

Efficient 3D Tracking for Motion Compensation in Beating Heart Surgery

Rogério Richa, Philippe Poinet, and Chao Liu

LIRMM - UMR 5506 CNRS - UM2
161 Rue Ada 34392 Montpellier, France
{rogerio.richa, philippe.poinet, chao.liu}@lirmm.fr

Abstract. The design of physiological motion compensation systems for robotic-assisted cardiac Minimally Invasive Surgery (MIS) is a challenging research topic. In this domain, vision-based techniques have proven to be a practical way to retrieve the motion of the beating heart. However due to the complexity of the heart motion and its surface characteristics, efficient tracking is still a complicated task. In this paper, we propose an algorithm for tracking the 3D motion of the beating heart, based on a Thin-Plate Splines (TPS) parametric model. The novelty of our approach lies in that no explicit matching between the stereo camera images is required and consequently no intermediate steps such as rectification are needed. Experiments conducted on *ex-vivo* and *in-vivo* tissue show the effectiveness of the proposed algorithm for tracking surfaces undergoing complex deformations.

Keywords: tracking, stereo, thin-plate splines, robotic assisted surgery.

1 Introduction

In recent years, a great effort has been expended on the design of physiological motion compensation techniques for robotic-assisted surgery. Nowadays, miniature versions of mechanical stabilizers make total endoscopic coronary grafting possible but the residual motion due to insufficient immobilization has to be manually canceled by the surgeon. This consequently prolongs the operating time, resulting in greater costs. Robotic assistance could aid surgeons by creating a virtually stable environment where the robotic instruments move in synchrony with the heart motion, giving them the feeling to be operating on a static organ. Recent studies display the potential improvements of the precision and repeatability of their gestures when working on a stabilized environment [1].

For the construction of a virtually stable environment for the surgeon, the motion of the beating heart must be accurately measured. A practical solution for this problem is the use of vision-based techniques, which eliminates the need for introducing additional sensors in the surgeon's limited workspace. However, due to the heart's complex dynamics and surface characteristics, tracking the motion of the beating heart is a complicated task. The feasibility of vision-based systems for motion compensation was first explored in the work of Nakamura

et al. [2], where fiducial markers were fixed on a pig's heart and tracked using high-speed cameras. However, for practical reasons the use of markers are not desirable in real surgical procedures. An alternative solution is the use computer vision techniques for inferring the heart motion based solely on natural structures on the heart surface.

In the literature, several approaches for tracking the heart surface in 3D using stereo endoscopes can be found. Tracking salient features on the heart surface has been a popular choice and although several works [3,4,5] explore feature tracking, this class of techniques has limited performance when features undergo large deformations. Another class of methods based on region tracking [6,7] has the potential to display better results if a suitable motion model for the heart is considered. In this paper, we describe an efficient algorithm for tracking the beating heart using a Thin-Plate Spline (TPS) based parametric model for the heart surface. Inspired by the integration of stereo correspondence and surface reconstruction [8], we formulate the tracking problem without the need for explicit matching between stereo pairs and therefore no intermediate steps such as rectification are required. Tracking is formulated as an iterative registration algorithm and solved using the Efficient Second-order Minimization (ESM) algorithm [9]. The method is validated on experiments conducted on *ex-vivo* and *in-vivo* data.

2 Methods

Thin-Plate Spline (TPS) transformations have been successfully applied to model non-rigid deformations [10,11,12] for numerous applications. In this paper, we extend the TPS transformations for 3D tracking in a stereo framework. In our development, we follow the formulation proposed by Lim [12].

The Thin-Plate Spline is a radial basis function (RBF) that specifies an approximation function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ that minimizes its 'bending energy'. The function f for a point $\mathbf{x} = (x, y)$ in a plane is defined by the Thin Plate Spline basis function $U(s) = s^2 \log(s^2)$, a $(n+3)$ parameter vector $\mathbf{t} = (w_1, \dots, w_n, r_1, r_2, r_3)^T$ and a set of n control points $\mathbf{c} = (\check{x}, \check{y})$, such that:

$$f(\mathbf{x}) = r_1 + r_2x + r_3y + \sum_{i=1}^n w_i U(\|\mathbf{c}_i - \mathbf{x}\|) \quad (1)$$

Side conditions $\sum_{i=1}^n w_i = \sum_{i=1}^n w_i \check{x}_i = \sum_{i=1}^n w_i \check{y}_i = 0$ must be considered for $f(\mathbf{x})$ to have square-integrable second derivatives. To write a $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ mapping m we stack two RBFs f^x and f^y sharing their control points:

$$m(\mathbf{x}) = \begin{bmatrix} f^x \\ f^y \end{bmatrix} = \begin{bmatrix} r_2^x & r_3^x & r_1^x \\ r_2^y & r_3^y & r_1^y \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + \sum_{i=1}^n \begin{bmatrix} w_i^x \\ w_i^y \end{bmatrix} U(\|\mathbf{c}_i - \mathbf{x}\|)$$

A linear system can be written to estimate the parameter vectors \mathbf{t}^x and \mathbf{t}^y that define f^x and f^y based on the mapped values $\mathbf{c}' = m(\mathbf{c})$ of each control point \mathbf{c} . Stacking the control point correspondences \mathbf{c}' into a matrix \mathbf{P}'_c yields:

$$\begin{bmatrix} \mathbf{L} & \mathbf{P}_c \\ \mathbf{P}_c^T & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{t}^x & \mathbf{t}^y \end{bmatrix} = \begin{bmatrix} \mathbf{P}'_c \\ \mathbf{0} \end{bmatrix} \tag{2}$$

where $L_{ij} = U(\|\mathbf{c}_j - \mathbf{c}_i\|)$, \mathbf{P}_c are the stacked coordinates of the control points $(1, \check{x}, \check{y})$ on the original plane and \mathbf{O} and $\mathbf{0}$ are 3×3 and 3×2 zero matrices respectively. Denoting the leftmost matrix as \mathbf{K} , we invert (2) to solve for \mathbf{t}^x and \mathbf{t}^y :

$$\begin{bmatrix} \mathbf{t}^x & \mathbf{t}^y \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} \mathbf{P}'_c \\ \mathbf{0} \end{bmatrix} \tag{3}$$

For simplification purposes, we denote the $(n+3) \times n$ sub-matrix of \mathbf{K}^{-1} as \mathbf{K}_* . For the problem of aligning two images, the transformed pixel coordinates $\mathbf{x}' = (x', y')$ of the reference image can be calculated as a function of control point correspondences \mathbf{c}' on the current image. The q transformed pixel coordinates \mathbf{x}'_j can be stacked in a matrix \mathbf{P}'_x such that $\mathbf{P}'_x = [\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_q]^T$:

$$\mathbf{P}'_x = [\mathbf{V} \ \mathbf{W}] \mathbf{K}_* \mathbf{P}'_c \tag{4}$$

where $V_{ji} = U(\|\mathbf{x}_j - \mathbf{c}_i\|)$ and $\mathbf{W}_j = (1, x_j, y_j)$. For a more compact notation, the matrix $[\mathbf{V} \ \mathbf{W}]$ is denoted as \mathbf{M} . If the position of the control points \mathbf{c} in the reference plane does not change, \mathbf{M} and \mathbf{K}_* can be pre-computed.

2.1 Modeling the Surface Depth

In our problem, we track a template image \mathbf{T} representing the target surface which undergoes non-rigid 3D motion. The TPS warping presented above defines a mapping between each pixel position on \mathbf{T} to the current image \mathbf{I} of the surface, as defined in (4). Nevertheless, the standard TPS warping does not have a division in its formulation and therefore it cannot capture projective deformations with a finite number of control points. Even though techniques such as dynamic control point insertion [10] can be used to reduce the warping imprecision, the warping model remains insufficient.

In a recent work, Malis [13] proposed the parameterization of the projective depths of the tracked surface in order to capture more general 3D motion. In a similar fashion, the mapping m proposed in (2) can be extended to model surface depth using an additional TPS function f^z :

$$m(\mathbf{x}) = \begin{bmatrix} f^x \\ f^y \\ f^z \end{bmatrix} = \begin{bmatrix} r_2^x & r_3^x & r_1^x \\ r_2^y & r_3^y & r_1^y \\ r_2^z & r_3^z & r_1^z \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + \sum_{i=1}^n \begin{bmatrix} w_i^x \\ w_i^y \\ w_i^z \end{bmatrix} U(\|\mathbf{c}_j - \mathbf{x}\|) \tag{5}$$

With this new formulation, the transformed pixel coordinates can be calculated in the same fashion as equation (4) by considering \mathbf{x}' and \mathbf{c}' in homogeneous coordinates $\mathbf{x}' = (sx', sy', s)$ and $\mathbf{c}' = (s\check{x}', s\check{y}', s)$.

To fully constrain the 3D tracking problem, a calibrated stereo rig with the world coordinate frame centered on the left camera is considered. The design of our efficient tracking algorithm comes from one important observation: if the same target image \mathbf{T} is tracked on left and right images of the stereo pair, the position of the control points correspondences \mathbf{c}' on both images are consequently the projections of the same 3D points of the target surface. By directly estimating the coordinates of these 3D points, the problem of tracking the target image \mathbf{T} on both images of the stereo pair can be efficiently solved. The control point correspondences \mathbf{P}'_c can be understood as the projections of 3D points \mathbf{X} imaged by a camera with a 3×4 calibration matrix \mathbf{C} :

$$\mathbf{P}'_c = (\mathbf{C} [\mathbf{X} \mathbf{1}]^T) \quad (6)$$

where $\mathbf{1}$ a n -vector of ones. Finally, replacing (6) in (4), the mapping becomes a function of \mathbf{X} :

$$\mathbf{P}'_x = \mathbf{M} \mathbf{K}_* [\mathbf{X} \mathbf{1}] \mathbf{C}^T \quad (7)$$

From the formulation above, we can define the warping function $w(\mathbf{x}_i, \mathbf{h}, \mathbf{C})$ that maps the i -th pixel \mathbf{x}_i of the target image \mathbf{T} onto the image \mathbf{I} (with camera matrix \mathbf{C}) based on a parameter vector \mathbf{h} , which is a $3n$ column vector composed by the stacked columns of the matrix $\mathbf{X} = (\mathbf{h}^x, \mathbf{h}^y, \mathbf{h}^z)$. Denoting $\mathbf{h} = (\mathbf{h}^{xT}, \mathbf{h}^{yT}, \mathbf{h}^{zT})^T$, we have:

$$w(\mathbf{x}_i, \mathbf{h}, \mathbf{C}) = [sx'_i \quad sy'_i \quad s_i] = \mathbf{M}_i \mathbf{K}_* [\mathbf{h}^x \mathbf{h}^y \mathbf{h}^z \mathbf{1}] \mathbf{C}^T \quad (8)$$

where \mathbf{M}_i is the i -th row of the matrix \mathbf{M} corresponding to \mathbf{x}_i .

The tracking problem is the estimation of the optimal warping parameter vector \mathbf{h} that minimizes the alignment error ϵ between the reference image \mathbf{T} and both left and right images of the stereo pair \mathbf{I}_l and \mathbf{I}_r simultaneously:

$$\min_{\mathbf{h}} \epsilon = \sum_{\mathbf{x} \in \mathbf{A}} \left[[\mathbf{I}_l(w(\mathbf{x}, \mathbf{h}, \mathbf{C}_l)) - \mathbf{T}(\mathbf{x})]^2 + [\mathbf{I}_r(w(\mathbf{x}, \mathbf{h}, \mathbf{C}_r)) - \mathbf{T}(\mathbf{x})]^2 \right] \quad (9)$$

where \mathbf{A} is the set of the template coordinates and $\mathbf{I}(w(\mathbf{x}, \mathbf{h}, \mathbf{C}))$ is an image transformed by the warping function $w(\mathbf{x}, \mathbf{h}, \mathbf{C})$.

For solving the minimization problem above we use the efficient second-order minimization (ESM) algorithm [9]. The ESM is applied because it displays a faster convergence rate and larger convergence basin than traditional optimization techniques such as Gauss-Newton or Leverberg-Marquardt.

Since we propose tracking in the form of an iterative registration algorithm, we assume that for the beating heart image sequence the inter-frame motion is small. In order for this assumption to hold, we use in our experiments a high-speed acquisition system appropriate for the beating heart speed.

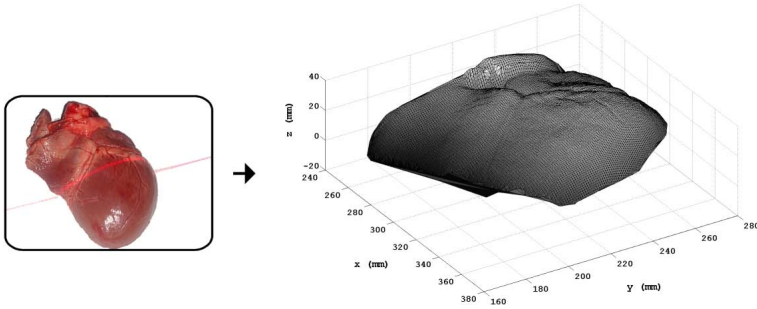


Fig. 1. Ex-vivo experiment The 3D shape of an ex-vivo pig's heart retrieved by laser profilometer (0.2mm depth precision) to serve as ground truth for the experiment

3 Experimental Results

3.1 *Ex-vivo* Experiment

In order to assess the validity of the proposed TPS model for the heart surface, an experiment on an *ex-vivo* heart was performed. We compared the 3D surface shape of an *ex-vivo* pig's heart approximated by a TPS surface and the ground truth provided by a laser profilometer (0.2mm depth precision). Figure 1 illustrates the reconstructed heart surface by the laser profilometer and figure 2 displays the proposed TPS surface representation and the corresponding approximation error. For representing a target region on the heart surface, we used a TPS surface with 9 control points evenly spaced on a grid.

Relating to the requirements of cardiac surgery where the surgeon manipulates fine structures such as veins of 2mm of diameter, the average error and standard deviation of the TPS representation of 0.66 mm and 0.36 mm respectively. Figure 2 (bottom) indicates the small approximation error and suggests the TPS model is a good representation of the heart surface. The error peaks are due to the fact that the control points are chosen evenly spaced on a grid. The TPS surface can be refined to better represent complex regions of the heart surface by adding additional control points for the representation of these specific regions.

3.2 *In-vivo* Experiment

To evaluate the clinical value of the proposed method, we performed an experiment on *in-vivo* beating heart image sequence. Since tracking is proposed as an iterative registration algorithm, the interframe motion has to be small to prevent the minimization procedure from failing to converge. For best capturing the heart dynamics, two high-speed cameras (DALSA 1M75) mounted on a small baseline were used. The acquisition speed was set to 100 Hz and special LED illumination was used. Figure 3 illustrates the experimental setup.

In this experiment, we tracked a small region on the heart surface corresponding to the coronary artery tree. To model the deformation of the target region on

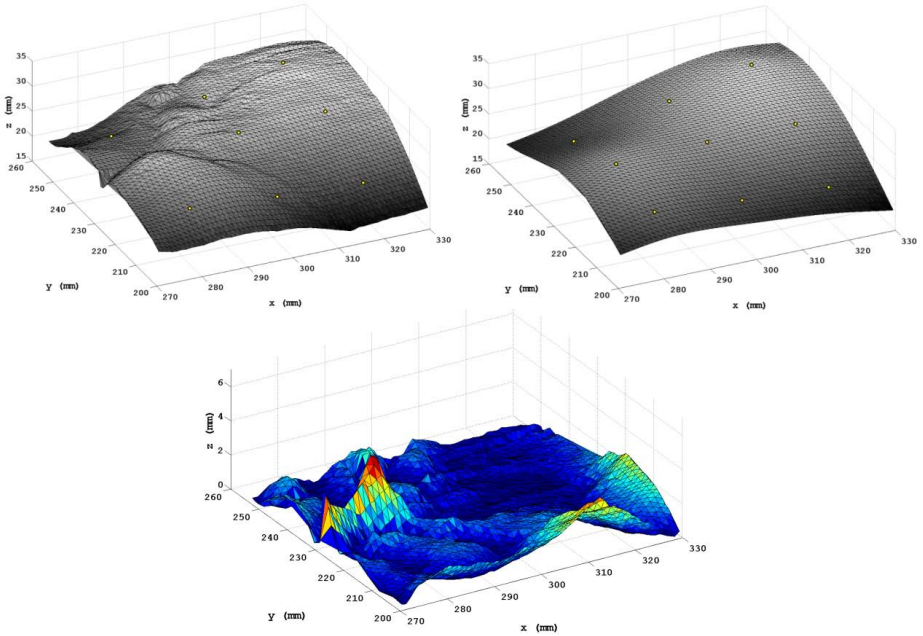


Fig. 2. The TPS model validation (Top-left) A detail of heart surface shape retrieved by the laser-profilometer (Top-right) The surface approximation by the TPS model (Bottom) The approximation error

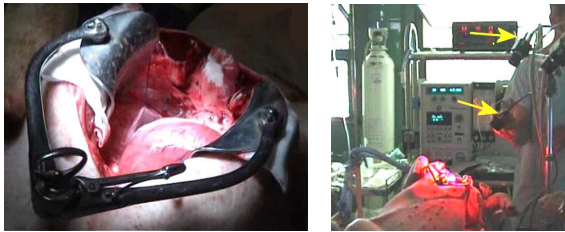


Fig. 3. The *in-vivo* experimental setup (left) Monochromatic images of a pig's beating heart were acquired after a sternotomy using two high speed cameras and special red LED lighting (highlighted) (right) The surgical setup in detail

the heart surface, 5 control points were used. Figure 4(a) illustrates the deformation the target surface undergoes on both images of the stereo pair. The 3D displacement of a fixed point on the center of the target region is plotted in figure 4(b). From the experimental results we verify that the TPS model successfully copes with the real heart surface deformation and the plotted 3D coordinates of a fixed point on the heart illustrate the high accuracy of the tracking algorithm. Although for each iteration, the ESM algorithm recalculates the Jacobian matrix of the warping, tracking only takes 4 to 6 iterations in average to converge.

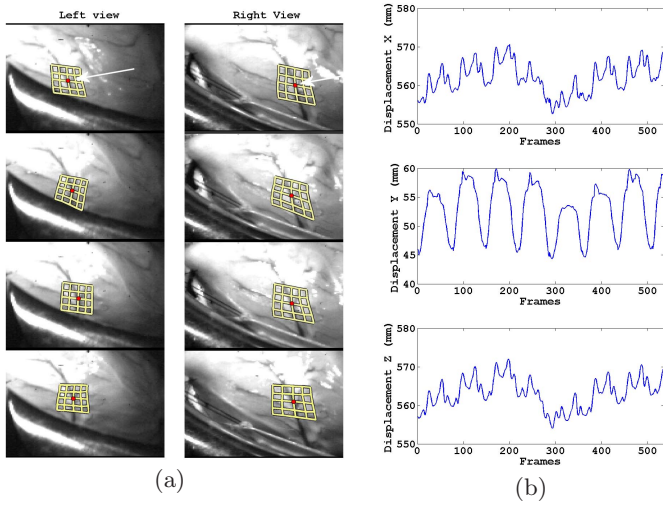


Fig. 4. Tracking the beating heart (a) The target region tracked on the left and right images of the stereo pair (first and second columns respectively) at different instants of the heart cycle. The TPS model successfully copes with the real tissue deformation. (b) The 3D coordinates of a fixed point on the heart surface (circle in the center of the grid) retrieved using the proposed method. The world coordinate frame is centered in the left camera. In our experimental setup, the stereo rig was mounted at approximately 50 cm from the target region.

4 Conclusion and Future Work

In this paper, we propose an efficient algorithm for 3D tracking the beating heart in the context of robotic-assisted cardiac MIS. The proposed method represents the first step towards the virtual stabilization of the heart. Experiments on both *ex-vivo* and *in-vivo* data attest its potential for providing a stabilized view of a target region on the heart surface. Compared to existing 3D tracking techniques, our method is simpler but still very efficient once no intermediate steps such as image rectification or matching is required. In fact, the approach presented in this paper could be generalized for tracking continuous deformable surfaces. Our current work focuses on the real time implementation of the algorithm for later deployment in a control scheme.

Acknowledgements. This work is supported by the AccuRobAs project, under the 6th Framework Program (FP6) of the European Union.

References

1. Stoyanov, D., Yang, G.Z.: Stabilization of image motion for robotic assisted beating heart surgery. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 417–424. Springer, Heidelberg (2007)

2. Nakamura, Y., Kawakami, K., Heartbeat, H.: synchronization for robotic cardiac surgery. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2001), Seoul, Korea, vol. 2, pp. 2014–2019 (2001)
3. Stoyanov, D., Mylonas, G.P., Deligianni, F., Darzi, A., Yang, G.Z.: Soft-tissue 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)
4. Noce, A., Triboulet, J., Poignet, P.: Efficient tracking of the heart using texture. In: Proceedings of IEEE International Conference of the Engineering in Medicine and Biology Society (EMBS 2007), Lyon, France, vol. 1, pp. 4480–4483 (2007)
5. Mountney, P., Lo, B., Thiemjarus, S., Stoyanov, D., Yang, G.Z.: A probabilistic framework for tracking deformable soft tissue in minimally invasive surgery. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 34–41. Springer, Heidelberg (2007)
6. Lau, W., Ramey, N.A., Corso, J.J., Thakor, N.V., Hager, G.D.: Stereo-based endoscopic tracking of cardiac surface deformation. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3217, pp. 494–501. Springer, Heidelberg (2004)
7. Stoyanov, D., Darzi, A., Yang, G.: A practical approach towards accurate dense 3D depth recovery for robotic laparoscopic surgery. *Computer Aided Surgery* 4(10), 199–208 (2005)
8. Akgul, Y.S., Kambhamettu, C.: Recovery and tracking of continuous 3d surfaces from stereo data using a deformable dual-mesh. In: Proceedings of IEEE International Conference on Computer Vision (ICCV 1999), Kerkyra, Greece, pp. 765–772 (1999)
9. Benhimane, S., Malis, E.: Homography-based 2d visual tracking and servoing. *International Journal of Computer Vision* 26 (7), 661–676 (2007)
10. Bartoli, A., Zisserman, A.: Direct estimation of non-rigid registrations. In: Proceedings of the 15th British Machine Vision Conference, Kingston, UK, vol. 2, pp. 899–908 (2004)
11. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 11(6), 567–585 (1989)
12. Lim, J., Yang, M.: A direct method for modeling non-rigid motion with thin plate spline. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005), Washington, USA, vol. 1, pp. 1196–1202 (2005)
13. Malis, E.: An efficient unified approach to direct visual tracking of rigid and deformable surfaces. In: Proceedings of IEEE Conference on Intelligent Robots and Systems (IROS 2007), San Diego, USA, pp. 2729–2734 (2007)