Stereo-Based Endoscopic Tracking of Cardiac Surface Deformation

William W. Lau[1], Nicholas A. Ramey[1], Jason J. Corso[2], Nitish V. Thakor[1], and Gregory D. Hager[2]

[1] Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
{wlau, nramey, nthakor}@bme.jhu.edu
[2] Computational Interaction and Robotics Laboratory, Johns Hopkins University, Baltimore, Maryland, USA
{jcorso, hager}@cs.jhu.edu

Abstract. We propose an image-based motion tracking algorithm that can be used with stereo endoscopic and microscope systems. The tracking problem is considered to be a time-varying optimization of a parametric function describing the disparity map. This algorithm could be used as part of a virtual stabilization system that can be employed to compensate residual motion of the heart during robot-assisted off-pump coronary artery bypass surgery (CABG). To test the appropriateness of our methods for this application, we processed an image sequence of a beating pig heart obtained by the stereo endoscope used in the da Vinci robotic surgery system. The tracking algorithm was able to detect the beating of the heart itself as well as the respiration of the lungs.

1 Introduction

Stereo imaging systems are commonly employed in today's operating rooms. However, these systems usually only provide basic stereoscopic images to the surgeon. Our goal is to process these images to produce a time-varying model of the surgical field. By computing such models, it will be possible to develop control methods that improve safety and provide guidance or enhanced dexterity to the surgeon. For example, one of the major revolutions in medicine is the introduction of cardiopulmonary bypass (CPB), which provides a motionless and bloodless surgical field optimal for anastomosis construction. Arrested heart on-pump surgery has been the gold standard for the last 50 years. However, a number of studies show that CPB is associated with high incidences of morbidity, including systemic inflammatory response and cerebral thromboembolism [1-3]. Driven by the idea of decreasing the trauma related to CPB, recent advancements in stabilization techniques have made coronary artery bypass grafting (CABG) on beating heart technically feasible. In essence, the stabilizer attaches to the beating heart by means of suction to locally immobilize the coronary artery to be bypassed while the rest of the heart beats and supplies blood to the body. Currently approximately one fourth of all CABG

performed in the United States are done by beating-heart approaches [4]. With the aid of computer-assisted telemanipulation, totally endoscopic off-pump bypass surgery has also been successfully carried out by several groups [5-6].

Although sufficient stabilization can be achieved, improvement is desired to reduce prolonged operating time due to incomplete immobilization [7]. In robot-assisted surgery, this problem can potentially be addressed by a virtual stabilization system. A virtually arrested heart can be created by dynamic adjustments of the robotic arm to compensate the residual motion so that the orientation between the heart surface and the robot's end-effector remains constant. This leads to the problem of tracking the non-rigid motion of the heart surface in real time.

Traditional brute-force stereo matching techniques using local match measures are time consuming and often have limited accuracy. To provide accurate, real-time visualization of the surgical scene in 3D, we have developed a surface tracking algorithm that can recover a dense set of depth estimates of the objects in the field of view. The algorithm directly infers the 3D structure of the surface by locally adjusting an algebraic surface description. This makes the algorithm both numerically stable and computationally efficient. The ability to process the images online allows our algorithm to be used in a number of surgical applications, including establishing safety regions, developing virtual fixtures, and measuring mechanical properties of various tissues and organs. Our previous work has demonstrated the utility of the algorithm to track respiration with sub-pixel accuracy. In this paper, we focus on investigating its application to tracking the deformation of a beating heart.

2   Materials and Methods

2.1   Experimental Procedures

A cross-bred domestic pig (weight, 19.5 kg) was anesthetized with telazol-ketamine-xylazine (TKX, 4.4 mg T/kg, 2.2 mg K/kg, and 2.2 mg X/kg) and mechanically ventilated with a mixture of isoflurane (2%) and oxygen. Heart rate was continuously monitored by a pulse oximeter (SurgiVet, Waukesha, WI). The da Vinci tele-manipulation system (Intuitive Surgical, Sunnyville, CA) was used for endoscopic visualization. Three small incisions were made on the chest to facilitate the insertion of a zero-degree endoscope and other surgical tools. The pericardium was opened and video sequences of the beating heart from the left and right cameras were recorded at 30 frames/sec. The recording lasted approximately two minutes.

2.2   Surface Representation

B-spline representation is efficient in that the surface can be evaluated continuously at each pixel coordinate using a substantially smaller set of parameters than the size of the surface. Moreover, its spatial uniqueness, continuity, local shape controllability, and parameter linearity make it the best model to represent the surface for our problem.

Consider a collection of scanline locations α and row locations β, with m parameters per scanline and n parameters for row locations, a pth by qth degree tensor B-spline is a disparity function of the form

$$D(\mathrm{p};\alpha,\beta) = \sum_{i=0}^{m} \sum_{j=0}^{n} N_{i,p}(\alpha)N_{j,q}(\beta)\mathrm{p}_{i,j} \ , \tag{1}$$

where $\mathbf{p}$ is the set of control points for the B-spline surface. Furthermore, we can formulate B as a matrix of basic functions so that the disparity can be represented by the matrix product of B and $\mathbf{p}$:

$$D(\mathrm{p}) = B\mathrm{p} \ . \tag{2}$$

Let k donate an indexing linear enumeration of the mn evaluated basis functions, and define $B_{i,k} = N_{k,p}(\alpha_i) * N_{k,q}(\beta_i)$. It follows that B can be defined as

$$B \equiv \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,mn} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,mn} \\ & & \vdots & \\ B_{N,1} & B_{N,2} & \cdots & B_{N,mn} \end{bmatrix} . \tag{3}$$

If we have the intrinsic and extrinsic parameters of the cameras, we can obtain the depth information in metric coordinates using

$$Z(\mathrm{p},u,v) = f\frac{T}{D(\mathrm{p},u,v)} \ , \tag{4}$$

where u and v are image coordinates, f the common focal length, and T the baseline of the stereo system.

2.3   Stereo Matching

The algorithm has been discussed in detail in [8]. Briefly, stereo matching is computed through minimization of zero-mean sum of squared differences (ZSSD). Correct matching is dependent on the availability of texture on the surface. In addition to the global dc offset that results from differences between the optical characteristics of the two cameras, local dc offsets at parts of the images arise from the fact that the object is viewed from two slightly different angles. Regional illumination difference of the surface is therefore treated by subtracting the image by its convolution with a small Gaussian filter. As the algorithm iteratively updates disparity estimates, it requires some initial disparity estimates to start with. The initial estimates can be the disparity of a sampled point on the surface or they can be obtained by traditional stereo fitting with a complete correspondence search. The cameras are calibrated so that image pairs can be rectified to reduce the dimensionality of the problem.

Let $L(u,v,t)$ and $R(u,v,t)$ denote the left and right rectified image pair at time t, respectively. The disparity map can be considered as a lookup table consisting of a scalar offset for each pixel in the left image so that $L(u,v,t)$ and $R(u+D(u,v),v,t)$ are the projection of the same physical point in 3D space. Our objective is to estimate a set of parameters $\mathbf{p}$ that minimizes the ZSSD of the left and right warped images. Let $\overline{L}$ and $\overline{R}$ be the left and right zero-mean images, respectively, the optimization criterion can be written as

$$O(\mathrm{p}) = \sum_{(u_i,v_i)} (\overline{L}(u_i,v_i,t) - \overline{R}(u_i + D(\mathrm{p};u_i,v_i),v_i,t))^2 \ . \tag{5}$$

After some mathematical manipulations, the optimal $\Delta\mathbf{p}$ is the solution to the overdetermined linear system

$$\Delta\mathrm{p} = J(\mathrm{p},t)^t \, \mathrm{E}(\mathrm{p},t) \ , \tag{6}$$

where

$$\mathrm{E}(\mathrm{p},t) \equiv \overline{L}(t) - \overline{R}(\mathrm{p},t) \ , \tag{7}$$

and the Jacobian

$$J(\mathrm{p},t) = diag(\overline{L}_{\chi}(t)) \delta D / \delta \mathrm{p} \ . \tag{8}$$

It is obvious that $\delta D/\delta\mathbf{p}$ is equivalent to the N x mn matrix B in the previous section. Since B is a constant matrix and J does not change within an image frame, by using B-spline surface presentation, we avoid the problem of recomputing the Jacobian of the disparity function at runtime, thus speed up the processing significantly.

3   Results

The algorithm was implemented in matlab/mex. We processed the images on a desktop with a 3-Ghz P4 processor and one gigabyte of memory running Linux. A bi-quadratic B-spline surface with three control points in each direction was used. To initiate processing, we employed a standard stereo algorithm with a complete correspondence search.

The subject was ventilated at 15 breaths/min and the mean heart rate was found to be 104 beats/min. Examples of the 3D reconstruction of the heart surface are shown in figure 1. The coronary artery tree provided some natural landmarks for more accurate tracking. Figure 2a shows the motion of the heart surface measured at one pixel point for 12 seconds. The figure consists of two principal components. One component has a frequency of 60 x 3 / 12 = 15 cycles/min, and an amplitude of approximately 15 units of disparity. In figure 2b we subtract off the lower frequency component from the signal. The remaining component has a frequency of 21 x 60 / 12 = 105 cycles/min. The amplitude ranges from 6 to 8 units of disparity.

4   Discussion

Motion of the heart was contributed by the beating of the heart itself as well as the respiration of the lungs. The low-frequency, high-magnitude component represented the subject's respiration, and was consistent with the ventilator setting. The other component resembled the heart beats, and corresponded well to the oximeter measurement. Each heart beat consisted of three peaks. The dominant peak in the middle was caused by the contraction of the left ventricle. When the left ventricle started to relax, the left atrium contracted, which led to the smaller peak at the end of the cycle. We are not certain what the cause of the first peak was. In the future, we will include electrocardiography in our experiments. This will help us verify our data with higher confidence and identify the source of the first peak.
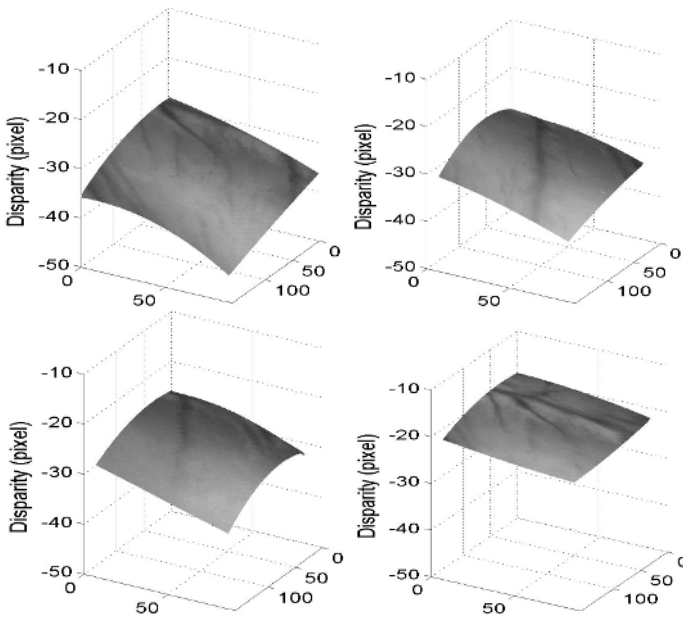

Fig. 1. 3D reconstructions of the heart surface at different phases of the heart beat cycle. The texture is mapped on a bi-quadratic B-spline surface with nine control points. The coronary artery tree serves as natural landmarks for the tracking system.

We measure the performance of the algorithm by its ability to correctly register the left and the right images. A relatively low-order B-spline surface was chosen to construct the disparity map because it sufficiently represented the smooth surface of the selected heart region. Since the region we were tracking was small (158x282 pixels), only nine control points were needed for optimal tracking. For each frame, the initial disparity was adopted from the disparity calculated in the previous frame. The system adjusted this disparity map to fit the current frame until the error or the change

of error went below a certain predefined threshold, or the maximum number of iterations had been reached. We plot in figure 3a the percentage improvement of image ZSSD difference between the left and the warped right image for each frame on a representative sequence. The error improvement ranges from about 0% to as high as 70%. Small improvements, e.g. the frame at line 1 in figure 3, do not equate to low performance. The derivative of disparity in figure 3c is indicative of how fast the heart moves. If the movement of the heart is small between two consecutive frames, the initial disparity would be very close to the optimal disparity, thus minimal change is needed. Likewise, substantial improvements (line 2) are often due to the large movement of the heart. Figure 3b shows that the final error in each frame is comparable, falling within a narrow range.
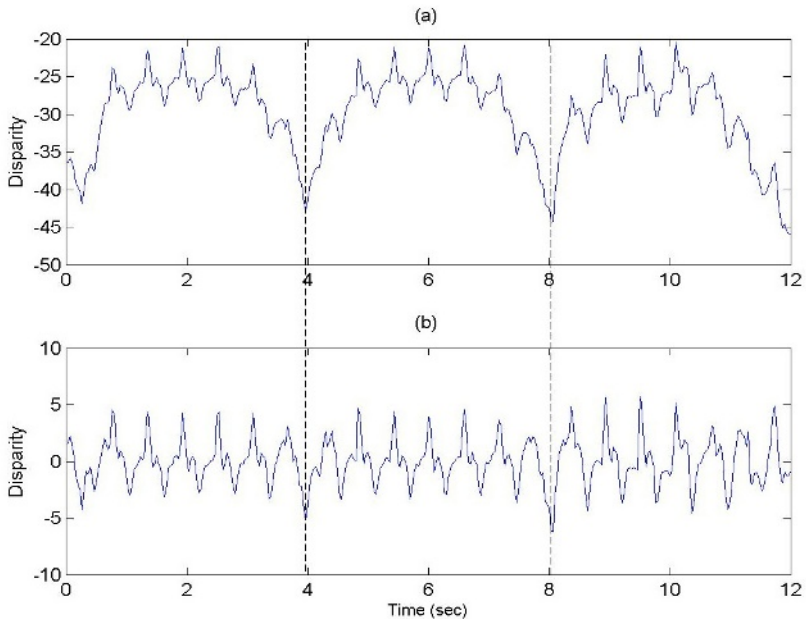

Fig. 2. (a) Shows the tracking of a pixel point on the heart surface. (b) shows the same tracking with the lower-frequency component filtered out.

Because the system essentially tracks the motion of the coronary artery tree, when the tree move away from the region of interest, the ability of the system to accurately track the surface reduces. This situation explains the larger errors in some image pairs. Nevertheless, since we do not explicitly track the motion of points on the surface, the system can be improved by projecting structured light onto the surface. Another practical issue is that during an actual bypass operation various tools will come into the scene and the heart will be occluded from the endoscope's view. Although the effect of occlusion on the stability of heart motion tracking is not assessed here, we have previously showed that the system can handle C^0 discontinuities well [8]. In addition, specular reflections can potentially cause instability to the system because of their inconsistency between the left and right images. Currently we do not account for the specularity in the images as it does not

seem to significantly affect the results. In the future, we will incorporate a scheme to detect and mask out specular reflections from the tracking.
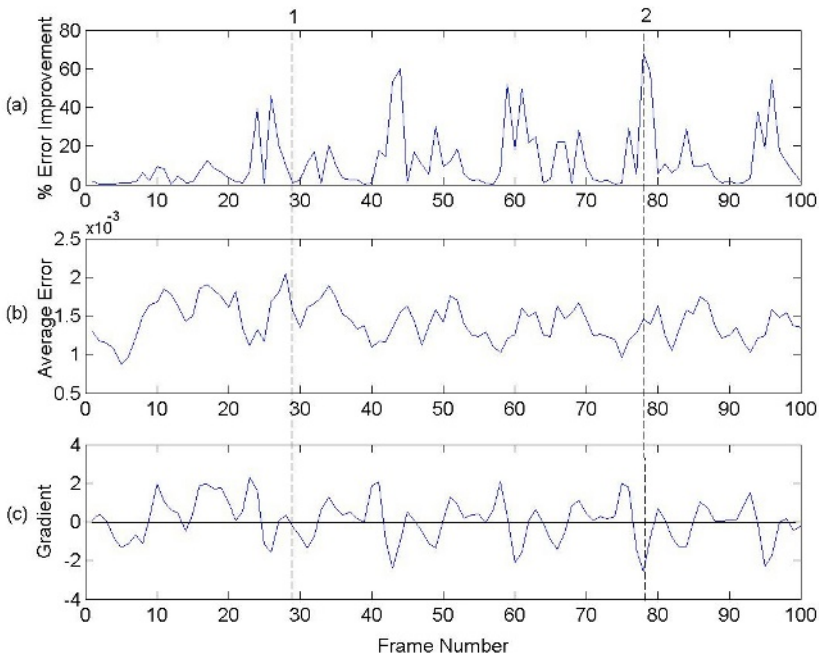

Fig. 3. (a) % Error improvement from the initial disparity to the final disparity in each image pair for 100 frames. (b) The mean image difference between the left and right warped images in the corresponding sequence. (c) Derivative of the heart motion indicates that large error improvements are partly due to considerable difference between the current and the previous frames.

5   Conclusion

The algorithm has demonstrated its ability to track the coherent motion of the heart. We pose the tracking problem as a time-varying optimization of a parametric disparity map describing the depth. The algorithm computes the surface geometry directly from image intensity data. It is robust, accurate, and the image processing can be done in real-time. When couple with the region tracking algorithms described in [9], the package can be used as a virtual stabilization system so that CABG can potentially be performed safely without the need of a heart-lung machine or a mechanical stabilizer. The main issue lies on the robot's ability to react accurately to relatively small motions at high frequencies. In the future, we will also explore the utility of this technique in other clinical areas, with one of them being the assessment of regional myocardial functions.

References

1. Kirklin, J.K., Westaby, S., Blackstone, E.H., et al. Complement and the damaging effects of cardiopulmonary bypass. *J Thorac Cardiovasc Surg* 86:845–57, 1983.
2. Roach, G.W., Kanchuger, M., Mangano, C.M., et al. Adverse cerebral outcomes after coronary bypass surgery. *N Engl J Med* 335:1857-63, 1996.
3. Mack, M.J., Pfister, A., Bachand, D., et al. Comparison of coronary bypass surgery with and without cardiopulmonary bypass in patients with multivessel disease. *J Thorac Cardiovasc Surg* 127(1):167-73, 2004.
4. Mack, M.J. Advances in the treatment of coronary artery disease. *Ann Thorac Surg* 76(6):S2240-5, 2003.
5. Boyd, W.D., Rayman, R., Desai, N.D., et al. Closed-chest coronary artery bypass grafting on the beating heart with the use of a computer-enhanced surgical robotic system. *J Thorac Cardiovasc Surg* 120:807–9, 2000.
6. Mohr, F.W., Falk, V., Diegeler, A., et al. Computer-enhanced robotic cardiac surgery—experience in 148 patients. *J Thorac Cardiovasc Surg* 121:842-53, 2001.
7. Detter C., Deuse T., Christ F., et al. Comparison of two stabilizer concepts for off-pump coronary artery bypass grafting. *Ann Thorac Surg* 74:497-501, 2002.
8. Ramey, N.A., Corso, J.J., Lau, W.W., et al. Real time 3D surface tracking and its applications. Accepted to *Proceedings of Workshop on Real-time 3D Sensors and Their Use (at CVPR 2004)*, 2004.
9. Hager, G.D., and Belhumeur P.N. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans Pattern Analysis and Machine Intelligence* 20(10):1025-39, 1998.