Stabilization of Image Motion for Robotic Assisted Beating Heart Surgery

Danail Stoyanov and Guang-Zhong Yang

Institute of Biomedical Engineering,
Imperial College London, London SW7 2AZ, UK
{danail.stoyanov,g.z.yang}@imperial.ac.uk
http://vip.doc.ic.ac.uk

Abstract. The performance of robotic assisted minimally invasive beating heart surgery is a challenging task due to the rhythmic motion of the heart, which hampers delicate tasks such as small vessel anastomosis. In this paper, a virtual motion compensation scheme is proposed for stabilizing images from the surgical site. The method uses vision based 3D tracking to accurately infer cardiac surface deformation and augmented reality for rendering a motion stabilized view for improved surgical performance. The method forgoes the need of fiducial markers and can be integrated with the existing master-slave robotic consoles. The proposed technique is validated with both simulated surgical scenes with known ground truth and *in vivo* data acquired from a TECAB procedure. The experimental results demonstrate the potential of the proposed technique in performing microscale tasks in a moving frame of reference with improved precision and repeatability.

Keywords: Robotic Assisted Surgery, Beating Heart Surgery, Motion Compensation, Soft-Tissue Tracking.

1 Introduction

The use of master-slave robotic systems for minimally invasive surgery has made it possible to perform totally endoscopic beating heart surgery. This approach has a number of well documented patient benefits including reduced surgical trauma and avoidance of cardiopulmonary bypass. However, handling respiratory and cardiac motion under which these procedures must be performed is demanding even for skilled surgeons. Thus far, delicate tasks such as small vessel anastomosis during Totally Endoscopic Coronary Artery Bypass (TECAB) are normally performed under mechanical stabilization. Despite this, residual motion is still prominent, which is compounded by limited epicardial exposure and the complexity of instrument control. A potential solution to this is to introduce real-time *in situ* adaptive motion compensation.

Motion compensation involves synchronizing the movement of a robotic instrument with the deformation of the soft-tissue, hence allowing the surgeon to operate on a *virtually* stabilized operating field-of-view. Previous research has

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part I, LNCS 4791, pp. 417–424, 2007.

demonstrated that by synchronizing instrument movement, it is possible to improve the accuracy of simulated surgical tasks [1]. Control requirements for a practical robotic device to operate in sync with the heart have also been investigated, illustrating the requirement for predictive control loops [2, 3, 4]. The pre-requisite of all motion stabilization techniques is accurate reconstruction of tissue deformation *in situ* and in real-time. Thus far, a number of approaches have been proposed to infer 2D deformation of the heart *in vivo* either by using projected fiducial markers [2] or by optical tracking techniques [3]. It has also been shown that the epicardial surface can be reconstructed in metric 3D space based on stereo-laparoscopes [5]. There is, however, limited investigation on the psychovisual and hand-eye coordination of the surgeon under such motion compensation schemes. The purpose of this paper is to present a vision based image stabilization scheme for motion compensation that combines real-time computational stereo tissue deformation recovery with 3D motion stabilization. Our study focuses on the visualization of the motion stabilized stereo-laparoscope images and does not describe a robotic tool synchronization system. The proposed technique forgoes the need of fiducial markers and can be integrated with an Augmented Reality (AR) scheme on a master-slave robotic console. The method is validated with both simulated surgical scenes with known ground truth and *in vivo* data acquired from a TECAB procedure.

2 Method

2.1 3D Tissue Deformation Recovery

For real-time motion compensation, accurate 3D tissue deformation recovery is essential. In order to compute 3D measurements using a stereo-laparoscope, the device is first calibrated using an existing planar object calibration algorithm [6]. This computes the intrinsic camera parameters and the spatial relationship between the stereo rig. The camera matrices and mapping between a world point $\mathbf{M} = [X \quad Y \quad Z]^{\mathsf{T}}$ and an image point $\mathbf{m} = [x \quad y]^{\mathsf{T}}$ can be represented by the following equation:

$$\lambda \mathbf{m}_n = \mathbf{K}_n [\mathbf{R}_n \mid \mathbf{t}_n] \mathbf{M} \tag{1}$$

where $\mathbf{K}$ is the intrinsic parameter matrix, $\mathbf{R}$ and $\mathbf{t}$ are the extrinsic parameters, and λ is a scale factor. When these parameters are known, the 3D position of a landmark visible from both stereo-laparoscope views can be computed by finding the intersection of the rays back-projected from each camera. In practice, the true intersection point may not exist and a surrogate measure of identifying the mid-point of the shortest line segment between the rays is usually used.

For this purpose, the correspondence between image points in the two views must be established. To obtain stereo correspondence and subsequently track temporal tissue motion, we used a technique that extends previous work proposed in [7]. With this method, salient features are first selected on the epicardial surface based on the intensity gradient information [8]. They are then tracked in real-time using stereo-temporal constraints using the Lucas-Kanade (LK) algorithm. By denoting a warp

function $\mathbf{W}_n(\mathbf{m};\mathbf{p})$, which maps point $\mathbf{m}$ according to the parameter vector $\mathbf{p}$, the alignment of an image template $T_n(\mathbf{m})$ with an input image I_n in more than one camera can be formulated as the following minimization problem:

$$\sum_n \sum_{\mathbf{m}} [I_n(\mathbf{m} + \mathbf{P}_n(\mathbf{p})) - T_n(\mathbf{m})]^2 \tag{2}$$

The selected parameterization of the warp function represents the 3D location of the target point $\mathbf{p} = \mathbf{M}$ and has an analytical *Jacobian* [9]. From a given starting solution, the LK parameter update term can be derived iteratively by approximating Eq. (2) with a first-order Taylor polynomial and obtaining the partial derivative with respect to $\Delta\mathbf{p}$. This leads to the following update equation:

$$\Delta\mathbf{p} = \mathbf{H}^{-1} \sum_n \sum_{\mathbf{m}} \left[\nabla I_n \frac{\partial \mathbf{W}_n}{\partial \mathbf{p}} \right]^{\mathsf{T}} [T_n(\mathbf{m}) - I_n(\mathbf{W}_n(\mathbf{m};\mathbf{p}))]$$
$$\text{where } \mathbf{H} = \sum_n \sum_{\mathbf{m}} \left[\nabla I_n \frac{\partial \mathbf{W}_n}{\partial \mathbf{p}} \right]^{\mathsf{T}} \left[\nabla I_n \frac{\partial \mathbf{W}_n}{\partial \mathbf{p}} \right] \tag{3}$$

To obtain a starting solution, the same method as above was used but implemented in a pyramidal fashion to match across the stereo pair. This approach is effective as the hierarchical scheme allows for large disparity values. It is also particularly useful for improving the convergence of the algorithm for stereo-laparoscopes, where the cameras are arranged in a verged configuration. For feature detection, gradient based landmarks were used for improved computational performance suitable for real-time implementation [8]. To circumvent the effect of specular highlights, a threshold filter based on the intensity and saturation of the image was used to remove the associated outliers.

2.2 Motion Compensated Imaging

Once the motion of the target cardiac surface is reconstructed, virtual motion compensation is applied. To compensate for physiological motion, we use an AR scheme by creating a virtually moving camera, which renders the surgical field-of-view in a moving frame of reference.

For the calibrated stereo-laparoscope cameras, a ray $\mathbf{q}_n(\mathbf{m}_n)$ defining the line of sight for an image point $\mathbf{m}_n$ can be derived by using the camera matrix and the optical centre of the camera as:

$$\lambda \mathbf{q}_n(\mathbf{m}_n) = \mathbf{c}_n + \mathbf{P}_n^{\dagger} \mathbf{m}_n \tag{4}$$

where $\mathbf{c}_n$ denotes the optical centre and $\mathbf{P}_n^{\dagger}$ the Moore-Penrose pseudo inverse of the projection matrix. A virtual surface rendered with a texture taken from the stereo-laparoscope will appear as the real scene in a virtual camera representing the stereo-laparoscope if all points on the surface satisfy Eq. (4). This effectively ensures the alignment of the virtual scene with the real images.

Compensating for the motion of a moving target on the cardiac surface can be achieved by directly moving the virtual camera with the respective motion of the

target. Denoting the target motion as $\mathbf{J}^t = \mathbf{M}^t - \mathbf{M}^{t-1}$ and using Eq. (1), the position of the compensated camera must satisfy the following equation:

$$\mathbf{P}_n \mathbf{M}^t = \mathbf{P}'_n \left(\mathbf{M}^t + \mathbf{J}^t \right) \tag{5}$$

By using the decomposition of the camera matrix, it is clear that the intrinsic parameters remain constant and the above equation leads to an adjustment of the camera's extrinsic translation parameter as $\mathbf{t}_n^t = \mathbf{R}_n \mathbf{J}^t - \mathbf{R}_n \mathbf{c}_n^{t-1}$.

For an exact virtual scene representation from the new camera position, it is necessary to have the knowledge about the 3D geometry and photometric properties of the object. The scene geometry can be obtained with dense computational stereo techniques but this can be computationally demanding. In this study, an approximation of the cardiac surface was used instead in the vicinity of the compensated point by using a thin plate spline. The sparse set of landmarks around the compensated target was determined in §2.1 and their corresponding distance along rays defined in Eq. (4) was used as the control point to recreate the entire surface. The remainder of the surface was then interpolated back to a virtual plane satisfying Eq. (4). It should be noted that this interpolation does not necessarily represent the true perspective projection.

3 Experimental Setup and Results

In order to validate the proposed motion compensation approach, both simulated and *in vivo* data was used. In this study, the experimental environment consisted of a daVinci® console used for 3D visualization, a SenseAble Omni® haptic device for instrument manipulation, and a 3 GHz Pentium-D workstation with two NVIDIA GeForce 7600Gs graphics cards and 2 Gb of RAM for real-time processing. The proposed method was implemented using C++ and D3D for real-time 3D tissue-deformation recovery and motion compensated AR visualization. When motion compensation was applied the virtual instrument controlled by the haptic device was synchronized with the same motion signal as applied to the virtual camera.

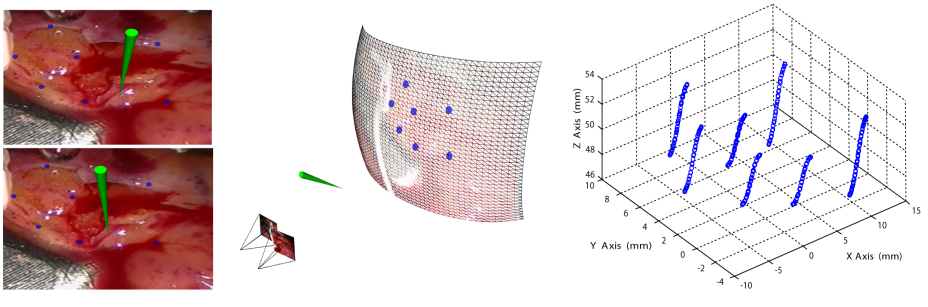


Fig. 1. The simulation environment used in this study, illustrating a stereo pair generated by the virtual camera (left), the corresponding virtual 3D scene and surface mesh (middle), and the 3D motion trajectories (right) of the highlighted control points used for validation

Ten volunteers (2 surgeons and 8 computer scientists) were asked to perform two separate tasks with and without motion compensation, respectively. *Task I* represents a simple targeting exercise and involves touching a sequence of markers on the moving surface as shown in Fig. 1. *Task II* simulates an incision required to expose an anastomosis site, which requires the user to follow a smooth path through a series of markers on the beating heart surface.

3.1 Validation with Simulated Data

For validation with simulated data, a 3D surface mesh textured with an *in vivo* image was used to represent the deforming cardiac surface. The dynamics of the motion were modeled by a sinusoidal Gaussian mixture function mimicking the respiratory and cardiac induced epicardial deformation. An example rendition of the virtual surface as observed from the stereo cameras is shown in Fig. 1.

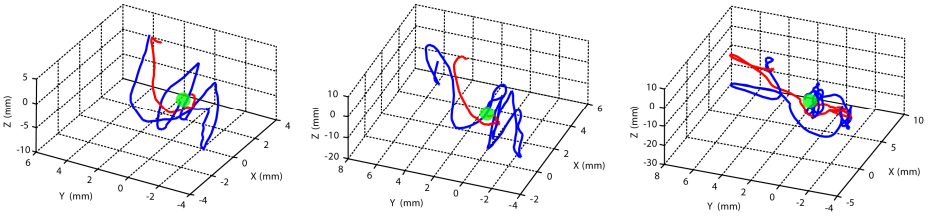


Fig. 2. Example trajectories in the moving frame of reference of a single target (green) for the instrument tip with (red) and without (blue) motion compensation for three of the subjects studied

The results obtained for the ten subjects performing *Task I* are summarized in Table 1 using the total path length (TPL) of the tooltip's motion, which is representative of the amount of movement required to reach the target. Example trajectories presented in the moving frame of reference of the target are shown in Fig. 2. It was found that with motion compensation, all subjects could reach the target with a faster and smoother trajectory compared to that without motion compensation.

For *Task II*, Table 1 summarizes the results for the ten subjects, where the mean path error (MPE) and its standard deviation (SPE) are used as performance measures. In this study, the path error was defined as the shortest Eucledian distance from the tooltip to the optimal path between the control points. It is evident that with motion compensation, there is a significant reduction in the MPE for most subjects along with a significant reduction in SPE. This indicates that subjects are much more confident and accurate in following the prescribed path along the cardiac surface, which is essential for exposing the anastomosis site.

To verify the statistical significance of the results presented in Table 1 we used a Kolmogorov-Smirnov test to find that the data did not adequately fit a Normal distribution. Therefore, we used the non-parametric Wilcoxon test for paired samples, which showed that the difference in results observed with and without motion compensation was significant for all performance metrics. The *p-value* given by the Wilcoxon test for TPL, MPE and SPE was 0.0273, 0.002 and 0.0098 respectively.

Table 1. Results obtained for the ten sujects performing *Tasks I* and *Task II*. The prefix (c) is used to distiguish the compensated version of the experiment and all measurements are in millimeters.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Median [quartiles]
TPL	433.6	586.0	944.6	608.5	281.6	551.5	510.6	346.3	474.2	520.6	**515.6 [412, 592]**
cTPL	239.6	244.9	312.0	325.0	179.2	324.1	755.7	193.5	207.1	324.0	**278.4 [204, 324]**
MPE	2.43	2.79	2.45	2.20	1.80	2.48	3.41	1.85	1.48	1.95	**2.32 [1.84, 2.56]**
cMPE	1.48	1.67	1.72	0.97	1.07	1.42	3.04	1.13	0.82	1.59	**1.45 [1.04, 1.68]**
SPE	1.31	1.87	1.55	1.46	0.97	1.21	2.04	1.04	0.78	1.22	**1.27 [1.02, 1.63]**
cSPE	0.60	1.14	1.05	0.58	0.57	0.87	2.12	0.73	0.57	1.30	**0.8 [0.58, 1.18]**

3.2 Validation with *In Vivo* Data

For demonstrating the practical value of the proposed framework, an *in vivo* dataset from a TECAB surgery was used for further validation. In this experiment, the ground truth geometry was derived by tracking landmarks on the epicardial surface for the duration of the *in vivo* sequence (2100 frames) with the aforementioned 3D surface reconstruction technique. The reconstructed dense surface map and the corresponding 3D motion trajectory of an example tracked landmark are illustrated in Fig. 3. This result was then used as the ground truth 3D geometry as in previous experiments.

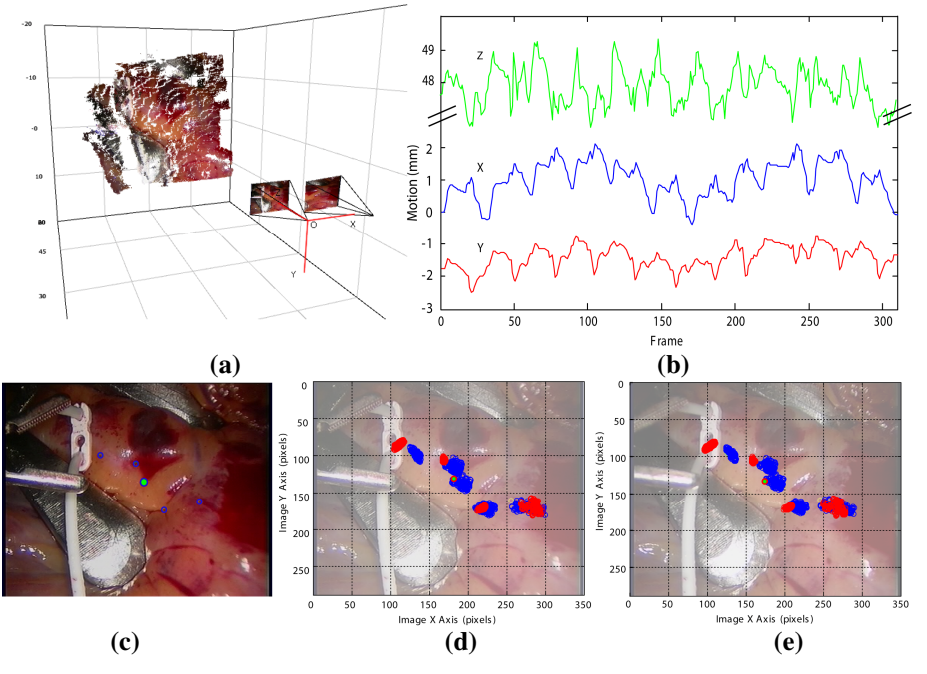
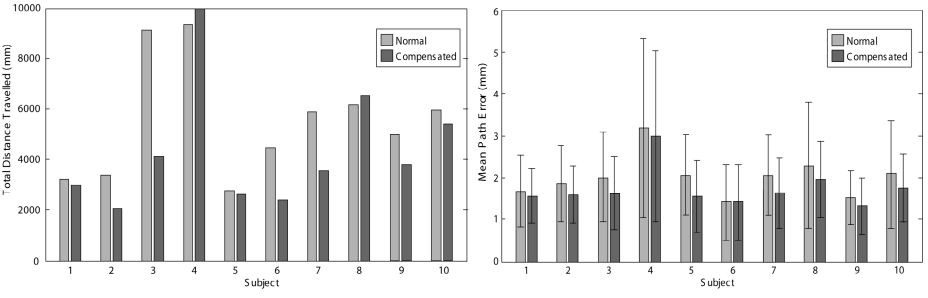


Fig. 3. (a-b) Dense cardiac surface reconstruction in 3D and the corresponding 3D motion components of the stabilization target along the *x*, *y*, and *z* axes **(c)** The reconstructed ground truth control points (blue circles) and the zero motion target (green). **(d-e)** Stereo pair from **(c)** showing the motion of each control point with (red) and without (blue) motion compensation.

Fig. 3 demonstrates the image motion of the control points back-projected to the camera view prior to (shown in blue) and after (shown in red) applying the proposed motion compensation scheme. It is apparent that the motion of the highlighted target point is close to zero, whereas the motion of the neighboring region is significantly reduced. The overall deformation of the surface is highly non-linear and thus there is residual motion.

Fig. 4 illustrates the results for the ten subjects performing *Task I* based on the *in vivo* data. It is evident that there is a marked improvement in performance for most of the subjects. This difference is statistically significant as demonstrated by a non-parametric Wilcoxon test for paired samples where the *p-value* was computed as 0.02 and 0.001 for the TPL and MPE metrics respectively.



(a) **(b)**

Fig. 4. (a) Bar chart illustrating the difference in the total path length in order to hit all targets during *Task I* with and without motion compensation. **(b)** The mean and standard deviation of the error between the tooltip and the ideal path for *Task II* for all the studied subjects.

4 Discussion and Conclusions

In conclusion, we have presented in this paper a method for motion compensation in robotic assisted beating heart surgery. The method uses vision based 3D tracking to accurately infer cardiac surface deformation and AR for rendering a motion stabilized view for improved performance of micro-scale surgical tasks such as vessel anastomosis. The experiments carried out in this study demonstrate the potential of the proposed technique in performing these delicate tasks in a moving frame of reference with greater precision and repeatability. An important area of future work is to develop methods that can adaptively select motion compensation target and seamlessly integrate the proposed motion compensation scheme into the current robotic assisted MIS workflows. Potential techniques include the use of tooltip tracking or to use real-time eye tracking to locate the surgeon's fixation point, thus leading to a gaze contingent adaptive motion compensation framework.

References

1. Trejos, A.L., Salcudean, S.C., Sassani, F., Lichtenstein, S.: On the Feasibility of a Moving Support for Surgery on the Beating Heart. In: Taylor, C., Colchester, A. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 1088–1097. Springer, Heidelberg (1999)

2. Ginhoux, R., Gangloff, J., De Mathelin, M., Soler, L., Sanchez, M.A., Marescaux, J.: Active Filtering of Physiological Motion in Robotized Surgery Using Predictive Control. IEEE Trans. Robotics 21(1), 67–79 (2005)
3. Ortmaier, T., Groger, M., Boehm, D.H., Falk, V., Hirzinger, G.: Motion Estimation in Beating Heart Surgery. IEEE Trans. Biomedical Engineering 52(10), 1729–1740 (2005)
4. Nakamura, Y., Kishi, K., Kawakami, H.: Heartbeat Synchronization for Robotic Cardiac Surgery. In: Proc. ICRA 2001, vol. 1, pp. 2014–2019 (2001)
5. Stoyanov, D., Darzi, A., Yang, G.-Z.: Dense 3D Depth Recovery for Soft Tissue Deformation During Robotically Assisted Laparoscopic Surgery. Computer Aided Surgery 10(4), 199–208 (2005)
6. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. on Pattern Analysis and Machine Intelligence 22(11), 1330–1334 (2000)
7. Stoyanov, D., Mylonas, G., Deligianni, F., Yang, G.-Z.: Soft-Tisse Motion Tracking and Structure Estimation for Robotic Assisted MIS Procedures. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3750, pp. 139–146. Springer, Heidelberg (2005)
8. Shi, J., Tomasi, C.: Good Features to Track. In: CVPR 1994, vol. 1, pp. 593–600 (1994)
9. Devernay, F., Mateus, D., Guilbert, M.: Multi-Camera Scene Flow by Tracking 3D Points and Surfels. In: CVPR 2006, vol. 2, pp. 2203–2212 (2006)