2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Real-Time Visual Servoing of a Robot Using Three-Dimensional Ultrasound

Paul M. Novotny*, Jeffrey A. Stoll†,
Pierre E. Dupont†,and Robert D. Howe*
*School of Engineering and Applied Sciences
Harvard University, Cambridge, MA USA
Email: howe@deas.harvard.edu
†The Department of Aerospace and Mechanical Engineering
Boston University, Boston, MA USA

Abstract— **This paper presents a robotic system capable of using three-dimensional ultrasound to guide a surgical instrument to a tracked target location. Tracking of both the surgical instrument and target was done using image based algorithms on the real-time 3D ultrasound data. The tracking techniques are shown to be especially amenable for execution on powerful graphics processor units. By harnessing a graphics card, it was possible to detect both a surgical instrument and a surgical target at a rate of 25 Hz. The high update rate permits the use of tracked instrument and target locations for controlling a robot. Validation of the system was done in a water tank, where the robot moved the instrument to the target site with a mean error of 1.2 mm.**

I. Introduction

Real-time three-dimensional ultrasound (3DUS) has been demonstrated as a viable tool for guiding surgical procedures [1]. This visualization technique may enable a range of new minimally invasive techniques in cardiac surgery. For example, beating heart intracardiac procedures are now possible with the use of 3DUS and minimally invasive instruments [2]. Ultrasound permits visualization through the opaque blood pool in the heart, and the advent of real-time 3DUS overcomes difficulties with 3D spatial perception in conventional 2D ultrasound [1]. These procedures eliminate the need for a cardiopulminary bypass and its well documented adverse effects, including delay of neural development in children, mechanical damage from inserting tubing into the major vessels, increased stroke risk, and significant decline in acute and chronic cognitive performance [3] [4] [5]. However, initial animal trials highlighted several obstacles to clinical implementation of ultrasound-guided intracardiac surgery [2]. The 3D ultrasound provides limited spatial perception due to distorted appearance of tissue and instruments (Fig. 1). The low quality imaging when combined with the confined dynamic environment of a beating heart makes complex tasks difficult for surgeons to perform.

Surgical robotics is a promising tool for working under these conditions. By allowing a robot to perform a set of surgical tasks autonomously, the procedure would be greatly simplified. A common task in surgical procedures is inserting an anchor into tissue. For example, in atrial septal defect repair a patch must be attached to the atrial septum to prevent blood flow between the two atria. With ultrasound guidance, a surgical robot could perform this task by autonomously moving to the target site and deploying an anchor. Several researchers have reported success using visual servoing techniques to control a robot under ultrasound guidance. Early work done by Stoll *et al.* [6] used 2D ultrasound to visually servo a robot end effector to positions within the ultrasound image. More recently, Vitrani *et al.* [7] tracked the fingers of a endoscopic grasper with 2D ultrasound for robot control. In addition, there have been numerous reports of guiding surgical needles with ultrasound [8] [9]. These techniques, however, use 2D ultrasound in homogeneous static tissue, which limits the applicability to the heterogeneous dynamic environment found in intracardiac procedures. Recently it

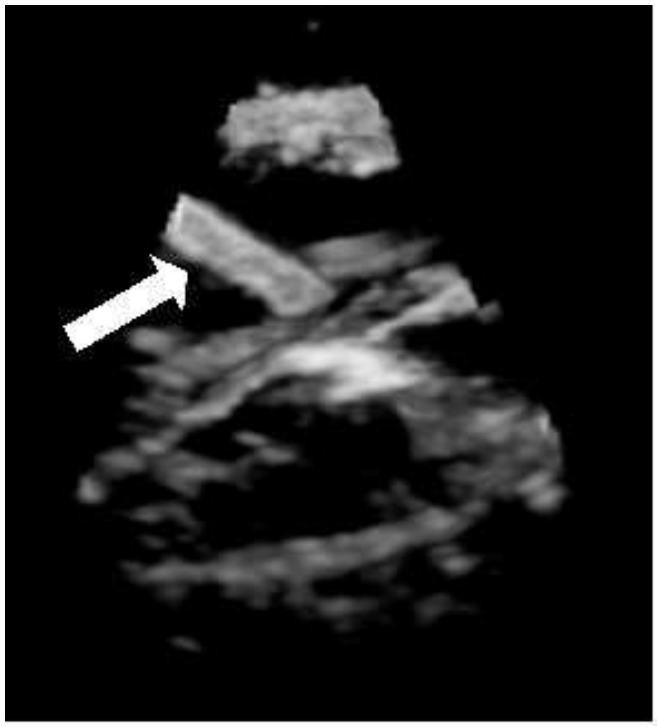

Fig. 1. *In vivo* 3D ultrasound image of a porcine heart. Arrow points to a 5 mm instrument inserted through the heart wall into the right atrium.

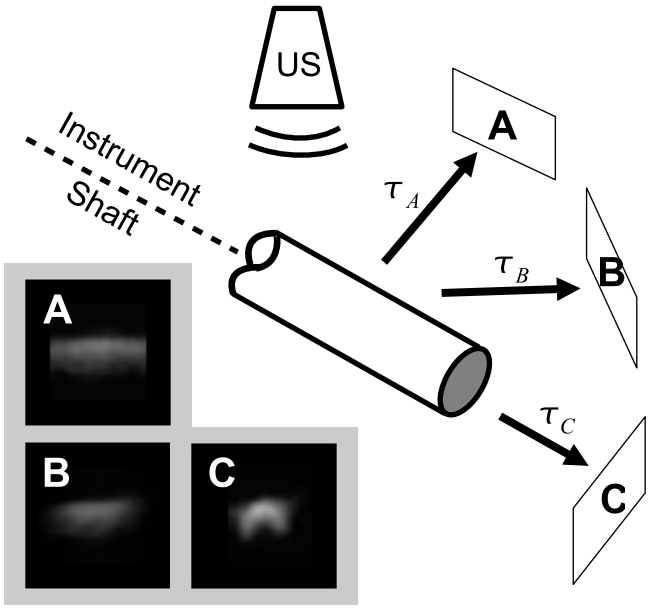

Fig. 2. Example of the modified Radon transform performed an ultrasound image of an instrument. Each image (A-C) is a projection of the ultrasound image along the corresponding direction shown in the schematic. The projection along the axis of the instrument (C) is the brightest. Note that this diagram omits out-of-plane projections that are part of the implementation.

has been shown that 3D ultrasound can be used for visual servoing in a water tank [10]. However, tracking speeds were limited to 2Hz.

The work presented here uses 3D ultrasound for visual servoing of a robot for surgical procedures. It builds upon the previous work in 3D ultrasound based visual servoing and addresses its shortcomings. First, the speed of the instrument tracking is addressed by using faster tracking algorithms that run at the 25 Hz update rate of the ultrasound machine. In addition to tracking the instrument, as others have done, we describe a technique to use the ultrasound volume to track a target in parallel with the instrument tracking. With both a tracked surgical instrument and a target site, we verify the utility of these tracking techniques by visually servoing a small robot in a water tank.

II. Methods and Materials

A. Instrument Tracking

To attain real-time tracking of instruments within an 3DUS image, we utilized techniques that are especially amenable to high performance implementations. This consideration is especially important due to the high data rates of the real-time 3DUS, where volumes (204x48x148 voxels) are produced at a rate of 25 Hz. This high data rate requires that real-time algorithms must handle 30 MB of data every second. For this challenge, we build upon work presented by Novotny *et al.* [11], that introduces a high performance instrument shaft detection algorithm. We also incorporate passive markers similar to those presented in [12].

1) Instrument Shaft Tracking: The first step is to detect the axis of the instrument. Generally, minimally invasive instruments such as those used in intracardiac procedures are fundamentally cylindrical in shape (Fig. 1). Graspers, scissors, needle drivers, and clip deployment devices are all attached to a long cylindrical tube. A modified form of the Radon transform has been shown to be capable of identifying these long straight cylinders [11], a feature that is not found in cardiac ultrasound images. The algorithm identifies maximums of the the equation:

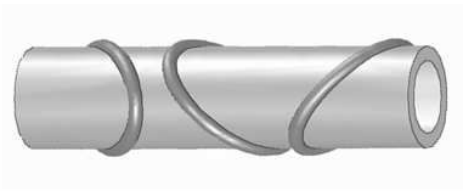

Fig. 3. Schematic of passive marker placed on a surgical instrument. The relative positions of the three bumps encode the roll angle and tip position of the surgical instrument.

$$\breve{g}\left(P,\theta,\phi\right) = \int_{-r}^{r} g\left(\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} + \begin{bmatrix} cos\theta cos\phi \\ sin\theta cos\phi \\ sin\phi \end{bmatrix} s\right) ds \qquad (1)$$

where $\breve{g}$ is referred to as the modified Radon space of the ultrasound volume, g. Each point in $\breve{g}$ corresponds to the integral along a three dimensional line defined by a position, P, and two angles, θ and ϕ. Identifying lines in a 3D volume now becomes a problem of finding local maximums of $\breve{g}$. In other words, we integrate the image volume, g, along a direction defined by (θ, ϕ), through the point P and identify maximums. This scheme is illustrated in Fig. 2 where integrations are illustrated for multiple directions. Fig. 2c contains points with a high integral, or simply, the image is brighter than the other two images (Fig. 2a, 2b). This is a result of the correspondence of the integration and the object's axis. As a result, by finding the maximum value of $\breve{g}$, the axis of the instrument in 3D space is implicitly defined by the parameters (P, θ, ϕ).

This algorithm is especially well suited for implementation on parallel architectures such as modern graphics processing units [11]. To enable real-time tracking, the instrument axis tracking algorithm is calculated on the GPU where the 3DUS volume is stored in texture memory. Each calculation of Eqn. 1 is performed by a pixel shader on the GPU. These pixel shaders take advantage of high performance volume interpolation and vector manipulation built into the GPU. In addition, the GPU used in the implementation (7800GT, nVidia Corp, Santa Clara, CA) has 16 pipelines that calculate Eqn. 1 simultaneously for 16 different positions in the $\breve{g}$ space. Once the $\breve{g}$ space has been sufficiently sampled, the maximum identifies the position of the instrument axis.

2) Passive Marker Tracking: Once the axis of the instruments is found, it is necessary to detect the final two degrees of freedom of the instrument (tip position and roll angle) to fully define its position and orientation. An instrument marker, shown in Fig. 3, was fixed to the instrument shaft

Fig. 4. 3D ultrasound image of a surgical instrument in a water tank. The results of the passive marker tracking is indicated by the white dots.

Fig. 5. 3D ultrasound image of the target X. The target was composed of two intersecting strands of nylon. This fiducial was designed to be attached to surgical targets, such as a patch.

near the tip. It consisted of raised ridges on the surface of the instrument, following prescribed paths, such that, when imaged, they uniquely indicated the remaining degrees of freedom.

Finding the markers begins with the image volume already loaded into texture memory from the Radon transform algorithm. The built-in tri-linear interpolation is then used to quickly render a slice through the instrument axis, orientated so that the instrument axis is horizontal. In this slice, the marker ridges appear as bumps on the instrument surface. The position of these bumps along the instrument shaft axis was determined via a template matching algorithm. The positions of the three best matches found in the slice are used to determine the tip position of the instrument and the roll angle (Fig. 4). The instrument tip position is found by an offset of 1 cm from the marker center, defined as the average location of the marker bumps. The instrument roll angle is found by comparing the relative bump spacing to the known marker shape.

Previous work showed in tank test that the algorithm performed with RMS positional errors of 0.7 mm and angular errors of 0.5 degrees [11].

B. Target Tracking

In addition to tracking the surgical instrument, the target patch must also be tracked. In the case of atrial septal defect repair, the target is a location on a patch. By adding distinctive markers to this patch, as done with instrument tracking, tracking a target structure is greatly simplified. In this case, we investigated tracking an X, a structure that is easily added to a surgical patch and imaged by 3D ultrasound (Fig. 5).

In addition, this fiducial was specifically chosen because detecting two intersecting lines is ideally suited for the modified Radon transform already implemented for instrument axis detection. To identify this marker, it is first necessary to distinguish it from the surgical instrument that is also in the ultrasound image. This is done by identifying the two lines in the ultrasound image that are both intersecting and perpendicular. These conditions ensure that the instrument shaft and target marker are not not mis-identified by the tracking algorithm. The intersection of the lines defines the position of the target, and the cross product defines its orientation.

The target tracking algorithm uses the same Radon based algorithm as the instrument tracking. As a result, we are again able to take advantage of the parallel architecture found on consumer graphics cards to reach real-time performance.

C. Robot Control

The instrument and target tracking algorithms were used for robot control. To autonomously control the instrument tip, the target position must be tracked in the robot coordinate frame. To this end it is necessary to first track the target position in the ultrasound image, p_I and calculate the transformation matrix, T_R^I that converts the target position from image coordinates, to robot coordinates, p_R.

$$p_R = T_R^I p_I \tag{2}$$

As shown in Fig. 6, T_R^I in Eqn. 2 can be broken into three components:

$$T_R^I = T_R^W T_W^M T_M^I \tag{3}$$

T_R^W is the transformation from the robot coordinates to the robot wrist. This transformation is calculated using the joint encoders of the robot and is updated at 1 kHz. T_W^M describes the tip position in relation to the wrist position. This transformation is calculated *a priori* by measuring the instrument length. The final transformation, T_M^I specifies the relationship between the instrument tip and the ultrasound

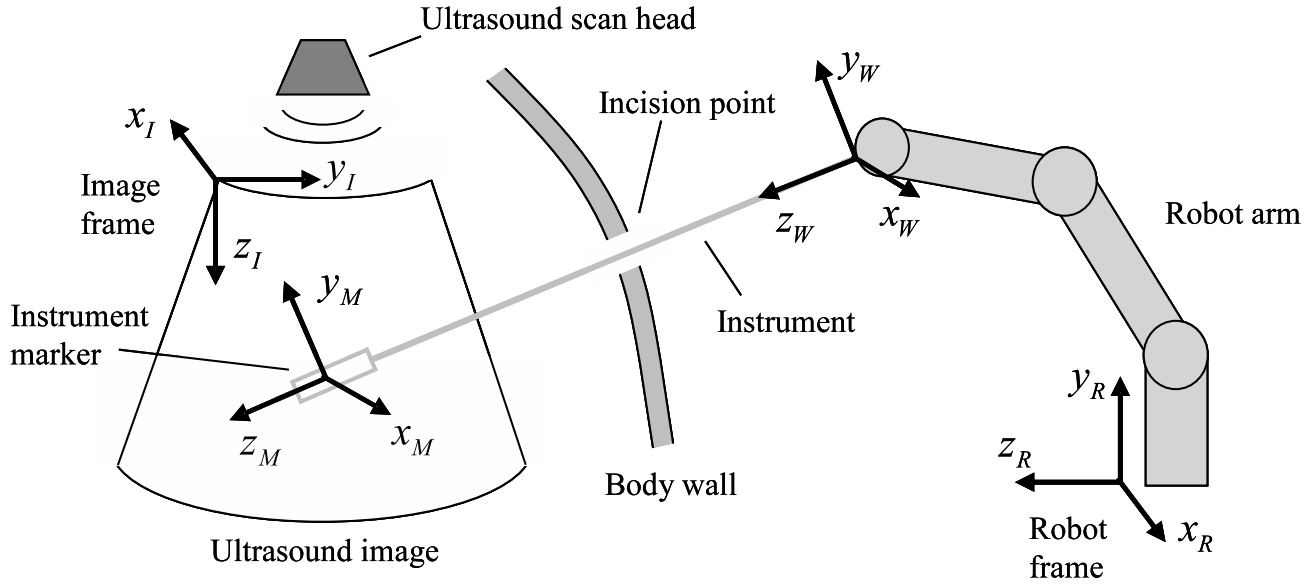

Fig. 6. Schematic of experimental setup. The robot controls an instrument inserted through a port into the patient. The instrument tip is imaged by the 3D ultrasound probe.

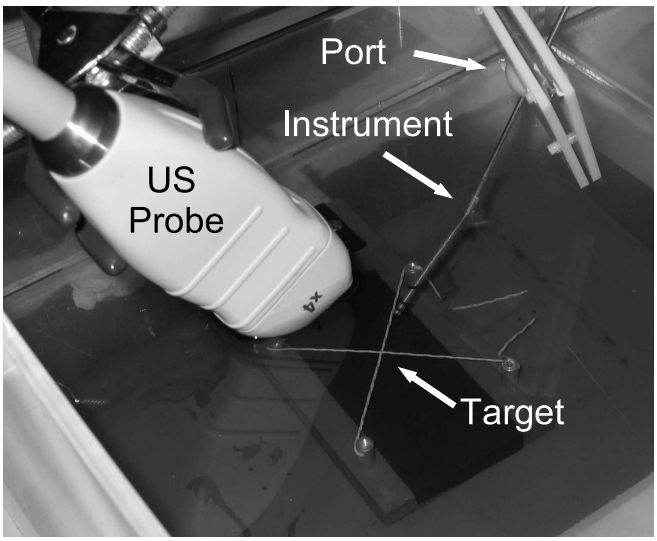

Fig. 7. Image of the test setup. The 3D ultrasound probe is shown imaging the target and the surgical instrument in a water tank. The surgical instrument is attached to a small robot and inserted through a port.

volume. This transformation is calculated by the instrument tracking algorithm and updated at a rate of 25 Hz as the instrument moves within the ultrasound volume.

A PD controller is used to minimize the error between the instrument tip and the target position, p_R, in the robot space. To remove large velocities caused by large differences in instrument tip and target location, a maximum velocity was set to 2 mm/s. This maximum was also chosen as fast speeds are not appropriate during surgery.

D. Experimental Setup

Visual servoing using 3D ultrasound was tested in a water tank. A surgical instrument was attached to the end of a small robot (Phantom, Sensible Technologies, Woburn, MA). The robot was used to move the surgical instrument from a home position to target position. Target positions were defined as offset positions from the target fiducial. The target fiducial was constructed by crossing two lengths of nylon and was held statically at the bottom of the water tank.

The surgical instrument was a custom anchor driver that consists of a 1.5 mm stainless steel tube. Inside the tube is a push rod that inserts the anchors into the target site. The coatings and instrument markers were mounted to the exterior of the instrument. To produce distinct elements, wire was wrapped around the instrument shaft as shown in Fig. 3. Uncoated metals such as the stainless steel used for surgical instruments are highly reflective in ultrasound. To ensure that the instrument is visible in ultrasound, a more diffusive interaction with the ultrasound pulse is desired. As a result, electrical shrink tubing was applied to the instrument in order to improve the appearance.

To simulate an instrument passing through a port in the heart wall, the instrument passes through a sliding ball joint. The target and instrument are both imaged by a real-time 3D ultrasound probe (SONOS 7500, Philips Medical, Andover, MA) in a water tank. (Fig. 7).

In this setup, the ultrasound data is streamed from the ultrasound machine to a personal computer over a gigibit LAN using TCP/IP. The data stream is captured on the target PC and passed to both the instrument tracking and target tracking algorithm running in separate asynchronous threads. Both the tracking algorithms calculate the modified Radon transform on the graphics processing unit (7800GT, nVidia

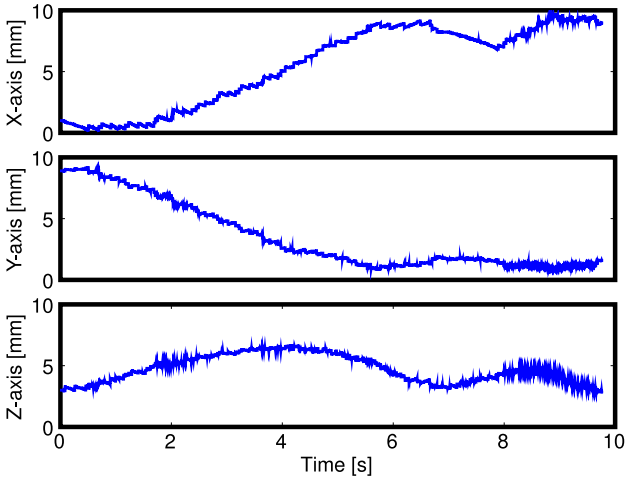

Fig. 8. Example trajectory of the instrument tip as it moves toward the target location. x, y, and z trajectories are shown separately for the length of the trial.

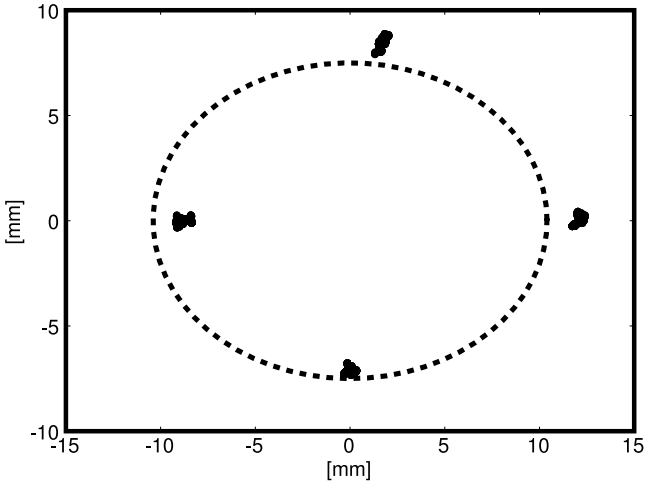

Fig. 9. The instrument was commanded to move to 4 locations. The plot shows the final tip positions for the four commanded locations as circles. The dotted circle indicates desired location.

Corp, Santa Clara, CA). The instrument and target positions are then passed to a third thread that contains the control loop for the robot. The entire system runs on a Pentium 4 3 GHz personal computer with 2 GB of RAM.

To validate the effectiveness of the setup, the robot was commanded to move to 4 distinct target positions. These 4 positions were offset from the center of the tracked target position, separated by 90 degrees around a circle of diameter 1.5 cm. These offset positions are analogous to placing anchors around the circumference of a patch in surgical procedures.

The robot started 1 cm above the center of the target and was commanded to go to one of the offset target positions. The tip position of the robot was monitored and recorded using the joint encoders of the robot. Each of the 4 trials was run until the robot converged on the target location. Data was recorded for an additional 2 seconds while the robot reached steady state at the target location. This procedure was repeated for the 4 target positions.

III. Results

In our experimental setup, the instrument tracking technique required 2 seconds to initially detect the instrument and target in the entire ultrasound volume. For subsequent tracking, the algorithm tracked the instrument and target at 25 Hz, matching the speed of the incoming ultrasound volumes generated by the ultrasound machine.

Fig. 8 shows the trajectory of the instrument tip as it moves from the initial starting point to the target position. The total trajectory is for 10 seconds with a total travel of the instrument tip of 1.2 cm.

In Fig. 9 the final tip position for the 4 positions is shown. For each of the 4 target locations the tip postion is shown for the last second, which results in about 1000 measurements of the tip locations, shown as circles in Fig. 9. The tip locations show good correspondence to the target circle, shown as dashed. Fig. 10 displays the radial error

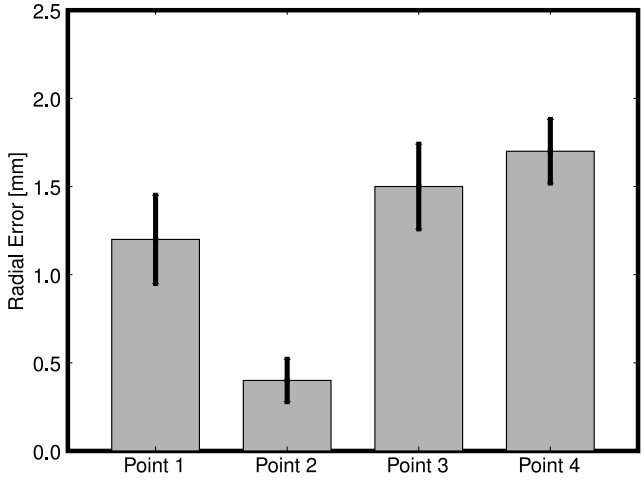

Fig. 10. Mean error and standard deviation of the final tip position from the desired target location.

for each of the four target positions shown in Fig. 9. The mean and standard deviation of the radial error is shown for each point. For the 2 seconds the instrument maintained the position at the target position, it had a mean radial error and standard deviation of 1.2±0.25 mm. Point 2, 3, and 4 had similar radial errors of 0.4±0.12 mm, 1.5±0.24 mm, and 1.7±0.18 mm, respectively. Across all trials the instrument tip converged to on average 1.2 mm from the desired final position.

IV. Discussion

A technique for real-time instrument and target tracking in 3D ultrasound is presented. This technique takes advantage of the computational power of consumer graphics processing units to achieve real-time speeds. To the authors knowledge this is the first such implementation where tracking an instrument and target was performed at the update rates of the 3D ultrasound probe (25 Hz).

With the ability to track an instrument and target at real-time speeds, controlling a robot with ultrasound guidance is now possible. We demonstrated that ultrasound based visual servoing can be used to position an anchor driver with respect to a target location. This autonomous task is potentially useful in surgical procedures with limited visual and physical access, such as intra cardiac surgery. The ability to automatically place anchors accurately greatly reduces the complexity of the procedure and decreases the mental workload on the surgeon.

Tank trials validated the robots ability to accurately move to target locations using the ultrasound for guidance. The robot successfully moved to locations representing the circumference of a circular patch. Average steady state positional error was 1.2 mm and did not exceed 2 mm. This error, while minimal, was largely due to the instrument tracking algorithm. Previous research showed that instrument tracking had positional of 0.7 mm [11]. In the work presented here, when the instrument tracking is combined with target tracking for robot control, the positional error increased by 0.5 mm to an average error of 1.2 mm. This addition 0.5 mm error is most likely due to the target tracking algorithm as it uses the same algorithm as the instrument tracking, and therefore has similar accuracy.

This study did not encapsulate all the challenges that are inherent in performing a procedure inside a beating heart. It has been shown that Radon based tracking algorithm is effective inside a beating heart [11]. However, *in vivo* validation is still necessary for the total system and will be addressed in future animal trials. One major challenge is working with moving target locations, as seen in a beating heart. For these circumstances it will be necessary to incorporate more sophisticated state-space based control algorithms that can adjust to dynamic conditions. In addition, characterizing the target movements will allow for prediction of the current position in between updates from the tracking algorithm.

Additional work will focus on the target fiducial. While the target X used in this study was successfully tracked with the Radon based tracking algorithm, it is still an early design. Incorporating this fiducial in a clinically viable patch will be addressed by future work.

ACKNOWLEDGMENT

The authors would like to thank Jinlan Huang for her help with instrument coatings and markers. This work is supported by the US National Institutes of Health under grant NIH R01 HL073647-01.

REFERENCES

[1] J. W. Cannon, J. A. Stoll, I. S. Salgo, H. B. Knowles, R. D. Howe, P. E. Dupont, G. R. Marx, and P. J. del Nido, "Real time 3-dimensional ultrasound for guiding surgical tasks," *Computer Assisted Surgery*, vol. 8, no. 2, pp. 82–90, 2003.

[2] Y. Suematsu, G. Marx, J. Stoll, P. DuPont, R. Cleveland, R. Howe, J. Triedman, T. Mihaljevic, B. Mora, B. Savord, I. Salgo, and P. del Nido, "Three-dimensional echocardiography-guided beating-heart surgery without cardiopulmonary bypass: a feasibility study." *Journal of Thoracic Cardiovascular Surgery*, vol. 128, pp. 579–587, 2004.

[3] J. M. Murkin, W. D. Boyd, S. Ganapathy, S. J. Adams, and R. C. Peterson, "Beating heart surgery: Why expect less central nervous system morbidity?" *Annals of Thoracic Surgery*, vol. 68, pp. 1498–1501, 1999.

[4] J. Zeitlhofer, S. Asenbaum, C. Spiss, A. Wimmer, N. Mayr, E. Wolner, and L. Deecke, "Central nervous system function after cardiopulmonary bypass," *European Heart Journal*, vol. 14, no. 7, pp. 885–890, 1993.

[5] D. Bellinger, W. D, K. Kuban, L. Rappaport, P. Hickey, G. Wernovsky, R. Jonas, and J. Newburger, "Developmental and neurological status of children at 4 years of age after heart surgery with hypothermic circulatory arrest or low-flow cardiopulmonary bypass," *Circulation*, vol. 100, pp. 526–532, 1999.

[6] J. Stoll, P. Dupont, and R. Howe, "Ultrasound-based servoing of manipulators for telesurgery," in *Telemanipulator and Telepresence Technologies VIII Conference, Proceedings of the SPIE*, vol. 4570, Newton, MA, USA, Oct. 2002, pp. 78–85.

[7] M. Vitrani, G. Morel, and T. Ortmaier, "Automatic guidance of a surgical instrument with ultrasound based visual servoing," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'05)*, Barcelona, Spain, Apr. 2005, pp. 510–515.

[8] K. Draper, C. Blake, L. Gowman, B. Downey, and A. Fenster, "An algorithm for automatic needle localization in ultrasound-guided breast biopsies," *Medical Physics*, vol. 27, pp. 1971–9, 2000.

[9] J. H. et al, "An ultrasound-driven needle-insertion robot for percutaneous cholesystectomy," *Phys. Med. Biol.*, vol. 49, pp. 441–45, 2004.

[10] J. Stoll, P. Novotny, R. Howe, and P. Dupont, "Real-time 3d ultrasound-based servoing of a surgical instrument," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'06)*, Orlando, FL, USA, May 2006.

[11] P. Novotny, J. Stoll, P. Dupont, and R. Howe, "Gpu based real-time instrument tracking with three dimensional ultrasound," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'06)*, Copenhagen, Denmark, Oct. 2006.

[12] J. Stoll and P. Dupont, "Passive markers for ultrasound tracking of surgical instruments," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'05)*, Palm Springs, CA, USA, Oct. 2005.