}{4r^3} & \delta < r \\ 0 & otherwise \end{cases} \tag{5}$$

The main advantage of this ridge like operator is that it favors reliable correspondence and effectively discards less confident measurements [6].

IV. Deformation Tracking Framework

A. Feature Matching

To match features across the image set, we used the method developed in [12] but only considered 2D temporal motion. In principle, one of the strengths of the proposed method is that it can be integrated with any matching technique. This can facilitate difficult matching tasks in the presence of deformation and large camera motion. The matching process produces a set of temporal 2D-2D correspondences that are used to propagate to the entire soft-tissue surface according to (2).

B. Specular Reflections

Specular reflections during MIS can significantly affect the tracking performance. This is because the observed motion is not representative of the physiological tissue motion. Typically, highlights can be detected by using thresholding of the image intensity and saturation levels. While this technique has limited performance and relies on a threshold value dependent on the optical configuration, the minimal computational effort involved makes the method attractive for real-time applications. Subsequently, a dilation operation using a circular kernel is performed to compensate for dark ridges around the highlights. The resulting specular map is then used to remove features lying within these regions from subsequent feature correspondence.

V. Experiments and Results

For quantitative performance evaluation, the proposed method was implemented in C++ and a phantom heart model (The Chamberlaine Group, MA, USA) was scanned using a Siemens Somatom Sensation 64 CT Scanner. For registration between the dynamic CT frame and the video data, the model was embedded with 15 high contrast CT markers for deriving the rigid transformation. The performance of the algorithm with different numbers of mesh nodes is presented in Table 1. It is evident that the technique is fast and suitable for real-time performance on a laptop PC with a 2.2 GHz Intel® T7500 processor. The reported timings for feature tracking are stable with respect to the number of mesh nodes as the tracking was implemented over the full image space.

TABLE I
Performance Timing

Mesh Nodes	Total Time (ms)	Feature Tracking Time (ms)
25	7.8±1.7	15
49	27.4±6.0	15
81	80.4±23.7	15
121	228.1±53.8	15

A. Phantom Experiments

The proposed algorithm was applied to a video sequence of the phantom heart for tracking the motion of an observed fiducial. For quantitative comparison with the ground truth data, the coordinates of the fiducial trajectory from the CT image were reprojected back into the camera space and compared with the motion estimated from the video sequence. The results are illustrated in Fig 3 where it is evident that the proposed method can correctly estimate the motion of the fiducial. However, with time, there is incremental drift in the tracked coordinates. This is the result of minor error propagations in the mesh coordinates that can form with small misalignments due to feature matching error and noise. These are not explicitly handled with the proposed approach but can be addressed with more sophisticated corrective schemes. Other errors in Fig 3 are due to data segmentation and volumetric model extraction.

B. In Vivo Experiments

To demonstrate the potential clinical value of the technique, the proposed method has also been applied to *in vivo* sequences form different MIS procedures. The results for a TECAB procedure are shown in Fig 2, in which the algorithm is compared with the feature tracking approach presented in [12]. It is evident that the results are almost identical and the surface regularization term does not seem to cause large deviations.

In Fig 4, the result of tracking multiple triangulation patches by using the proposed technique is shown. This illustrates the capability of the method in dealing with multiple tissue surfaces and handling surface topological change due to dissection.

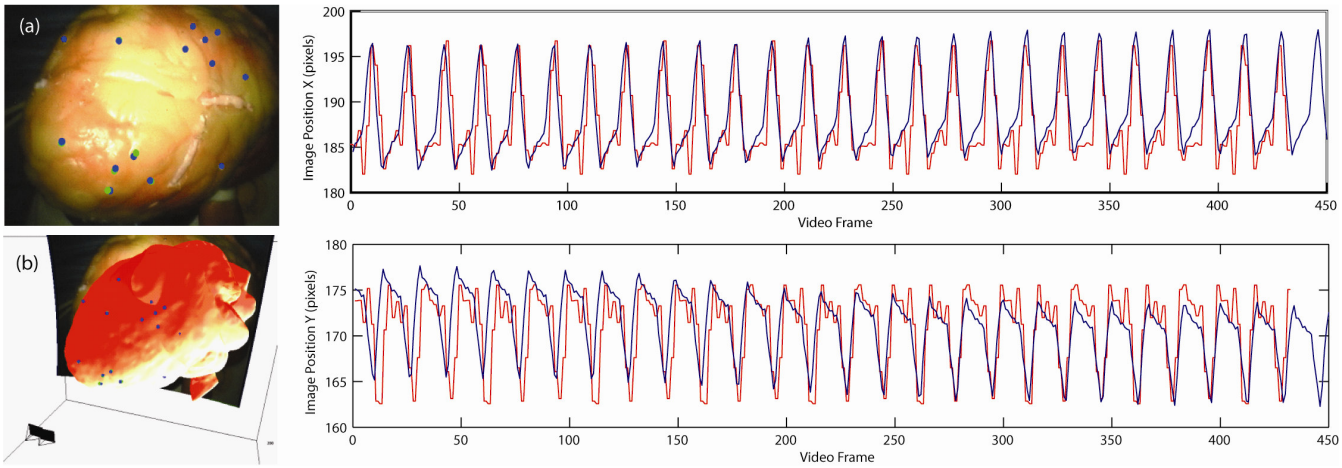

Fig. 3. (a) Laparoscopic image of the phantom heart model with CT fiducials rendered onto the image to align with the observed points; (b) 3D rendition of the phantom model and fiducials within the camera's coordinate system; (c-d) trajectory motion of a fiducial recovered using the proposed surface tracking approach shown in blue and compared to the data obtained from the CT ground truth shown in red.

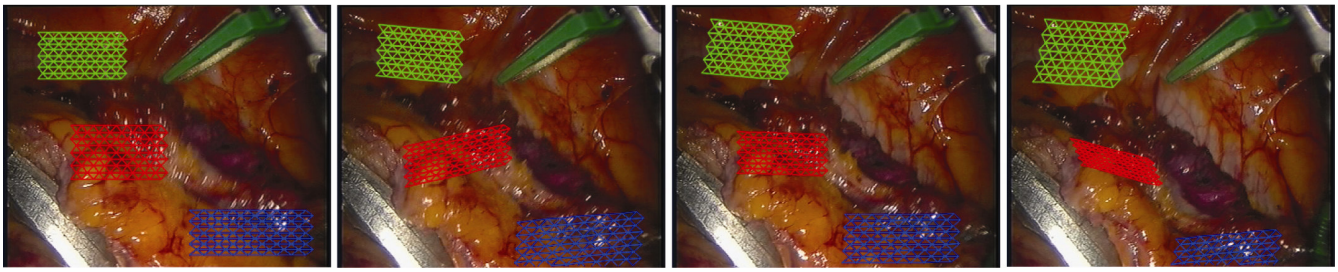

Fig. 4. This figure shows a series of laparoscopic images from a TECAB procedure using the daVinci® surgical system with multiple surface patches shown in different colors tracked as the tissue is affected by cardiac and respiratory motion. This illustrates the concept of tracking multiple surfaces using disparate mesh triangulations.

VI. Conclusion

In this study, we have presented a method for tissue deformation tracking by sparse salient features combined with geometric surface parameterization. Experimental results have shown that the technique is efficient and suitable for real-time tissue tracking *in situ*. Detailed error analysis indicates that the proposed technique can achieve performance comparable to other feature tracking techniques but with a larger surface coverage. Future directions we are pursuing include the use of probabilistic prediction for repetitive motions and the use of biomechanical tissue properties incorporating statistical deformation models [13].

References

[1] D. Stoyanov, G. Mylonas, A. J. Chung, M. Lerotic and G.-Z. Yang, " Intra-operative Visualisations: Perceptual Fidelity and Human Factors," *IEEE/OSA Journal of Display Technologies*, 2008.

[2] C. N. Riviere, J. Gangloff, and M. de Mathelin, "Robotic compensation of biological motion to enhance surgical accuracy," *Proceedings of the IEEE,* vol. 94(9), pp. 1705-1716, 2006

[3] D. Stoyanov, A. Darzi and G.-Z. Yang, "Dense 3D depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery," in *Medical Image Computing and Computer Assisted Intervention*, vol. II, pp. 41-48, 2004.

[4] W. W. Lau, N. A. Ramey, J. Corso, N. V. Thakor, and G. D. Hager, "Stereo-Based Endoscopic Tracking of Cardiac Surface Deformation," in *International Conference on Medical Image Computing and Computer Assisted Intervention*, vol. 3217, pp. 494-501, 2004.

[5] R. Richa, P. Poignet and C. Liu, "Efficient 3D Tracking for Motion Compensation in Beating Heart Surgery," in *International Conference on Medical Image Computing and Computer Assisted Intervention*, vol. II, pp. 684-691, 2008.

[6] J. Pilet, V. Lepetit, and P. Fua, "Real-time Non-Rigid Surface Detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 822-828, 2005.

[7] M. Salzmann, J. Pilet, S. Ilic, P. Fua, "Surface Deformation Models for Nonrigid 3D Shape Recovery," *IEEE Trans Pattern Anal Mach Intell*, vol. 29(8), pp. 1481-1487, 2007

[8] F.Moreno-Noguer, M.Salzmann, V.Lepetit, P.Fua, "Capturing 3D Stretchable Surfaces from Single Images in Closed Form," *to appear in IEEE Comp Society Conference in Computer Vision and Pattern Recognition*, 2009

[9] P. Mountney, B. P. L. Lo, S. Thiemjarus, D. Stoyanov, G.-Z. Yang, "A Probabilistic Framework for Tracking Deformable Soft Tissue in Minimally Invasive Surgery," in *International Conference on Medical Image Computing and Computer Assisted Intervention*, vol. 2, pp. 34-41, 2007

[10] P. Fua and Y. G. Leclerc, "Object-Centered surface reconstruction: Combining multi-image stereo and shading," in *Image Understanding Workshop*, pp. 1097-1120, 1993.

[11] M. Salzman, "Learning and Recovering 3D Surface Deformations," *PhD Thesis, Ecole Polytechnique Federale de Lausanne*, 2009.

[12] D. Stoyanov, G. Mylonas, F. Deligianny, A. Darzi and G.-Z. Yang, "Soft-Tissue Motion Tracking and Structure Estimation for Robotic Assisted MIS Procedures," in *Medical Image Computing and Computer Assisted Intervention*, vol. II, pp. 139-146, 2005.

[13] H. Delingette, X. Pennec, L. Soler, J. Marescaux, and N. Ayache, "Computational Models for Image-Guided Robot-Assisted and Simulated Medical Interventions," *Proceedings of the IEEE,* vol. 94, pp. 1678-1688, 2006.