3D HEART MOTION ESTIMATION USING ENDOSCOPIC MONOCULAR VISION SYSTEM

Mickaël Sauvée [*,**] **Philippe Poignet** [**]
Jean Triboulet [**] **Etienne Dombre** [**] **Ezio Malis** [***]
Roland Demaria [****]

[*] *Sinters, 5 rue Paul Mesple, 31000 Toulouse, France*
[**] *LIRMM, 161 rue Ada, 34292 Montpellier, France*
[***] *INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France*
[****] *Service de chirurgie Cardiovasculaire, CHU Arnaud de Villeneuve, 371 av. G. Giraud, 34295 Montpellier, France*

Abstract: In robotic assisted beating heart surgery, motion of the heart surface might be virtually stabilized to let the surgeon work as in on-pump cardiac surgery. Virtual stabilization means to compensate physically the relative motion between instrument tool tip and point of interest on the heart surface, and to offer surgeon a stable visual display of the scene. In this way, motion of the heart must be estimated. This paper focus on motion estimation of heart surface. Classical computer vision method has been applied to reconstruct 3D pose of interest point. Experimental results obtained on *in vivo* images show the feasibility of this approach for estimating motion of heart surface points.

Keywords: Medical applications - Motion estimation - Vision - Robot control - Spectral analysis

1. INTRODUCTION

One of the most widely spread intervention in cardiac surgery is coronary artery bypass grafting (CABG). Currently most of them are performed using heart lung machine and arrested heart, which allows the surgeon to achieve complex and fine suture on motionless heart surface. However cardiopulmonary bypass (CPB) have deleterious effects. For instance, clinical studies (Picone *et al.* (1999)) have shown a systemic inflammatory response resulting from blood contact with extra-corporeal circuit of the CPB machine. To avoid the CPB problems, solutions for operations on the beating heart should be proposed. Among proposed solutions, passive mechanical stabilizers have been used on the heart surface to cancel the motion (e.g. Octopus™ by Medtronic, Jansen *et al.* (1997)). The idea is to apply a mechanical constraint on the heart surface to stabilize the working area. To ensure contact between heart surface and mechanical device, vacuum suction or suture technique have been developed. However remaining motion inside the stabilized area still exists. Lemma *et al.* (2005) have observed excursion up to 2.4 mm on the stabilized area. So surgeons have to manually cancel the movement of the target coronary. Another way to perform beating heart surgery will be to compensate the organ motion with a robotized assistive device.

Initially proposed to overcome minimally invasive surgery (MIS) difficulties (such as decreased mobility, limited point of view), the introduction of

telerobotic systems (e.g. Da Vinci™ system by Intuitive Surgical, Guthart and Salisbury (2000)) in operating room offer the possibility to provide tools for motion compensation. Thus the surgeon can concentrate on his task. In robotic assisted beating heart surgery, the surgeon desired movement might be superimposed to the trajectory generated by the motion compensation algorithm. Hence the surgeon could work on a virtually stabilized heart as in on-pump surgery. In this situation, the key issues are listed bellow:

(1) **Heart motion estimation**
Movement must be extracted from visual feedback from endoscopic view of highly deformable and non structured surface, so development of specified algorithms must be done
(2) **Control system design**
Robot manipulator must compensate for high bandwidth motion (up to 5Hz), with high precision (suture tasks are performed on 2 millimeter diameter vessels). Moreover, stability must be guaranteed for motion in both free space and in interaction with environment.
(3) **Visual stabilization**
A stable view of suturing area must be displayed to the surgeon screen.

Related work have been proposed in the literature. First solution was proposed by Nakamura *et al.* (2001). They used a high speed camera which tracked artificial markers placed on the heart. Visual feedback control was used to control a light mini-robot. *In vivo* experiments performed on porcine model show good 2D trajectory tracking, but errors of about 1 mm had been observed. More recently, motion estimation based on natural landmark has been presented by Ortmaier *et al.* (2005). In this paper, measurements are based on natural landmark tracking, with a prediction algorithm based on ECG signals and respiratory pressure signal to improve robustness of landmark detection. Nevertheless, experimental evaluation is restricted to tracking of landmarks inside a mechanically stabilized area and presented results were expressed in image coordinate, without 3D reconstruction. In Ginhoux *et al.* (2005), active markers are placed on the heart and pointing laser system observed by a 500 Hz video sensor allows to compute the distance between the instrument tool tip and the reference surface, used in an adaptive model predictive controller. *In vivo* evaluation on porcine model exhibits motion cancellation with error variance about 1.5 mm along X and Y direction. However the given estimation only reflects a global motion of the heart surface if we consider the large dimension of their active marker system. Recently Cavusoglu *et al.* (2005) proposed to combine biological signals (ECG signal, arterial and ventricular blood pressures) and heart motion measurement in a model based predictive algorithm to add feedforward path to robot motion control. To analyze heart motion, measurements are obtained by Sonomicrometric system (manufactured by Sonometric). This technique is based on ultrasound signals, transmitted by small piezoelectric crystals fixed to the heart surface. To perform this experiment, the pericardial sac has to be filled with a saline solution. This constrain does not rely on feasibility procedure during CABG.

Estimation of displacement and acceleration of heart surface is a key issue in robotic assisted surgery for motion compensation control design. In this paper, we focus on 3D heart motion estimation. We proposed to use available information, i.e. endoscopic image, to perform 3D motion reconstruction based on small passive artificial landmarks with known dimension placed on the heart surface. In the remainder of this paper, first the method applied on endoscopic images to estimate motion of interest points is presented. Experimental evaluation of this method is detailed in a second section. Conclusions and discussion about on going work are exposed in the last section.

2. METHODS

The chosen approach for extracting visual information and estimating motion of the interest region could be described as follows :

(1) Image acquisition using calibrated endoscopic vision system,
(2) Tracking of given and geometrical patterns, small enough to be assumed planar,
(3) Pose estimation of the template using metric of the object.

2.1 Endoscopic vision model

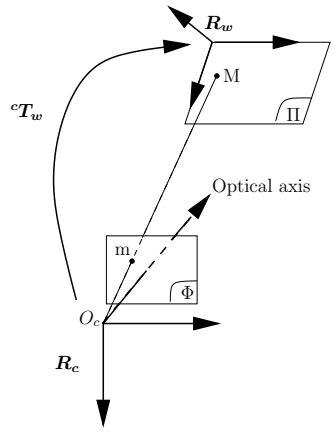


Fig. 1. Projection model : frame description.

A pinhole model is applied to describe image formation in classical vision system, based on thin lens hypothesis. Although endoscope is a long rigid tube composed of a succession of lenses that provide image from extremity of the tube to camera lens, pinhole model is also assumed in literature to model endoscopic image (Marti *et al.* (2004)). Nevertheless pinhole model is a linear approximation of the real camera projection. Therefore with endoscopic system which induces high radial distortion, it is necessary to improve accuracy of the model by adding nonlinear compensation (Zhang and Payandeh (2002)).

The pinhole model is split into two parts. First part (see Figure 1) maps the coordinates of point M defined in world frame R_w (attached to the object plane Π) to coordinates of point m on plane Φ at a distance equal to 1 from camera projection center O_c (center of camera frame R_c) along optical axis (Z direction of frame R_c) (equation (1)). The coordinates $(x_n, y_n)^t$ of point m are the normalized coordinates, obtained from perspective projection without considering camera intrinsic parameters. The perspective projection matrix $\boldsymbol{H}$ results from the combination of (3*4) projection matrix and the rigid matrix transform ${}^{\boldsymbol{c}}\boldsymbol{T_w}$ defining world frame R_w w.r.t. camera frame R_c, s is a scale factor induced by perspective projection. $\boldsymbol{r_i}$ are the columns of the rotation matrix and $\boldsymbol{t}$ is the translation vector from the rigid transformation ${}^cT_w$.

$$s\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \boldsymbol{H}\begin{pmatrix} {}^wX \\ {}^wY \\ {}^wZ \\ 1 \end{pmatrix} = [\boldsymbol{r_1 r_2 r_3 t}]\begin{pmatrix} {}^wX \\ {}^wY \\ {}^wZ \\ 1 \end{pmatrix} \quad (1)$$

The next step takes into account the intrinsic parameters of the vision system by applying affine transform from camera frame to image frame (equation (2)).

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \boldsymbol{K}\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} \quad (2)$$

$\boldsymbol{K}$ is the (3*3) intrinsic parameter matrix, composed of optical center coordinates (u_0, v_0) and focal length in X and Y direction (f_{c1}, f_{c2}).

Distortion model adds extra displacements on the normalized coordinates of point m. The model is composed of two components. First one takes into account radial distortion (equation (3)) , and second one approximates tangential distortion (equation (4)):

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = (1 + k_1 r^2 + k_2 r^4)\begin{pmatrix} x_n \\ y_n \end{pmatrix} + dx \quad (3)$$

$$dx = \begin{pmatrix} 2k_3\ x_n y_n + k_4(r^2 + 2x_n^2) \\ k_3(r^2 + 2y_n^2) + 2k_4 x_n y_n \end{pmatrix} \quad (4)$$

where r stand for $\sqrt{x^2 + y^2}$. The intrinsic parameters and coefficients of distortion model have been computed using Matlab Calibration Toolbox [1] (based on algorithm proposed by Zhang (1998)). Since model is an approximation of reality, workspace calibration must include volume created by heart motion interest point.

2.2 Pattern tracking algorithm

For each image acquired, homography matrix at time k, $\boldsymbol{G(k)}$, has to be computed between the reference and the current pattern, to extract image coordinates of the four corners points of the square pattern $(u_i(k), v_i(k))^t$ using image coordinates in the reference pattern $(u_i(0), v_i(0))^t$ through equation (5). The coordinates of the four corners in the reference pattern are manually selected during off line procedure and automatically refined by detecting nearest corner location at subpixel level.

$$\lambda\begin{pmatrix} u_i(k) \\ v_i(k) \\ 1 \end{pmatrix} = \boldsymbol{G(k)}\begin{pmatrix} u_i(0) \\ v_i(0) \\ 1 \end{pmatrix} \quad \forall i = 1 \dots 4 \quad (5)$$

A pattern based algorithm able to track the whole image of the square object was chosen to compute $\boldsymbol{G}(k)$. We used Efficient Second order Method (ESM) proposed by Benhimane and Malis (2004). The application of the ESM algorithm to visual tracking allows an efficient homography estimation with high inter-frame displacements. In this algorithm, an iteratively estimate procedure find the optimal homography which minimizes the Sum of Square Differences between the reference pattern (defined off line) and the current pattern (which has been reprojected in the reference frame using the current homography). Because initial prediction of the homography is not available, algorithm starts with initial estimation equal to the identity matrix. Both the image derivatives of the template and the image derivatives of the current pattern are used to obtain an efficient second-order update. It is an efficient algorithm since only first derivatives are used and the Hessians are not explicitly computed. The whole details of the computation and the software can be found on the authors' website [2] .

[1] available on : http://www.vision.caltech.edu/bouguetj/calib_doc/ (last updated : 2005)

[2] http://www-sop.inria.fr/icare/personnel/malis/software/ESM.html (last updated 2005)

At the end of this step, only information in image space has been computed. We now have to integrate metric information about the pattern to estimate 3D pose.

2.3 3D Pose estimation

The method used here was inspired from Zhang (1998) with known intrinsic parameters and implemented with OpenCV Library [3]. Assuming world frame attached to object plane, ${}^wZ = 0$ for the 4 points and equation (1) can be rewritten in equation (6) without loss of generality.

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = s^{-1}[\boldsymbol{r_1 r_2 t}] \begin{pmatrix} {}^wX \\ {}^wY \\ 1 \end{pmatrix} = \boldsymbol{H^*}(k) \begin{pmatrix} {}^wX \\ {}^wY \\ 1 \end{pmatrix} \tag{6}$$

Image plane coordinates related to world coordinates are used to compute perspective projection matrix estimation $\hat{\boldsymbol{H}}^*(k)$ between $(x_n, y_n, 1)^t$ to $({}^wX, {}^wY, 1)^t$ (see Appendix A of Zhang (1998)). Scale factor s is retrieved computing the mean of norms of the first two columns of matrix $\hat{\boldsymbol{H}}^*(k)$ ($\hat{s} = 0.5 * (||\boldsymbol{h_1}|| + ||\boldsymbol{h_2}||)^{-1}$). Then from equation (6), we get:

$$\hat{\boldsymbol{r_1}} = \hat{s}\boldsymbol{h_1}, \quad \hat{\boldsymbol{r_2}} = \hat{s}\boldsymbol{h_2}, \quad \hat{\boldsymbol{t}} = \hat{s}\boldsymbol{h_3} \tag{7}$$

The third vector of rotation matrix is obtained using orthogonality property of this matrix :

$$\hat{\boldsymbol{r_3}} = \hat{\boldsymbol{r_1}} \times \hat{\boldsymbol{r_2}} \tag{8}$$

3. RESULTS

3.1 Precision evaluation

Before applying this approach to estimate heart motion, accuracy of the method has been evaluated using calibrated measurement system. Different error sources can be listed: camera model approximation, pattern tracking accuracy, 3D reconstruction algorithm.

The endoscopic vision system is composed of a rigid endoscope (Hopkin's II from Karl Storz Inc), mounted on 35 mm focal lens optic and CMOS camera Dalsa 1m75. 300 W xenon light source is connected to the rigid endoscope. Frame rate is adjusted at 125 Hz, and 512*512 pixel image size is selected.

First NDI Polaris system was used (accuracy = 0.35 mm) to evaluate static estimation. Object have been placed on the passive marker tool of the Polaris system, and we track displacement of this tool to measure motion induced to the object. We work with relative motion estimation to avoid computation of rigid transformation between Polaris reference frame and camera frame. Results obtain in this evaluation show precision along X and Y direction ten times smaller than error along Z direction. Secondly, from this first observation, we focus on Z direction and we use laser measurement device (precision = 0.05 mm) to evaluate the precision of our approach in dynamic motion. Object was manually displaced constraining the object plane to be parallel to image plane. Maximum error of 0.8 mm is observed w.r.t. laser measures (standard deviation : 0.34 mm). So we consider that our complete system (endoscopic calibrated camera + pattern tracking + pose estimation) give us estimation of motion with precision of 1 mm.

Table 1. Intrinsic parameters (in pixel).

Pinhole model	f_{c1}	f_{c2}	u_0	v_0
	593.6	594.5	279.6	223.7
Distortion model	k_1	k_2	k_3	k_4
	-0.241	0.207	0.005	0.008

3.2 In vivo *experimentation*

The proposed approach has been tested on porcine model of 25 kg. The animal has been anaesthetised before intervention. Pig was under respiratory machine which constrains respiratory cycle to 20 cycles per minute. A thoracotomy has been performed to facilitate heart access. We assume that chest opening modify heart motion, but we think that in endoscopic situation, heart motion must be more constrained and observed displacement amplitude might be less important in real endoscopic context. Images have been acquired using endoscopic vision system presented in section 3.1. Endoscopic vision system is calibrated within a workspace of 20*15*15 mm, centered at a distance of 55 mm from the endoscope tip. Intrinsic parameters are given in Table 1.

Artificial passive markers have been placed on the heart surface as shown in Figure 2. Each planar marker is square of 5 mm side, drawn on white paper. To avoid lack of information for tracking algorithm, variable grey values have been used inside the pattern. Sequences of 1900 images have been acquired with a sampling time of 8 ms.

3.3 Motion analysis

Estimation of heart motion are presented in figure 3. Motion is estimated over the entire acquired sequence, that is 15 s.

[3] available on : http://www.intel.com/technology/computing/opencv/index.htm (last updated : 2005)

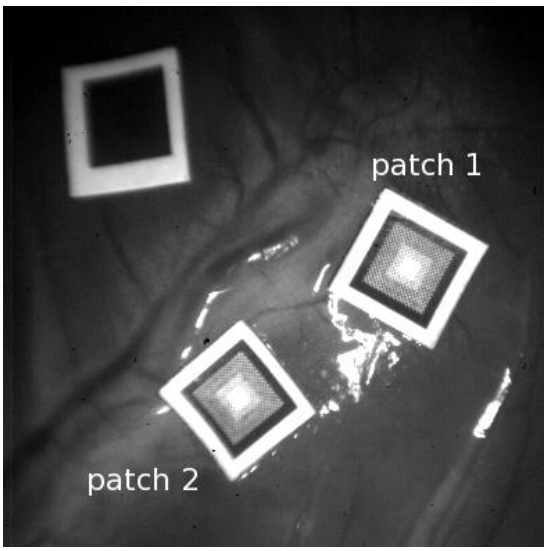


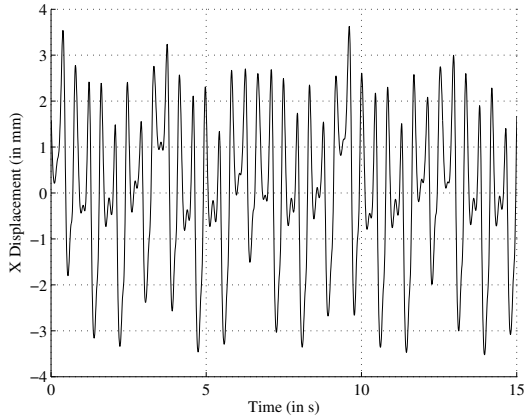
Fig. 2. Endocsopic image of the heart with planar object.

Table 2. Spectral analysis.

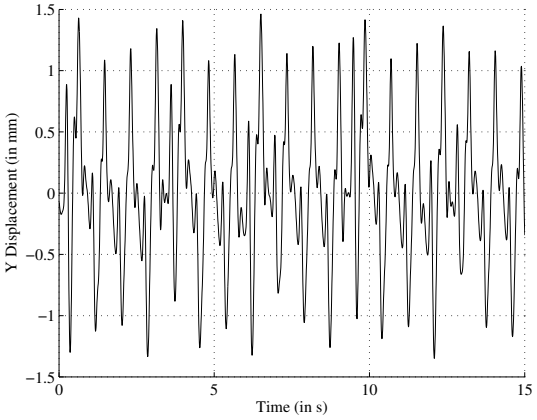
Frequency (in Hz)	0.34	0.67	1.19	2.38	3.57	4.76
Density in X direction	8.7%	0%	21.8%	60%	2.8%	6.7%
Density in Y direction	3.9%	1.9%	45.6%	33%	3.9%	11.7%
Density in Z direction	70.9%	11.5%	5.9%	8.7%	1.2%	1.8%

A spectral analysis have been performed to evaluate frequency component of the estimated signal (see table 2). The first two components (f_1 = 0.34 Hz, and f_2 = 0.68 Hz) must be associated with respiratory activity. f_1 is equal to frequency imposed by the respiratory machine (20 cycles per minute). The second one must be treated as harmonic component of the respiratory activity (f_2 = $2f_1$). Heart activity provides 4 other frequencies. f_3 = 1.19 Hz might represent the heart beat cycle (around 70 beats per minute), and the three others are harmonics of f_3. The spectral analysis shows that estimated motion are dominated by cardiac and respiratory activities. From power spectral density analysis, motion along X and Y direction (parallel to image plane) are mostly governed by cardiac activity, whereas amplitude of the motion along Z direction is greatly induced by respiratory activity. By applying low pass filter of 1Hz cutoff frequency, cardiac and respiratory components can be extracted from estimated motion in Z direction (see Figure 4). Thus maximum amplitude observed in Z direction around 11 mm is due for one part to respiratory displacement of 6 mm and for the other to cardiac displacement of 5 mm.

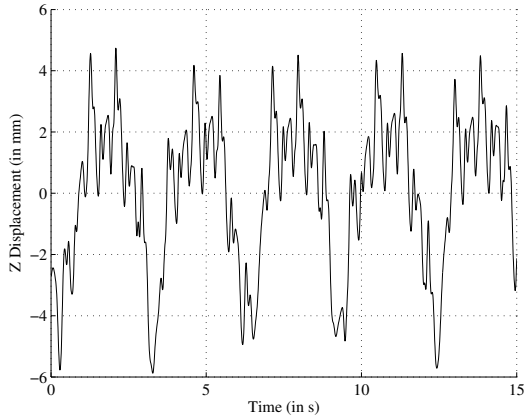
Acceleration have also been computed (see table 3). Heart motion exhibits maximum acceleration larger than 3 $m.s^{-2}$. Moreover, estimation obtained from other square objects placed at other



(a) X direction



(b) Y direction



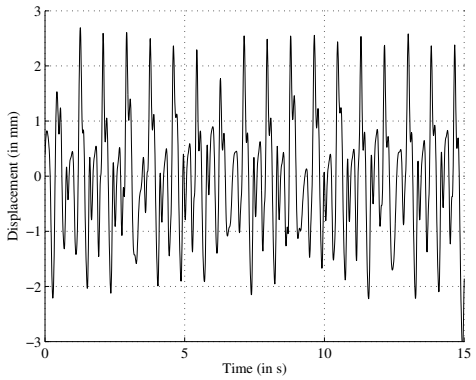
(c) Z direction

Fig. 3. Motion estimation of patch 1.

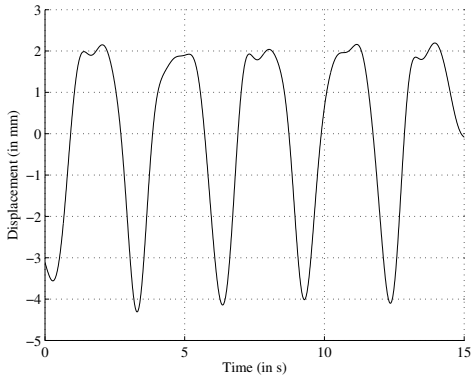
location on the heart exhibits acceleration around 4 $m.s^{-2}$.

4. CONCLUSIONS

In this paper, the estimation of heart motion using available endoscopic vision system in operating



(a) Cardiac component



(b) Respiratory component

Fig. 4. Decomposition of motion along Z direction.

Table 3. Maximum acceleration of the estimated motion.

	$\ddot{x}$ (in $m.s^{-2}$)	$\ddot{y}$ (in $m.s^{-2}$)	$\ddot{z}$ (in $m.s^{-2}$)
patch 1	1.73	1.09	3.33
patch 2	1.38	0.94	3.67

room have been presented. Proposed method exhibits result with precision better than 1 mm. *In vivo* experimentation has been performed on porcine model. Estimated heart motion is clearly governed by cardiac and heart cycle activities. Depending on location on heart surface, acceleration up to 4 $m.s^{-2}$ have been observed. Taking into account acceleration obtained, robotic system must be able to follow 10 $m.s^{-2}$ to ensure motion compensation in large security.

On going work is about improvement of the proposed approach. We currently work on in two different ways. On one hand, a real-time (sample time less than 10 ms) version of the proposed approach is under development to perform motion estimation during video acquisition. On the other hand, we work on the development of biomechanical model to describe local surface displacement based on estimation of few points around the working area (defining some limit conditions).

REFERENCES

Benhimane, S. and E. Malis (2004). Real-time image-based tracking of planes using efficient second-order minimization. In: *Proc. of Int. Conf. on Intelligent Robot and Systems*. Sendai, Japan. pp. 943–948.

Cavusoglu, M. Cenk, J. Rotella, W.S. Newman, S. Choi, J. Ustin and S.Shankar Sastry (2005). Control algorithms for active relative motion cancelling for robotic assisted off-pump coronary artery bypass graft surgery. In: *Proc. of the 12th Int. Conf. on Advanced Robotics*. Seattle, WA, USA.

Ginhoux, R., J. Gangloff, M. de Mathelin, L. Soler, M. M. Arena Sanchez and J. Marescaux (2005). Active filtering of physiological motion in robotized surgery using predictive control. *IEEE Trans on Robotics* **21**(1), 67–79.

Guthart, G. S. and J. K. Salisbury (2000). The intuitive telesurgery system: overview and application. In: *Proc. of Int. Conf. Of Robotics and Automation*. pp. 618–621.

Jansen, E. W.L., P. F. Gründeman, C. Borst, F. Eefting, J. Diephuis, A. Nierich, J. R. Lahpor and J. J. Bredée (1997). Less invasive off-pump cabg using a suction device for immobilization: the octopus method. *Eur. Jour. of Cardio-Thoracic Surgery* **12**, 406–412.

Lemma, M., A. Mangini, A. Reaelli and F. Acocella (2005). Do cardiac stabilizers really stabilize? experimental quantitative analysis of mechanical stabilization. *Interactive Cardiovascular and Thoracic Surgery* **4**, 222–226.

Marti, G., V. Bettschart, J.-S. Billiard and C. Baur (2004). Hybrid method for both calibration and registration of an endoscope with an active optical marker. In: *Proc. of Computer Assisted Radiology and Surgery*. pp. 159–164.

Nakamura, Y., K. Kishi and H. Kawakami (2001). Heartbeat synchronization for robotic cardiac surgery. In: *Proc. of Int. Conf. on Robotics and Automation*. Seoul, Korea. pp. 2014–2019.

Ortmaier, T., M. Groger, D. H. Boehm, V. Falk and G. Hirzinger (2005). Motion estimation in beating heart surgery. *IEEE Trans. on Biomedical Engineering* **52**(10), 1729–1740.

Picone, A. L., C. J. Lutz, C. Finck, D. Carney, L. A. Gatto, A. Paskanik, B. Searles, K. Snyder and G. Nieman (1999). Multiple sequential insults cause post-pump syndrome. *Annals of Thoracic Surgery* **67**, 978–985.

Zhang, X. and S. Payandeh (2002). Application of visual tracking for robotic-assisted laparoscopic surgery. *Journal of Robotics Systems* **19**(7), 315–328.

Zhang, Z. (1998). A flexible new technique for camera calibration. Technical Report MSR-TR-98-71. Microsoft Research.