

Accuracy Analysis of an Image Guided Robotic Urology Surgery System

Steve Thompson¹, Graeme Penney², Prokar Dasgupta³, and Dave Hawkes¹

¹ Centre for Medical Image Computing, University College London.

² Interdisciplinary Medical Imaging Group, Kings College London.

³ Department of Urology, Guys and St Thomas Hospital, London.

Abstract. We present an evaluation of the accuracy of a system for image guided radical prostatectomy using the daVinci telemanipulator. The system is split into components and ten sources of error identified. The magnitude of three of these error sources; segmentation of bone from MRI, registration to patient using intraoperative ultrasound, and endoscope tracking error is determined experimentally. The remaining errors are estimated from the literature. We demonstrate that the distribution of ultrasound slices used for registration can reduce the system error by up to 0.7mm. Our results show that our system can localise the prostate to within 3.7mm RMS, and that the largest component of this error is the segmentation of the pelvic bone from MRI.

1 Introduction

Introducing image guidance to laparoscopic prostatectomy has the potential to reduce the occurrence of positive margins and improve post operative potency by allowing the surgeon to intra operatively visualise the locations of tumours and neuro-vascular bundles. In this paper we present an image guidance system being developed for use during radical prostatectomy using the daVinci⁴ system.

This paper seeks to determine the accuracy with which pelvic anatomy can be overlaid with video data for this system and identify the main sources of overlay error. The system is described in Section 2 and the overall system error divided into ten discrete sources. Three of the main sources of error are investigated by experiments detailed in Sections 3, 4, and 5. Other error sources are obtained from the literature and discussed in Section 2.

2 System Description

Figure 1 details the system and defines the ten sources of error. Prior to surgery the prostate and surrounding structures are manually segmented from patient MRI scans. The resulting patient model which can be overlaid on the endoscope video by registering the model to the patient and tracking the endoscope. The

⁴ Intuitive Surgical Inc. Sunnyvale, CA

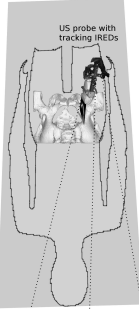
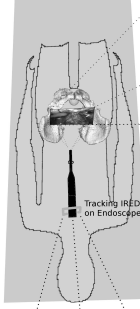
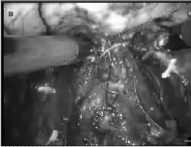
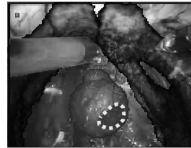
Preoperative	In Theatre	During Surgery
Make Model Manually segment prostate, seminal vesicles tumours, neuro-vascular bundles, urethra. Automatic segmentation of the pelvic bone.	Register Model Calibrate Scope  Calibrate endoscope cameras. Use tracked ultrasound probe to find pelvic bone and match to model.	Track Endoscope Project Video   Endoscope Video Data + Tracking Data + Calibration Data + Registered PreOp Model <hr/> Endoscope Video with Model Overlay 
Errors 1. Manual Segmentation Not Considered 2. Bone Segmentation See Section 3	3. Registration Error See Section 4 4. Ultrasound Calibration From Literature 5. Camera Calibration From Literature 6. Soft Tissue Motion From Literature	7. Endoscope Tracking See Section 5 8. Intraoperative Soft Tissue Motion Not Considered 9. Change in endoscope calibration Not Considered 10. Whole Patient Motion Not Considered
0mm	0.49mm	0.80mm
2.81mm	0.8mm	0mm
	1pixel	0mm
	2mm	0mm
Complete System Error = 3.7 mm		

Fig. 1. Schematic of image guided prostatectomy system. The system is divided into three sections. The ten component errors are also defined and quantified. The system error is the sum of squares of the component errors.

registration is achieved by locating the pelvic bone in theatre using tracked ultrasound. The pelvic bone is used in preference to skin based fiducial markers as it will not be affected by insufflation of the patient abdominal cavity, and is closely coupled to the prostate via the urethra. Furthermore, the pelvis is a large bone with the prostate positioned near its centre of gravity. The pelvis registration error will therefore be reduced when applied to the prostate. Using the pelvic bone for registration requires that the pelvic bone is segmented from the preoperative MRI data. It is not practical to manually segment the pelvic bone, so an automatic bone segmentation algorithm must form a key part of the system.

The pelvic bone is imaged in theatre using a B mode ultrasound probe tracked using an Optotrak Certus.⁵ The endoscope was also tracked using the Optotrak using a collar mounted around the endoscope body as shown in Figure 2. This was used in preference to the daVinci forward kinematic data to allow the system to work independently of the daVinci and to enable later comparison of the accuracy of both systems. It is assumed that each error is independent and

⁵ Northern Digital Inc. Ontario, Canada

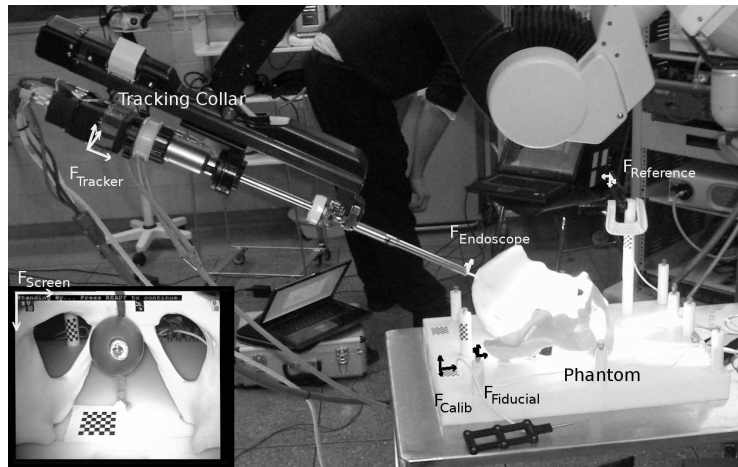


Fig. 2. The pelvic anatomy phantom in theatre, showing the fiducial markers on the phantom and the tracking collar on the endoscope. The six coordinate systems are also noted.

normally distributed, so the system error can be found by taking the root of the sum of squares of the component errors. A discussion of each of the errors defined in Figure 1 follows.

1. **Manual segmentation of soft tissues.** Determination of the accuracy of manual segmentation of soft tissues is difficult, requiring multi modality and cadaver studies. Applicable values are not in the literature and determination of these errors is beyond the scope of this report.
2. **Automatic bone segmentation.** See Section 3.
3. **Registration error.** See Section 4.
4. **Ultrasound calibration error.** The analysis in Section 4 is based on ultrasound data collected in a water bath. However it is known from Penney et al.[1] that the mismatch in the speed of sound between the water-based calibration and human tissue causes a registration error when used on real data. Penney et al. found that this mismatch increased registration error for the pelvic bone up to 0.8mm.
5. **Endoscope calibration error.** Endoscope calibration is performed by viewing a static calibration grid as previously performed on the daVinci by Mourgues et al. [2]. They report 1 pixel error, which corresponds to approximately 0.2mm.
6. **Soft tissue motion.** This refers to motion of the prostate relative to the pelvic bone between preoperative imaging and registration. Based on the results reported by Hoogeman et al.[3] we have estimated these movements to be 2mm Root Mean Square (RMS). By controlling the scan position and bladder/rectum filling this figure may be reduced.
7. **Camera tracking.** See Section 5.

8. **Intraoperative soft tissue motion.** Once the surgeon begins cutting the tissues of interest, the model will become obsolete. This system is intended to provide the surgeon with an accurate overlay of the model only when the prostate first becomes visible, (after navigation around the bladder) allowing the surgical plan to be updated prior to prostate removal. We expect that prostate will not have moved significantly at this stage due to its close coupling with the pelvis via the urethra.
9. **Change in endoscope calibration.** Multiple endoscope lenses are used during a typical prostatectomy. Separate calibrations can be performed on each lens preoperatively and the system updated each time the lens is changed. Furthermore, we have observed that while it is possible to alter the endoscopes focal length during the procedure, this is not typically done.
10. **Whole patient motion.** Registration is performed with the patient in the operative position. Based on observations from current procedures the patient does not move during surgery.

3 Effect of Bone Segmentation Errors

3.1 Method

The boundary between soft tissue and bone in pelvic MRI images is often indistinct and of variable appearance. Segmentation based solely on image appearance is therefore not robust. By constraining the segmentation, using prior information of the bone shape, a more robust segmentation is possible. Schmid et al.[4] and Thompson et al.[5] present promising results for segmenting the pelvic bone from MRI using shape models. We use the method proposed by Thompson et al. [5]. Our shape model is built from 21 CT data sets, where the twelve largest modes of variation are used for segmentation.

The downside of using a shape model to constrain the segmentation is that there will inevitably be segmentation errors caused by the mismatch between the best matching model shape and the actual shape. What follows is an investigation to quantify what effect these will have on prostate localisation error. This analysis excludes the optimisation error which will depend on the actual algorithm used to optimise the shape parameters of the model. We also assume that the model training set is representative of the patient population. This assumption will be tested as more data is collected.

3.2 Model Cross Validation

The 21 data sets used to build the model were used in a cross validation procedure to determine the shape modelling error. Segmentations for each image were first generated by propagating a manual segmentation from a single training set using the same non-rigid registration algorithm (Crum et al.[6]) as was used to construct the model. As this included registration errors we term it a bronze standard (which was checked visually) and a corresponding bronze standard deformation vector. Each training set was then segmented using a shape model

built from the 20 other training sets. This was done by solving $\mathbf{b} = \Phi^T(\mathbf{S} - \bar{\mathbf{S}})$ where \mathbf{b} is the optimal shape weight vector, Φ is the matrix of deformation modes, $\bar{\mathbf{S}}$ is the model mean and \mathbf{S} the bronze standard deformation vector. A manual segmentation was then propagated using \mathbf{b} , giving a model based segmentation.

3.3 Calculating Registration Error

The 21 bronze standard and model-based segmentation were now compared to determine the effect of the model-based segmentation on the Target Registration Error (TRE) at the prostate. For each training image a set of ideal ultrasound slices was created by slicing (using tri-linear interpolation) the bronze standard segmentation. Physically realistic ultrasound planes collected during the experiment described in Section 4 were used for this. These slices were then rigidly registered to the model derived segmentation using the registration algorithm described in Section 4. As the registration algorithm is sensitive to its starting position, the registrations were each repeated 100 times, starting from positions randomly and uniformly spread over a six dimensional (one dimension per degree of freedom) hypersphere of radius 10mm (rotations were scaled as $1^\circ = 4\text{mm}$). This radius should correspond with the initialisation distance that can be achieved by manual alignment of the data sets in theatre. The resulting registrations were then used to calculate the TRE at six points on the surface of a nominal prostate.

The registration algorithm described in Section 4 is sensitive to the number and spatial distribution of ultrasound slices used. As the objective function (Equation 1) is not normalised, the registration can be biased to be more accurate in a given region by simply including more ultrasound slices from that region. This is of particular importance when registering to an imperfect segmentation of the bone, as the registration can be biased to regions where the segmentation is known to be more accurate. Therefore, each of the model based segmentations were divided into three anatomical regions (left iliac:pubis:right iliac) and the mean surface to surface errors checked. On average the pubic regions were better segmented by the shape model than the iliac regions. To test whether this knowledge could be used to give a better registration result, two methods were used. They each used 420 ultrasound slices. Method A used two slices from each of the iliac regions for every pubic slice (164:82:164), while method B used three pubic slices for every four iliac slices (120:180:120). The actual slices used were randomly chosen from a set of 654 possible slices for each repeat registration.

3.4 Results

For each slice distribution method 2100 registrations were performed (21 segmentations \times 100 repeats from different starting positions). Histograms of the resulting TRE at the prostate surface are shown in Figure 3. It can be seen that weighting the registration to the pubis region creates a noticeable improvement

in the overall registration accuracy. With the registration so weighted the RMS registration error due to model based segmentation of the bone is 2.81mm. This figure was calculated after the removal failed registrations. A registration was classed as a failure if its TRE was greater than 3 standard deviations from the mean TRE.

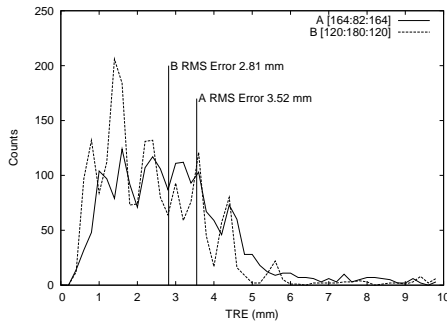


Fig. 3. Target registration error at the prostate surface for registrations using shape-model derived segmentations. Twenty one separate segmentations were tested 100 times each. 2 different distributions of ultrasound slices were used. Method B weights the registration to the pubic region, which results in a more accurate registration. Whilst there are some failed registrations with very high TREs, as this is a visual overlay system these would be spotted by the surgeon and the overlay system either turned off or the registration repeated.

4 Registration in Theatre

Using a tracked ultrasound probe, each pixel on the 2D ultrasound slices can be placed in 3D space via a probe calibration, determined preoperatively using an invariant point method. The ultrasound slices are first converted into “bone edge probability” images as per Penney et al.[1]. For each pixel a “distance to artefact” is calculated then bone edge probabilities are taken from a look-up table of “distance to artefact” and intensity vs bone edge probability. The look-up table is trained on a separate set of manually segmented ultrasound images. The bone edge images are thresholded, discarding pixels with a bone edge probability of zero. The remaining points are stored as a list of n 3D points x_i with corresponding bone edge probabilities EP_i . The segmented MRI data is stored as a binary (BoneEdge=1000, NotEdge=0) voxel image I_{MRI} ($0.71 \times 0.71 \times 2\text{mm}$).

The six degree of freedom registration between the two data sets is found using a gradient descent algorithm. For each potential transform T the objective function is defined by Equation 1 where $I_{MRI}(Tx_i)$ is the voxel value corre-

sponding to point Tx_i , found by nearest neighbour interpolation.

$$F = \sum_{i=1}^n EP_i \times I_{MRI}(T\mathbf{x}_i) \quad (1)$$

The optimiser starts with a step size of 4mm and 0.986 degrees. The step size is halved each time a local maximum is reached. The terminal step size is 0.0625mm. Smaller terminal step sizes did not improve registration accuracy.

At step sizes greater than 0.25mm I_{MRI} is blurred with a Gaussian kernel of 4mm to increase the capture range.

4.1 Registration Accuracy

The registration accuracy is the combination of the ultrasound calibration error, the tracking accuracy of the probe, the error in converting ultrasound intensity to bone edge probability, and the registration algorithm error. This was determined using a pelvis phantom in a water bath. Figure 2 shows the phantom.

A CT image of the phantom data was manually segmented to give a gold standard bone segmentation. Ultrasound slices were collected and converted to bone edge probability maps. The ultrasound data was then registered to the CT data as per Equation 1. The resulting registration was compared to a gold standard registration established using eight fiducial markers around the perimeter of the phantom. This was repeated 100 times from random starting points as discussed in Section 3. The resulting RMS TRE at the prostate was 0.49 mm.

5 Endoscope Tracking

5.1 Endoscope Tracking Error

The tracking accuracy for the endoscope tip was determined analytically using the method proposed by Fitzpatrick et al.[7], treating the Infra Red Emitting Diodes (IRED) localisation as the fiducial localisation error and the tip localisation error as the TRE. The individual IRED tracking error was found to be 0.02mm by tracking static IREDs over time. This matches the result reported by Barnes et al.[8]. The system is intended to be used while the endoscope is stationary, so it is valid to use the static error, rather than a dynamic measure. It was found that when surgical drapes were placed over the IREDs the error increased to 0.05mm. This figure was used together with the IRED geometry to solve equation 46 in Fitzpatrick et al. [7] for the endoscope tip error.

Due to technical and clinical limitations of where the IREDs can be attached to the daVinci and where the tracker can be positioned in the operating theatre, many of the fourteen IREDs are obscured at any given point, so the actual tip localisation error varies from frame to frame. The effect of this was assessed by performing a simulated prostatectomy on the phantom in theatre. Approximately 9000 frames of video were collected. Of these 1400 were discarded as their calculated tip error was greater than 2mm. The RMS error for the remaining frames was 0.80 mm.

6 Discussion

We have carried out an analysis of the component errors present in our image guidance system. The errors identified are summarised in Figure 1 and the system error is estimated to be 3.7mm. The largest contribution to this error is from segmentation of the pelvic bone. The accuracy of our shape-model segmentation method compares well with other published pelvic shape models, so to improve this segmentation will probably require extension beyond shape-model based segmentation.

Further research will be required to ascertain the size of error which can be tolerated in our system. Visual overlay systems can tolerate larger errors than blind navigation systems as the surgeon can mentally adjust for some error. We aim to carry out further phantom experiments to determine the directional distribution of the error which will determine how it appears on screen. We are also aiming to test the system live during surgery which will enable us to investigate the size of the overlay errors and to visualise a number of error sources (such as soft tissue deformation) which are not simulated in our phantom studies.

This is a first accuracy assessment for a complex image guided surgery system. Although a great deal of further analysis of the individual components and how they interact is required, we discuss the errors for a complete system from preoperative imaging to surgical video overlay.

References

1. Penney, G.P., Barratt, D.C., Chan, C.S., Slomczykowski, M., Carter, T.J., Edwards, P.J., Hawkes, D.J.: Cadaver validation of intensity-based ultrasound to CT registration. In: MICCAI. Volume 3750 of LNCS., Springer (2005) 1000–1007
2. Mourgues, F., Coste-Manière, E.: Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery. In: MICCAI. Volume 2488 of LNCS., Springer (2002) 25–34
3. Hoogeman, M.S., van Herk, M., de Bois, J., Lebesque, J.V.: Strategies to reduce the systematic error due to tumor and rectum motion in radiotherapy of prostate cancer. *Radiotherapy and Oncology* **74** (2004) 177–185
4. Schmid, J., Magnenat-Thalmann, N.: MRI bone segmentation using deformable models and shape priors. In: MICCAI. Volume 5241 of LNCS., Springer (2008) 119–126
5. Thompson, S., Penney, G., Buie, D., Dasgupta, P., Hawkes, D.: Use of a CT statistical deformation model for multi-modal pelvic bone segmentation. In: *Medical Imaging 2008: Image Processing*. Volume 6914 of Proceedings of the SPIE. (2008)
6. Crum, W.R., Tanner, C., Hawkes, D.J.: Anisotropic multi-scale fluid registration: evaluation in magnetic resonance breast imaging. *Physics in Medicine and Biology* **50** (2005) 5153–5174
7. Fitzpatrick, J.M., West, J.B., Maurer Jr, C.R.: Predicting error in rigid-body, point-based registration. *IEEE Transactions on Medical Imaging* **17**(5) (October 1998) 694–702
8. Barnes, P., Baldock, C., Meikle, S., Fulton, R.: Benchmarking of a motion sensing system for medical imaging and radiotherapy. In: *IEEE Nuclear Science Symposium Conference Record*. Volume 6. (2007) 4513–4520