

A Novel Algorithm for Heart Motion Analysis Based on Geometric Constraints

Mingxing Hu¹, Graeme Penney², Daniel Rueckert³, Philip Edwards⁴, Michael Figl⁴, Philip Pratt⁴, and David Hawkes¹

¹ Centre for Medical Image Computing, University College London

² Department of Imaging Sciences, King's College London

³ Visual Information Processing, Department of Computing, Imperial College

⁴ Department of Surgical Oncology and Technology, Imperial College
London, United Kingdom

{mingxing.hu,d.hawkes}@ucl.ac.uk, graeme.penney@kcl.ac.uk,
{d.rueckert,eddie.edwards,m.figl,p.pratt}@imperial.ac.uk

Abstract. Recently, much attention has been focused on heart motion analysis for minimally invasive beating-heart surgery. Unfortunately existing techniques usually require the camera(s) to be fixed during the motion analysis, which can restrict its usefulness during surgery. In this paper we present a novel method for heart motion analysis using geometric constraint, which can estimate the motion from a moving camera and employ multiple image features to improve robustness to noise. Our approach combines the benefits of geometry estimation for obtaining an accurate and robust solution with the proper treatment of respiratory motion. The proposed method can be applied to not only beating heart surgery, but also to other procedures involving periodic organ motion, such as lung and liver.

1 Introduction

Totally endoscopic coronary artery bypass (TECAB) surgery offers substantial benefits to the patient in terms of reduced trauma and blood loss, shorter hospitalisation time and lower risk of cardiac arrest [1]. Nevertheless, operating on a beating heart can be problematic, which slows down the surgical procedures, and demands high concentration from the surgeon. Motion analysis has received a lot of attention from both the research and clinical community over the last decade. It appears to be key to more accurate and safer TECAB surgery and an essential prior step for 2D-3D registration [1], surface reconstruction [2] and robot control [3-9].

In earlier studies, Nakamura et al. proposed a heartbeat synchronization method using a high speed camera to precisely sample heart motion [3]. An artificial marker was put on heart surface for motion tracking and the amplitude of the motion was evaluated based on the variations of the marker's coordinates. Thakral also investigated the use of a fiber-optic laser sensor to measure Z-axis motion and tested their method using a surgically prepared rodent [4]. Groeger et al. employed an affine model to measure the local motion of a beating heart and performed analysis of

measured trajectories for motion estimation and global prediction [5, 6]. Around the same time, Ginhoux et al. reported their in-vivo results of beating heart tracking and proposed two different predictive-control schemes for the compensation of respiratory and cardiac motion respectively [7, 8]. More recently, a model-based motion estimation was proposed by Bader et al., which describes motion behaviour using a partial differential equation and obtains the measurement by reconstructing 3D positions of markers from a stereo camera system [9].

However, as maybe observed, a common trait shared by all these methods is that they can only use image sequence from fixed camera(s). That is, the position and orientation of the camera(s) must be fixed during the video capture, allowing feature values in X, Y or Z directions to be measured in order to analyze the heart motion. In addition, these methods usually rely on only one image feature for the motion analysis, so it is difficult to employ some robust techniques to deal directly with image noise problem.

Therefore in this paper we seek a new heart motion analysis approach that works for moving camera systems and makes the following contributions: (1) The heart motion is estimated by employing the geometry distance between images, rather than insisting on using directly the feature values in X, Y or Z positions. Epipolar geometry encapsulates camera information (intrinsic and extrinsic parameters) and so it can handle the movement of the camera and even the change of the internal geometry. (2) The estimation is based on multiple feature points tracked in the image sequence, which make the proposed method more robust to image noise. If there are enough features present, some robust techniques can also be applied to solve the outlier problem. (3) With this robust technique, it is possible to extract images from the same position within each beating heart cycle automatically, and to convert the difficult 4D (3D+time) dynamic reconstruction problem back to 3D static reconstruction problem, which is easier to solve.

2 Materials and Methods

2.1 Epipolar Geometry

Consider the case of two images of a 3D scene taken from two cameras. The point in the left image is represented by the homogeneous vector $\mathbf{m} = (x, y, 1)^T$. Its corresponding point $\mathbf{m}' = (x', y', 1)^T$ in the right image is constrained to lie on the epipolar line l' derived from \mathbf{m} . This is known as the epipolar constraint, and algebraically it can be expressed as

$$\mathbf{m}'^T \mathbf{F} \mathbf{m} = 0 \quad (1)$$

where \mathbf{F} is a 3×3 fundamental matrix with rank 2. It incorporates information about the relative transformation and internal geometry of the camera.

The classic method to estimate epipolar geometry is Hartley's eight-point algorithm [10], which greatly improves the performance of the original work of Longut-Higgins [11] by applying a simple normalization to the image data. The method may be formulated in the following two steps:

- (1) Least squares minimization. After data normalization, rewrite the epipolar constraint as a linear and homogenous equation in terms of the nine unknown coefficients in \mathbf{F}

$$\mathbf{Z}\mathbf{f} = 0$$

where $\mathbf{f} = (F_{1,1}, F_{1,2}, F_{1,3}, F_{2,1}, F_{2,2}, F_{2,3}, F_{3,1}, F_{3,2}, F_{3,3})^T$ and

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{pmatrix}$$

Then a solution of \mathbf{F} can be obtained from the vector \mathbf{f} corresponding to the smallest singular value of \mathbf{Z} in the least squares sense.

- (2) Singularity enforcement. Replace $\mathbf{F} = \mathbf{V}_F \mathbf{\Lambda}_F \mathbf{U}_F^T$ ($\mathbf{\Lambda}_F = \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}\}$, λ_1 , λ_2 and λ_3 are in decreasing order) with the nearest rank 2 matrix $\mathbf{F}^+ = \mathbf{V}_F \mathbf{\Lambda}_F^+ \mathbf{U}_F^T$ ($\mathbf{\Lambda}_F^+ = \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, 0\}$).

Hartley's method delivers results comparable with those of some nonlinear methods, and is an good tool for initialization. Readers are referred to [10] for more details.

2.2 Heart Motion Analysis

The motion of the heart surface is a caused by a combination of cardiac and respiratory motion. Usually these two independent motions need to be separated by way of signal processing. In this paper, we avoid this problem using geometry estimation under an affine assumption, and analyze the heart motion directly from the image sequence.

Although the effect of respiration on the heart surface is nonlinear, it has been shown that an affine transformation can provide a good approximation to model the respiratory motion of the heart surface [12, 13]. This is especially true for an endoscope with a narrow field-of-view, which captures images covering only a small region of the heart surface. We therefore assume we can use the affine assumption in our method and consider respiration to be an extra affine transformation of the heart surface. We put this extra affine transformation into the fundamental matrix as a relative transformation between the camera and the heart surface. The result of this is that the movement of the heart surface can now be considered as being solely due to cardiac motion. Unless noted otherwise, the term "heart motion" refers to cardiac motion in the remainder of this paper.

The heart motion is a periodic (quasi-periodic) motion. For a 3D point on the heart surface, we can write $\mathbf{X}(t) = \mathbf{X}(t + kT)$, where T is the length of the heart cycle and $k = 1, 2, \dots$. If a camera captures two images of the heart surface at times t and $t + kT$, the image projections of \mathbf{X} can be denoted as

$$\mathbf{m}^t = \mathbf{P}^t \mathbf{X}(t), \quad \mathbf{m}^{t+kT} = \mathbf{P}^{t+kT} \mathbf{X}(t + kT) = \mathbf{P}^{t+kT} \mathbf{X}(t) \quad (2)$$

where \mathbf{P} is the projective matrix, including intrinsic and extrinsic parameters.

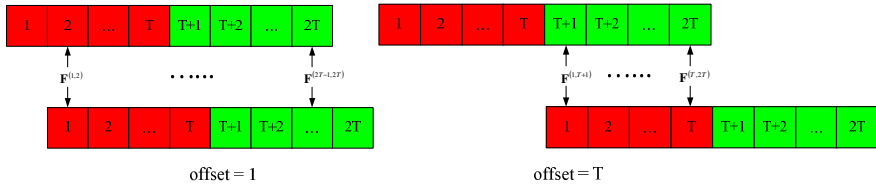


Fig. 1. Estimation of the heart cycle length based on geometric constraints. The image sequence moves from left to right with different offset and epipolar distances are computed between corresponding pairs.

Interestingly, we find that feature pair $(\mathbf{m}^t, \mathbf{m}^{t+kT})$ still satisfies the epipolar constraint, $(\mathbf{m}^{t+kT})^T \mathbf{F} \mathbf{m}^t = 0$, although they are taken at different time and using different projective matrices. But if the image points are not captured at the same position of the heart cycle, for example, at times t and $l \neq t + kT$. Eq. 1 will no longer hold for correspondence $(\mathbf{m}^t, \mathbf{m}^l)$ and this error can be evaluated by using an appropriate distance measure. The simplest measure is the algebraic distance, namely, the residual $r = (\mathbf{m}^l)^T \mathbf{F} \mathbf{m}^t$.

Unfortunately, r has no geometrical significance. Therefore in this paper we use the epipolar distance [10], which is the perpendicular distance of correspondence $(\mathbf{m}^l, \mathbf{m}^t)$ to the epipolar lines in frames t and l , defined as:

$$d_e = r \cdot \left(\frac{1}{(r'_x)^2 + (r'_y)^2} + \frac{1}{(r''_x)^2 + (r''_y)^2} \right)^{\frac{1}{2}} \tag{3}$$

where $r'_x = F_{1,1}x^t + F_{2,1}y^t + F_{3,1}$, $r'_y = F_{1,2}x^t + F_{2,2}y^t + F_{3,2}$, $r''_x = F_{1,1}x^l + F_{1,2}y^l + F_{1,3}$ and $r''_y = F_{2,1}x^l + F_{2,2}y^l + F_{2,3}$. It ensures that each image receives equal consideration.

Thus we can use the epipolar constraints to analyze the heart motion if a long enough image sequence is available, for example, with at least two or three heart cycles. As demonstrated in Fig. 1, the method is similar to self-correlation, but the criterion is based exactly on epipolar geometry. The image sequence moves from left to right with different *offset* frames, and for each offset, the fundamental matrix $\mathbf{F}^{(j-\text{offset}, j)}$ is computed between each frame pair $(j-\text{offset}, j)$ and the epipolar distance

$d_e^{(j-\text{offset}, j)} = \frac{1}{n} \sum_{i=1}^n d_{ei}$ between this pair is calculated accordingly. Then the average epipolar distance of all these $m - 2 \cdot \text{offset}$ pairs of frames can be written as

$$d_e = \frac{1}{m - 2 \cdot \text{offset}} \sum_{j=\text{offset}+1}^{m-\text{offset}} d_e^{(j-\text{offset}, j)} \tag{4}$$

When $\text{offset} = kT$, the error d_e can reach a minimum value according to the geometric constraint. The first time a minimum appears, the offset should be equal to the length of the heart cycle T . An overview of the complete algorithm is given in Fig. 2.

Here we want to emphasize that there are several advantages using geometric constraint in heart motion analysis. Firstly, it can deal with the movement and also the change of internal geometry of the camera, since we do not rely directly on the feature

Input: image sequence with n feature points from m frames
Output: length of heart cycle T

Algorithm:
For $offset = Min_offset : Max_offset$
 For $j = offset + 1 : m - offset$
 Estimate the fundamental matrix $\mathbf{F}^{(j-offset, j)}$ between frames $j - offset$ and j
 Compute the epipolar distance $d_e^{(j-offset, j)}$ using $\mathbf{F}^{(j-offset, j)}$ with Eq. (3)
 End
 Calculate the average epipolar distance with Eq. (4)
End
Find T , the first time a minimum average epipolar distance appears

Fig. 2. The algorithm for heart motion analysis based on geometric constraint

points' values in X, Y, or Z directions, which become meaningless if the camera parameters are changed. The fundamental matrix encapsulates both the intrinsic and extrinsic parameters, so we can concentrate on the heart cycle analysis using the geometry distance only. Secondly, the geometry estimation uses multiple image points and is based on a least squares minimization, so it is more robust to image noise compared with the traditional methods which use only one feature. Moreover, the geometric constraint helps us successfully separate the respiration from the heart motion under an affine assumption. The fundamental matrix encodes the respiration-induced motion as a relative movement of the camera, so the motion estimated is the heart motion itself without any respiration.

2.3 Sub-frame Accuracy

The accuracy of cycle length T we compute in Section 2.2 is limited to be an integer number of frames. In order to achieve sub-frame accuracy the following method is used.

Although heart tissue deforms nonlinearly, we can assume that the motion of a point from one frame to the next is approximately linear if the images are captured at a sufficiently high frame rate. As a result, sub-frame feature point $\mathbf{m}(t + \alpha)$ ($1 \geq \alpha \geq 0$) can be generated by using linear interpolation of image points either side of $\mathbf{m}(t + \alpha)$

$$\mathbf{m}(t + \alpha) = (1 - \alpha)\mathbf{m}^t + \alpha\mathbf{m}^{t+1} \quad 1 \geq \alpha \geq 0 \quad (5)$$

Thus it is possible to estimate the offset for a range of values around the expected value, T , at an arbitrary resolution. Furthermore, we can use the Gaussian-Newton algorithm or other optimization techniques to obtain the optimal solution with sub-frame accuracy.

2.4 Experimental Design

In order to assess the performance of the proposed method, a beating heart phantom made of silicone rubber (The Chamberlain Group, Great Barrington, MA, USA) was

employed to provide gold standard for evaluation. The phantom is continuously inflated and deflated using an air pump with an integrated controller. In order to create a beating heart model, it was scanned while beating at the rate of 55bpm with a Philips 64-slice CT scanner, producing 10 uniformly-spaced phases.

The first of these was manually segmented and converted into a tetrahedral mesh of 709 elements and 747 degrees of freedom. The Image Registration Toolkit [14] was then used to create a sequence of 3D tensor product cubic B-spline deformations, mapping the initial mesh onto each phase in turn.

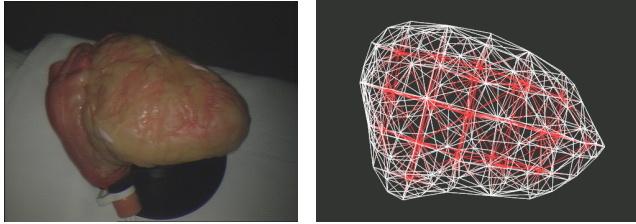


Fig. 3. Beating heart phantom and FEM mesh from CT data

To improve temporal stability of the FEM simulation 1D cubic B-splines were fitted through each node with respect to time, thus making the second order temporal derivatives continuous. Smoothness is also maintained across the coincident start and end of the cardiac cycle by arranging the basis functions. Fig. 3 shows the heart phantom and resulting mesh. Readers are referred to [15] for more details about the beating heart simulation.

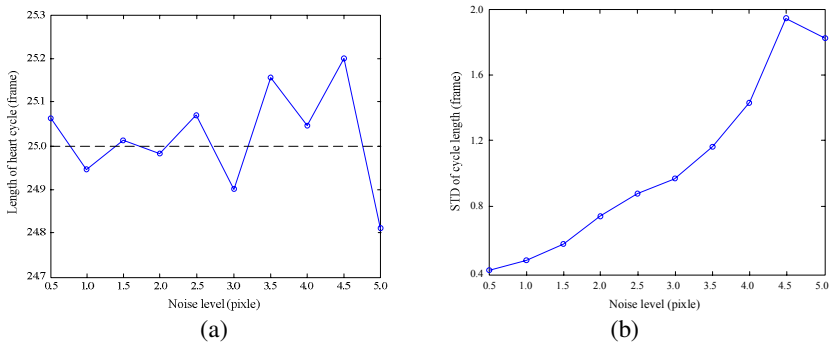


Fig. 4. Experimental results of synthetic data testing. (a) shows the estimated cycle length under different noise level (STD=0.5~5.0 pixels). (b) displays the standard deviation (STD) of estimated cycle length of 200 trials of each noise level.

3 Experimental Results

3.1 Synthetic Data

In synthetic data testing, 50 points were randomly selected from the beating heart model described in Section 2.4. For simplicity, we adjust the beating heart rate to

60bpm, that is, once per second. Then a moving camera was simulated with frame rate 25fps, and all these 3D points were projected to 50 image planes sequentially using the simulated camera, with different intrinsic and extrinsic (rotation and translation) parameters. So the ground truth for the length of heart cycle, T , is 25 frames per heart cycle.

Next, 10 different ranges of Gaussian noise were added to the image measurements, with a zero mean and standard deviation varying from 0.5 to 5.0 pixels. For each noise level, we ran 200 trials and the final results were the average of results from these 200 independent experiments. A graph of these results is shown in Fig. 4 (a). It can easily be noticed that the proposed method obtained an accurate heart cycle length over all the experiments, with error below 0.2 frame. In particular, when the noise reaches a high level ($\sigma=4,5$), the estimated results are still very close to the ground truth. Fig. 4 (b) illustrates the standard deviation of cycle length plotted against noise level. We can see that the standard deviation increases with noise, but errors are less than 2 frames. This strongly suggests that the proposed method can obtain an accurate result and is robust to image noise due to the use of multiple image features and data normalization for geometry estimation.

3.2 Phantom Model

The *da Vinci*TM robotic surgical system (Intuitive Surgical, Inc., Sunnyvale, CA, USA) was used to obtain images of the heart phantom. The heart phantom was beating with different rates, and the endoscopic camera was moving around the scene as well. The video endoscopic images were digitized at 25 frames per second (fps) using a frame grabber (LFG4 PCI64, Active Silicon, Uxbridge, U.K.). Therefore the ground truth for 55bpm testing was $60/55*25=27.27$ frames per heart cycle (fpc). 100 images were captured from the endoscope and the first and last frames are shown in Fig. 5 (a) and (b) respectively. A feature tracker based on optical flow was used to track 50 feature points over the image sequence. The proposed method was employed to calculate the length of heart cycle, and the estimated result was 27.37 frames per heart cycle, very close to the ground truth. The experiment was also repeated at higher frequency of 75bpm. In this situation, the ground truth was 20 frames per cycle. The experimental result was 20.18 frames per cycle, still very close to the ground truth.

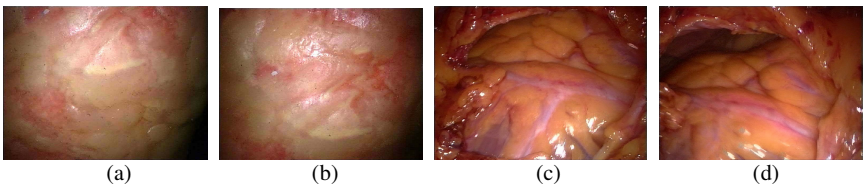


Fig. 5. Image sequences of heart phantom (a, b) and heart surface from TECAB surgery (c, d)

3.3 *In vivo* Data

Finally, the accuracy of the proposed method was evaluated using *in vivo* data. Our *in vivo* data set consisted of 50 images from a TECAB surgical procedure using the *da Vinci* system and the first and last frames are shown in Fig. 5 (c) and (d) respectively. These images were captured sequentially from the endoscope with a frame rate of 25fps and 10 feature points were manually detected and tracked from them. In order to provide a ground truth for evaluation, the electrocardiogram (ECG) signal was recorded as well during the beating heart surgery. In our experiment, the ECG reading was 76bpm and the computed cycle length was 19.57 frames per cycle, which is still close to the ground truth, i.e., $60/76 * 25 = 19.73$ fpc.

4 Discussion and Conclusions

This paper describes a novel technique for heart motion analysis using an image sequence from a moving camera. Epipolar geometry is employed to deal with the change of the camera parameters and to separate respiratory motion from heart motion under an affine assumption. Our approach has been evaluated using simulated data, a heart phantom and *in vivo* data from TECAB surgery. The results indicate the proposed method performs well in terms of both accuracy and robustness to noise, even when high-level noise is involved. Importantly our approach does not require any extra hardware, such as a high speed camera or a laser sensor, to assist the analysis.

Efforts in the near future will focus on further improvement of the robustness to image noise and outliers. Since the heart motion is estimated accurately using the proposed method, we can convert 4D dynamic surface reconstruction into 3D static scene reconstruction, and use the existing techniques i.e. [2] to solve it. Our long term goal is to automatically reconstruct the internal organ surfaces not only the beating heart surface, but also other organ surfaces with periodic motion, such as lung and liver) and register these with the preoperative data (CT or MRI).

Acknowledgments. The authors would like to thank Dr. Roberto Casula at St. Mary Hospital, London for his help in collecting the *in vivo* data. A special thanks also goes to Dr. Dean Barratt at CMIC, UCL for his helpful suggestions.

References

1. Szpala, S., Wierzbicki, M., Guiraudon, G., Peters, M.T.: Real-time fusion of endoscopic views with dynamic 3-D cardiac images a phantom study. *IEEE Trans. Med. Imaging* 24(9), 1207–1215 (2005)
2. Hu, M.X., Penney, G.P., Edwards, P.J., Figl, M., Hawkes, D.J.: 3D Reconstruction of Internal Organ Surfaces for Minimal Invasive Surgery. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 68–77. Springer, Heidelberg (2007)
3. Nakamura, Y., Kishi, K., Kawakami, H.: Heartbeat Synchronization for Robotic Cardiac Surgery. In: *Proc. ICRA*, pp. 2014–2019 (2001)

4. Thakral, A., Wallace, J., Tomlin, D., Seth, N., Thakor, N.V.: Surgical Motion Adaptive Robotic Technology (S.M.A.R.T): Taking the Motion out of Physiological Motion. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 317–325. Springer, Heidelberg (2001)
5. Gröger, M., Ortmaier, T., Sepp, W., Hirzinger, G.: Tracking local motion on the beating heart. In: Proc. SPIE Medical Imaging: Visualization, Image-Guided Procedures, and Display, vol. 4681, pp. 233–241 (2002)
6. Ortmaier, T., Groeger, M., Boehm, D.H., Falk, V., Hirzinger, G.: Motion Estimation in Beating Heart Surgery. *IEEE Transactions on Biomedical Engineering* 52, 1729–1740 (2005)
7. Ginhoux, R., Gangloff, J.A., de Mathelin, M.F., Soler, L., Arenas Sanchez, M.M., Marescaux, J.: Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In: Proc. ICRA, pp. 274–279 (2004)
8. Ginhoux, R., Gangloff, J., de Mathelin, M., Soler, L., Sanchez, M.M.A., Marescaux, J.: Active Filtering of Physiological Motion in Robotized Surgery Using Predictive Control. *IEEE Transactions on Robotics* 21, 67–79 (2005)
9. Bader, T., Wiedemann, A., Roberts, K., Hanebeck, U.D.: Model-Based Motion Estimation of Elastic Surfaces for Minimally Invasive Cardiac Surgery. In: Proc. ICRA, pp. 2261–2266 (2007)
10. Hartley, R.I., Zisserman, A.: *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge (2004)
11. Longuet-Higgins, H.C.: A computer algorithm for reconstructing a scene from two projections. *Nature* 293, 135 (1981)
12. Manke, D., Rosch, P., Nehrke, K., Bornert, P., Dossel, O.: Model evaluation and calibration for prospective respiratory motion correction in coronary MR angiography based on 3-D image registration. *IEEE Trans. Med. Imaging* 21, 1132–1141 (2002)
13. Shechter, G., Shechter, B., Resar, J.R., Beyar, R.: Prospective motion correction of X-ray images for coronary interventions. *IEEE Trans. Med. Imaging* 24, 441–450 (2005)
14. Image registration toolkit, <http://www.doc.ic.ac.uk/~dr/software/>
15. Pratt, P., Bello, F., Edwards, E., Rueckert, D.: Finite element simulation of the beating heart for image-guided robotic cardiac surgery. In: Proc. MICCAI workshop, Computational Biomechanics for Medicine, pp. 74–83 (2007)