Automatic Detection and Localization of da Vinci Tool Tips in 3D Ultrasound

Omid Mohareri, Mahdi Ramezani, Troy Adebar, Purang Abolmaesumi, and Septimiu Salcudean

Robotics and Control Laboratory, Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada

Abstract. Radical prostatectomy (RP) is viewed by many as the gold standard treatment for clinically localized prostate cancer. State of the art radical prostatectomy involves the da Vinci surgical system, a laparoscopic robot which provides the surgeon with excellent 3D visualization of the surgical site and improved dexterity over standard laparoscopic instruments. Given the limited field of view of the surgical site in Robot-Assisted Laparoscopic Radical Prostatectomy (RALRP), several groups have proposed the integration of Transrectal Ultrasound (TRUS) imaging in the surgical work flow to assist with the resection of prostate and sparing the Neuro-Vascular Bundle (NVB). Rapid and automatic registration of TRUS imaging coordinates to the da Vinci tools or camera is a critical component of this integration. We propose a fully automatic registration technique based on accurate and automatic localization of robot tool tips pressed against the air-tissue boundary of the prostate, in 3D TRUS. The detection approach uses a multi-scale filtering technique to uniquely identify and localize the tool tip in the ultrasound volume and could also be used to detect other surface fiducials in 3D ultrasound. Feasibility experiments using a phantom and two *ex vivo* tissue samples yield promising results with target registration error (defined as the root mean square distance of corresponding points after registration) of (1.80 mm) that proves the system's accuracy for registering 3D TRUS to the da Vinci surgical system.

Keywords: Robot-assisted surgery, da Vinci surgical robot, 3D ultrasound, fiducial detection.

1 Introduction

Laparoscopic surgery (also known as keyhole or minimally invasive surgery) has various advantages over traditional open surgery, including reduced blood loss, hospital stay, recovery time, and scar-tissue formation. Robot-assisted laparoscopic surgery using the da Vinci Surgical System (Intuitive Surgical, Sunnyvale, CA) is emerging as a new standard of treatment, particularly for urologic procedures such as radical prostatectomy. Augmented reality, implemented as the overlay of medical images data onto a stereoscopic camera view, is one research

P. Abolmaesumi et al. (Eds.): IPCAI 2012, LNAI 7330, pp. 22–32, 2012.

concept aimed at offsetting the visual and haptic limitations of da Vinci laparoscopic surgery. The displayed data can come from several medical imaging modalities such as ultrasound [2,10], fluoroscopy, CT [10] and MRI, depending on the tissue types involved in the procedure [5,7]. All augmented reality systems must include algorithms for accurately registering medical images to camera images in real time.

A current augmented-reality-based research in this area involves registering intraoperative 3D ultrasound images of the prostate to the da Vinci stereoscopic camera view during robot-assisted laparoscopic radical prostatectomy (RALRP). A robotic system for Transrectal ultrasound (TRUS) imaging during RALRP is designed [1] and is being used to capture two-dimensional and three-dimensional B-mode and elastography data. A method for 3D ultrasound to stereoscopic camera registration through an air-tissue boundary [14] is used to register the TRUS images to the da Vinci camera. The same registration technique is also implemented to allow the TRUS robot to automatically track the da Vinci tool tips with the TRUS imaging planes to provide guidance without any distraction to the surgeon. A schematic of this approach and the prostate anatomy during RALRP is shown in Figure 1. This registration method uses da Vinci tool tips or other fiducials pressed against the air-tissue boundary and requires knowledge of the position of the tool tips/fiducials in the ultrasound frame in real time. The da Vinci API or the da Vinci camera are used to provide the location of the tool tips/fiducials in the da Vinci frame or the camera frame. Once fiducials are localized in both frames, the homogeneous transformation between them could be solved using a least squares method.

The process of localizing the da Vinci tool tips or the surface fiducials in 3D ultrasound volumes is done manually and involves scrolling through 2D slices of the volume, finding the 2D slice with the tool tip inside and selecting the approximate center of each fiducial through mouse positioning and clicking. While the overall registration technique is a success, the process of manual fiducial selection

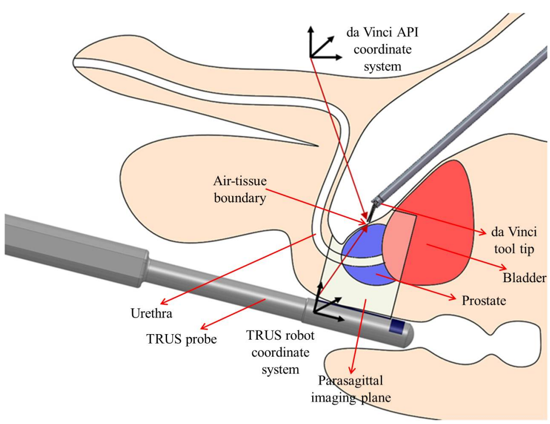


Fig. 1. Air-tissue boundary registration concept for automatic tool-tracking in RALRP

is time consuming and the interpretation of the fiducial centers in series of 2D ultrasound images in the volume is subjective and varies from user to user. An automatic fiducial localization algorithm would be easier to use than the manual selection and also may help to avoid disrupting the surgical workflow, moving the overall augmented reality and tool tracking systems further towards a real time implementation. This work will address the problem of automatic da Vinci tool tip localization in 3D TRUS but the detection algorithm could be easily used for other types of surface fiducials.

Localization of surgical tools in 3D ultrasound has been addressed in a few works [6,11], but to the best of authors' knowledge, there is no report on automatic tool tip localization in 3D Transrectal ultrasound. This problem can be divided into two sub-problems: (i) automatically detecting the presence of the tool tip in the ultrasound volume and finding its slice number, (ii) automatically locating the center of each detected tool tip in the 2D frame. Poon and Rohling [12], in a study on calibration of 3D ultrasound probes (a separate and unrelated type of calibration), used the centroid of an image region around a user-supplied location to semi-automatically detect the center of each fiducial. This simple concept solves the second sub-problem, but not the first. The goal of this study is to recommend a method to target both of the above problems and make the tool tip localization procedure fully automatic without compromising the previously achieved registration accuracy $2.37 \pm 1.15\ mm$ [1].

A multi scale filtering technique based on second order Gaussian derivative and a circular Hough transform is proposed and implemented for da Vinci too tip localization in 3D ultrasound. A 3D mask is created based on the background ultrasound volume (a volume that has been imaged before inserting the tool tip into the tissue) and applied to the ultrasound volume that includes the tool tip. This ultrasound volume is then filtered to find the edges representing the candidate tip locations in the remaining part of the image. Eventually, the tip of the tool is found by using a circular Hough transform. Hence, the tool location is both determined in the ultrasound volume and inside its 2D frame. The same method could be used to localize any surface fiducial pressed against the air-tissue boundary. Experiments have been performed to evaluate the registration accuracy using this automatic fiducial localization method and the results are compared with the manual method. To the best of authors' knowledge, this is the first implementation of an automatic algorithm for detecting da Vinci tools inside 3D ultrasound volumes.

2 Material and Methods

Several key factors are relevant to the selection of a detection algorithm for this problem. The ultrasound data is available in real time as a series of 2D images generated by the 3D ultrasound transducer scanning the body. Although the ultrasound considered in this study is three-dimensional, the actual volumes are created by an off-line scan conversion algorithm and therefore, the detection algorithm must be applied to the sequence of 2D images creating the 3D volume.

The appearance of the target objects is fairly consistent. Ultrasound images of da Vinci tool tips pressed against air-tissue boundaries all contain similar features: strong horizontal lines from the air-tissue boundary and approximately circular areas of high-intensity from the tool tips themselves. The scale of the ultrasound volumes is fixed by the spatial resolution of the ultrasound transducer, so scale invariance is not required. Prostate surgeries will involve fairly consistent relative orientations of transducer and probe, so rotational invariance is likewise not required. Finally, the detection must be very rapid. In an ideal real time augmented reality system, the medical image data displayed in the surgeon's stereo view would be updated at a rate close to that of normal video (i.e. 30 frames per second) and therefore, a detection algorithm that can scan an entire volume in a few seconds is necessary to avoid disrupting the surgical work flow.

2.1 Automatic Detection Algorithm

In this study, the automatic extraction of da Vinci tool tips is done in four steps: masking, filtering, circle detecting and removing the false positives. A schematic of the first three steps of this method is shown in the diagram of Figure 3. The idea behind creating the mask is to detect the air tissue boundary and remove everything which lies below this line corresponding to the air part in the ultrasound image. Figure 2 shows an example image of da Vinci tool tip pressed against the air-tissue boundary of an *ex vivo* liver tissue. To create a 3D mask, first series of 2D images of the background volume have been filtered using a Hessian based Frangi vesselness filter [4]. In this filtering approach, the principal directions in which the maximum changes occur in the gradient vector of the underlying intensity in a small neighborhood are identified. Eigenvalue decomposition of the Hessian matrix can give these directions. Therefore, the Hessian matrix at each pixel of the image for a particular scale (σ) of the Gaussian derivative operator is computed. Then, the eigenvalue decomposition is used to extract two orthonormal directions. Based on the computed eigenvalues (λ_1, λ_2)

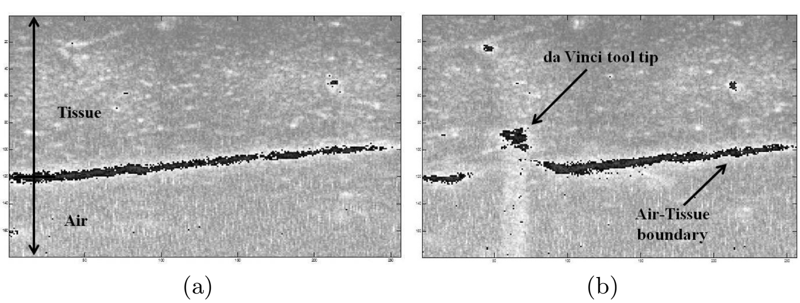


Fig. 2. (a) Example of an air-tissue boundary in an *ex vivo* liver phantom. (b) da Vinci tool tip pressed against an air-tissue boundary.

at each scale ($\Sigma = 1, 3, 5$) , a measure is defined as follows [4]:

$$S = \max_{\sigma \in \Sigma} S(\sigma) = \begin{cases} 0 & \text{if } \lambda_2 > 0 \\ \exp(\frac{R_b^2}{2\beta^2})(1 - \exp(-\frac{R_n^2}{2c^2})) & \text{if } \lambda_2 < 0 \end{cases} \tag{1}$$

$R_b = \frac{\lambda_1}{\lambda_2}$ is the blobness measure in 2D, and $R_n = \sqrt{\lambda_1^2 + \lambda_2^2}$ is a background or noise reduction term. R_b shows deviation from blob-like structures and R_n shows the presence of the structure using the fact that the magnitude of the derivatives (eigenvalues) is small in the background pixels. In this equation β is chosen as 0.5 and c is chosen to be the $\max_{x,y} R_n$. Images are further processed to remove the small areas of high intensity which are randomly scattered in the filtered image and are due to speckles. The image is first thresholded by a threshold obtained from the mean value of the intensities of pixels of the images. Next, parts that have less than M connected components are removed. M is a number obtained by trial and error for TRUS images. The remaining components, which represent the high intensity and relatively large areas corresponding to the tissue structures in the image, are dilated using morphological operators. To find the air-tissue boundary a line detection using Hough transform has been applied to the images. Lines with the minimum length of 30 and minimum gap of 15 pixels between them, have been extracted. A mask has been created using the

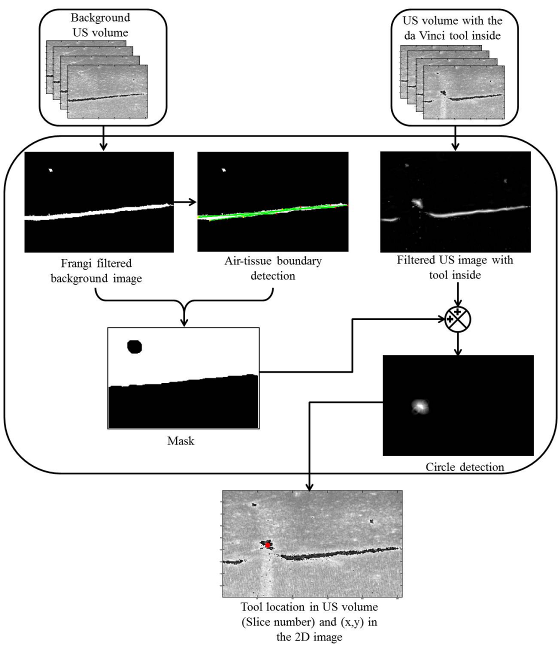


Fig. 3. The structure of the automatic tool detection algorithm. The background ultrasound volume and the volume with the tool inside are the inputs. Slice number and x,y position of the tool tip in the detected slice are the outputs.

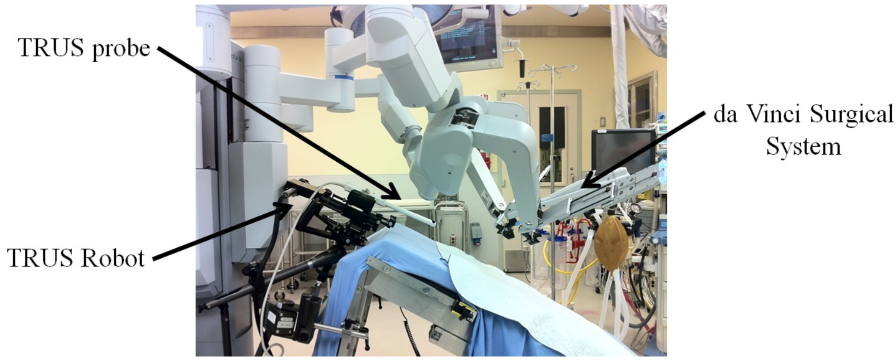


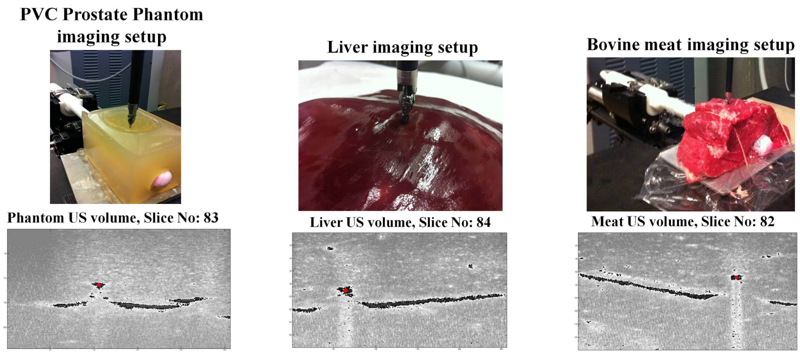
Fig. 4. Experimental setup used for ultrasound imaging: TRUS robot and the da Vinci surgical system

detected air-tissue boundary and artifacts. The mask is responsible to remove regions corresponding to the artifacts and regions outside the tissue boundary.

After applying the mask to the 3D ultrasound volume, the Frangi filter is applied to the series of 2D images and the relatively large components are extracted. A circular Hough transform is applied to the obtained components to find circles with radius of approximately 5 pixels and minimum pixels of 20. The mean location of these circles in each 2D image is computed as candidates of the tool location. Next, a hierarchical clustering algorithm [9] is performed to find the group of candidates which are in adjacent slices, considering the fact that the tool tip can be seen in a couple of consecutive slices. Euclidean distance between the identified candidates has been used as a similarity measure and those which are close to each other are linked to create the cluster. The linkage function continues until an inconsistency coefficient reaches its threshold [8]. Using this approach false positively detected candidates will be removed due to the fact that false positives do not necessarily occur in adjacent slices. Once the cluster of 2D images that have the tool inside them are found, the middle slice in the cluster is chosen to be the output slice for the detection algorithm.

2.2 Experimental Setup for Ultrasound Image Acquisition

The robotic system shown in Figure 4, which is designed for intra-operative TRUS imaging during RALRP, is used for ultrasound image acquisition in this study. The system has three main parts: a robotic ultrasound probe manipulator (robot), an ultrasound machine with a biplane TRUS probe, and control and image processing software. Ultrasound images are captured using a parasagittal/transverse biplane TRUS probe in combination with a Sonix TABLET ultrasound machine (Ultrasonix Medical Corp., Richmond, Canada). 3D ultrasound volumes are collected by rotating the 2D imaging planes and automatically recording the encoder positions for each image. Software running on the ultrasound console is responsible for directing the robot movements and the ultrasound data acquisition. Currently, a



(a) Prostate phantom (b) Liver dataset (c) Bovine meat dataset

Fig. 5. Sample tool detection results for three different data sets

simple graphical user interface (GUI) is being used to allow the user to position the probe, and automatically collect 2D B-mode images and radio-frequency (RF) data while rotating from -40 to +40 degrees. A standard brachytherapy stabilizer arm (Micro-Touch 610-911; CIVCO Medical Solutions, Kalona, IA) is mounted to the operating table and the robot is installed on the stabilizer as it is shown in Figure 4.

3 Results

The proposed automatic tool detection method has been tested on three different acquired data sets: a custom-made PVC prostate phantom, an *ex vivo* liver and *ex vivo* bovine meat. Some sample results of the tool detection algorithm along with the testing configuration for each of the data sets are shown in Figure 5.

The rigid point registration technique proposed by Umeyama in [13], is used in this study to evaluate the automatic detection algorithm's performance. The manual fiducial localization process is replaced by the developed automatic localization algorithm, and the registration accuracy is re-calculated. Da Vinci tool tips are pressed against 12 different points on the air-tissue boundary of each of the three collected data-sets ($N_t = 12$) and its positions (x^0) in the ultrasound robot frame $\{O_0, C_0\}$ is calculated using our automatic detection algorithm. The da Vinci API is used to provide the location of the tool tips (x^1) in the da Vinci frame $\{O_1, C_1\}$.

For each registration experiment, N_f number of points ($N_f \geq 3$) are randomly picked from the total points collected for each tissue type and the rigid point registration method explained in [14] is used to compute the transformation between the frames T_0^1 to minimize the Fiducial Registration Error (FRE). FRE is computed as the root-mean-square of distances between corresponding fiducials after this registration. Next, the rest of the points in each data-set are assumed to be the target points and the calculated transformation is used to transform

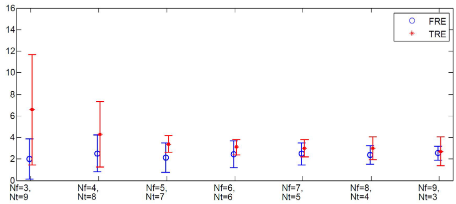


Fig. 6. FRE and TRE and their standard deviations for different numbers of fiducials (N_f) and target points(N_t)

them from the ultrasound robot frame to the da Vinci frame and the Target Registration Error (TRE) is estimated. Estimated TRE includes the real TRE and the target localization error (TLE). TRE is defined as the root-mean-square of distances between corresponding fiducials after registration, i.e., the distance between the localized position of each tool tip as transformed from the ultrasound robot space to da Vinci space and the position of that corresponding tool tip localized in the da Vinci space provided by the API. Because of the Fiducial Localization Error (FLE), the registration is inevitably inaccurate to some extent, and TRE is used as the available measure of registration error. Values of FRE and TRE are computed for different number of fiducials N_f chosen between the total points for each data-set N_t. The mean values of FRE and TRE and their standard deviations are then calculated for each combination of (N_f, N_t) for 100 iterations ($n = 100$) and the results are plotted for the liver dataset in Figure 6. As it can be seen from Figure 6, as (N_f) increases, both mean and standard deviation of TRE decreases and based on this analysis, the number of fiducials suggested for this registration is ($N_f = 5$). The mean and standard deviation values of the target registration error (TRE) for this number of fiducials and 100 iterations are reported in the anatomical frame of the patient in Table 1. Tool tip segmentation error is also calculated in (x, y, θ) directions. Three subjects were asked to independently identify the location of the tool in ultrasound images. The difference between the average user defined positions and the result of our algorithm, in addition to the inter-subject variations are calculated. For the liver dataset, automatic segmentation error is: $e(x, y) = 3.11 \pm 0.88\ mm$,

Table 1. Mean errors ($n = 100$) between tool tip location and predicted location based on registration. Errors are presented in the anatomical frame of the patient, along the superior-inferior (e_{S-I}), medial-lateral (e_{M-L}) and anterior-posterior (e_{A-P}) axes.

	$TRE_{A-P}\ mm$	$TRE_{S-I}\ mm$	$TRE_{M-L}\ mm$	Mean TRE mm
Phantom	0.62 ± 0.31	1.29 ± 0.22	0.82 ± 0.24	1.80 ± 0.32
Liver	0.72 ± 0.83	2.65 ± 0.51	1.18 ± 0.54	3.33 ± 0.81
Bovine meat	1.90 ± 0.79	1.24 ± 0.34	3.51 ± 0.70	4.54 ± 0.88

$e(\theta) = 0.79° \pm 0.39°$ and the inter-subject variation is: $e(x, y) = 4.25 \pm 1.45\ mm$, $e(\theta) = 2.05° \pm 0.8°$. For the pvc phantom dataset, automatic segmentation error is: $e(x, y) = 2.13 \pm 0.97\ mm$, $e(\theta) = 0.67° \pm 0.34°$ and the inter-subject variation is: $e(x, y) = 3.65 \pm 0.88\ mm$, $e(\theta) = 1.55° \pm 0.39°$. For the bovine meat dataset, automatic segmentation error is: $e(x, y) = 2.42 \pm 1.19\ mm$, $e(\theta) = 0.68° \pm 0.48°$ and the inter-subject variation is: $e(x, y) = 3.13 \pm 0.88\ mm$, $e(\theta) = 1.43° \pm 0.44°$.

4 Discussion

The overall average TRE previously reported using the manual fiducial localization of three fiducials on the PVC tissue phantom was $2.37 \pm 1.15\ mm$. The overall average TRE using the automatic detection technique in this study is $1.80 \pm 0.32\ mm$ for the recommended number of fiducial points, $N_f = 5$, on the PVC tissue phantom. As the number of fiducials is increased to 9, the TRE value reaches its minimum calculated value which is $1.61 \pm 0.39\ mm$, which is consistent with the previously reported theoretical analysis [3]. To further evaluate the algorithm, experiments have been done on *ex vivo* liver and bovine meat which are more similar to the human prostate tissue. The minimum calculated TRE value is $2.86 \pm 1.40\ mm$ for the liver and $4.15 \pm 0.61\ mm$ for the bovine meat for picking 9 fiducials on the air-tissue boundary. According to these results, the automatic detection algorithm could yield the same registration accuracy as the manual detection method does, and it could also compensate for errors resulting from mis-interpretation of tool location in the ultrasound volume. The goal of tracking is to have the tool tips appear in the TRUS images and errors in the axial and lateral ultrasound directions are irrelevant as long as the tool tips are within the image boundaries. And because the thickness of the TRUS beam at the anterior surface of the prostate is on the order of millimeters, small errors in the elevational direction likely are not critical. We propose choosing five tool positions on the tissue surface, $N_f = 5$, both to achieve an acceptable registration error and to have a reasonable number of tissue poking repetitions during the surgical procedure. Because this registration technique is designed for RALRP procedure, it is recommended to choose two points on the prostate apex, one point on the mid-gland and two points on the prostate base. In this process, da Vinci tool is slightly pressed on the surface of the tissue at different locations. Hence, there won't be a significant movement in the organ and only the air-tissue boundary is moved about $2 - 3\ mm$ at locations where the tool tip is placed. Also, only the tip of the tool (i.e., the area that touches the tissue) could be seen and detected in the ultrasound images. The instrument shaft could not be seen in the images and could not be used as an additional feature.

Replacing the process of manual detection of tool tips or surface fiducials in the ultrasound volume will accelerate the registration time and reduces the amount of surgical work-flow disruption. There will be no need for a sonographer to attend the surgery to find the tool tip in the ultrasound volume and the algorithm will reliably find the tool tips and send the points to the registration module to calculate the transformation. In addition, the results will not be dependent on the person who is choosing the points in the ultrasound volumes as it is the case in the manual detection.

5 Conclusions

In this study, we have addressed the problem of detecting da Vinci tool tips pressed against an air-tissue boundary, in a 3D ultrasound volume. A method based on multi-scale filtering and circle detection has been proposed. The tool tip localization accuracy is evaluated by analyzing the registration error between the TRUS robot frame and the da Vinci frame. Results show the equivalency of the proposed method and the previously reported manual detection procedure. As an overall comment on the proposed method, it is to be stressed that the method has a significant improving effect on both the duration,complexity and accuracy of the registration procedure. Future work will involve investigating *in vivo* studies to verify the accuracy and reliability of the proposed technique during RALRP procedure.

References

1. Adebar, T., Salcudean, S., Mahdavi, S., Moradi, M., Nguan, C., Goldenberg, L.: A Robotic System for Intra-operative Trans-Rectal Ultrasound and Ultrasound Elastography in Radical Prostatectomy. In: Taylor, R.H., Yang, G.-Z. (eds.) IPCAI 2011. LNCS, vol. 6689, pp. 79–89. Springer, Heidelberg (2011)
2. Cheung, C.L., Wedlake, C., Moore, J., et al.: Fusion of stereoscopic video and laparoscopic ultrasound for minimally invasive partial nephrectomy. In: Proceedings of SPIE Medical Imaging, vol. 7261, pp. 1–10 (2009)
3. Danilchenko, A., Fitzpatrick, J.M.: General approach to first-order error prediction in rigid point registration. IEEE Trans. on Medical Imaging 30(3), 679–693 (2011)
4. Frangi, A.F.: 3D model-based analysis of vascular and cardiac images (2001)
5. Fuchs, H., Livingston, M.A., Raskar, R., Keller, K., State, A., Crawford, J.R., Rademacher, P., Drake, S.H., Meyer, A.A.: Augmented Reality Visualization for Laparoscopic Surgery. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 934–943. Springer, Heidelberg (1998)
6. Gaufillet, F., Liegbott, H., Uhercik, M., et al.: 3d ultrasound real-time monitoring of surgical tools. In: IEEE Ultrasonics Symposium (IUS), pp. 2360–2363 (2010)
7. Grimson, E., Leventon, M., Ettinger, G., Chabrerie, A., Ozlen, F., Nakajima, S., Atsumi, H., Kikinis, R., Black, P.: Clinical Experience with a High Precision Image-Guided Neurosurgery System. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 63–73. Springer, Heidelberg (1998)
8. Guide, M.U.: The mathworks. Inc., Natick, MA 5 (1998)
9. Johnson, S.C.: Hierarchical clustering schemes. Psychometrika 32(3), 241–254 (1967)
10. Linte, C.A., Moore, J., Wiles, A.D., et al.: Virtual reality-enhanced ultrasound guidance: A novel technique for intracardiac interventions. Computer Aided Surgery 13(2), 82–94 (2008)
11. Novotny, P.M., Stoll, J.A., Vasilyev, N.V., Nido, P.J.D., Dupont, P.E., Zickler, T.E., Howe, R.D.: Gpu based real-time instrument tracking with three-dimensional ultrasound. Medical Image Analysis 11(5), 458–464 (2007)

12. Poon, T.C., Rohling, R.N.: Tracking a 3-d ultrasound probe with constantly visible fiducials. Ultrasound in Medicine & Biology 33(1), 152–157 (2007)
13. Umeyama, S.: Least-squares estimation of transformation parameters between two point patterns. IEEE Trans. on PAMI 13(4), 376–380 (1991)
14. Yip, M.C., Adebar, T.K., Rohling, R.N., Salcudean, S.E., Nguan, C.Y.: 3D Ultrasound to Stereoscopic Camera Registration through an Air-Tissue Boundary. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010. LNCS, vol. 6362, pp. 626–634. Springer, Heidelberg (2010)