Registration between ultrasound and fluoroscopy or CT in prostate brachytherapy

P. Fallavollita
School of Computing, Queen’s University, Ontario K7L 3N6, Canada

Z. Karim Aghaloo
Centre for Intelligent Machines, McGill University, Quebec H3A 2A7, Canada

E. C. Burdette
Acoustic MedSystems Inc., Champaign, Illinois 61822

D. Y. Song
Radiation Oncology & Molecular Radiation Sciences, Johns Hopkins Hospital, Baltimore, Maryland 21231

P. Abolmaesumib
Department of Electrical and Computer Engineering, University of British Columbia, Vancouver V6T 1Z4, Canada

G. Fichtinger
School of Computing, Queen’s University, Ontario K7L 3N6, Canada

(Received 17 November 2009; revised 16 March 2010; accepted for publication 7 April 2010; published 20 May 2010)

Purpose: In prostate brachytherapy, transrectal ultrasound (TRUS) is used to visualize the anatomy, while implanted seeds can be visualized by fluoroscopy. Intraoperative dosimetry optimization is possible using a combination of TRUS and fluoroscopy, but requires localization of the fluoroscopy-derived seed cloud, relative to the anatomy as seen on TRUS. The authors propose to develop a method of registration of TRUS images and the implants reconstructed from fluoroscopy.

Methods: A phantom was implanted with 48 seeds then imaged with TRUS and CT. Seeds were reconstructed from CT yielding a cloud of seeds. Fiducial-based ground-truth registration was established between the TRUS and CT. TRUS images are filtered, compounded, and registered to the reconstructed implants by using an intensity-based metric. The authors evaluated a volume-to-volume and point-to-volume registration scheme. In total, seven TRUS filtering techniques and three image similarity metrics were analyzed. The method was also tested on human subject data captured from a brachytherapy procedure.

Results: For volume-to-volume registration, noise reduction filter and normalized correlation metrics yielded the best result: An average of 0.54 ± 0.11 mm seed localization error relative to ground truth. For point-to-volume registration, noise reduction combined with beam profile filter and mean squares metrics yielded the best result: An average of 0.38 ± 0.19 mm seed localization error relative to the ground truth. In human patient data, C-arm fluoroscopy images showed 81 radioactive seeds implanted inside the prostate. A qualitative analysis showed clinically correct agreement between the seeds visible in TRUS and reconstructed from intraoperative fluoroscopy imaging. The measured registration error compared to the manually selected seed locations by the clinician was 2.86 ± 1.26 mm.

Conclusions: Fully automated registration between TRUS and the reconstructed seeds performed well in ground-truth phantom experiments and qualitative observation showed adequate performance on early clinical patient data. © 2010 American Association of Physicists in Medicine. [DOI: 10.1118/1.3416937]

Key words: prostate brachytherapy, registration, fluoroscopy, ultrasound

I. INTRODUCTION

Prostate cancer is a worldwide health problem and the most commonly diagnosed cancer among men in the US, where one of every six men receives a positive diagnosis in his lifetime. Low dose rate permanent prostate brachytherapy, or brachytherapy in this paper, entails permanent implantation of encapsulated radioactive sources (seeds) into the prostate in order to eradicate the cancer. The seeds are the size of a rice grain. Brachytherapy has emerged as a definitive treatment option for patients with early stage prostate cancer, a group that represents the majority of patients diagnosed in the modern era of prostate cancer screening. Brachytherapy has demonstrated the potential for excellent cancer control. The success of brachytherapy relies on the physician’s ability to tailor the spatial distribution of radiation dose with respect to the prostate and surrounding struc-
tures. A treatment plan consisting of precise seed distribution within the prostate is created prior to the procedure, but the actual implant shows deviations from the plan due to manifold reasons, including needle misplacement, edema, and migration of seeds during and after procedure. Suboptimal seed placement may cause an insufficient dose to the cancer and/or inadvertent radiation of the rectum, urethra, or bladder. The former may result in failure of treatment while the latter may cause adverse side effects like rectal ulceration, incontinence, and painful urination. According to a comprehensive review by the AAPM TG137 Report, interactive planning represents an improvement over intraoperative preplanning, however, in interactive planning the calculated dose distribution is based on the implanted needle position, and hence interactive planning might not account for seed movement after deposition. At the same time, continues the report, the dose distribution is updated dynamically based on the actual positions as the seeds are deposited. At this time, dynamic dose calculation is not available for permanent prostate brachytherapy because it is difficult to image individual seeds on TRUS.

The ability to perform dynamic dose calculation with visualization of underdosed areas during the procedure could significantly improve brachytherapy. The clinical usefulness of the dynamic planning approach has been demonstrated in limited clinical studies led by Zelefsky et al., Song et al., and most recently by Orio et al. However, no solution exists today that would allow for the deployment of dynamic implant dosimetry in average community care setting. In this paper, we investigate the last standing roadblock to dynamic brachytherapy implant dosimetry optimization: The registration between transrectal ultrasound and fluoroscopy.

I.A. Seed localization in ultrasound

Prostate brachytherapy is usually performed under transrectal ultrasound (TRUS) guidance (Fig. 1). While TRUS provides adequate imaging of the soft tissue anatomy, it does not allow for robust localization of the implanted brachytherapy seeds. It has been extremely difficult to localize seeds in TRUS, illustrated in Fig. 2. To date, no clinically robust and accurate method exists for this task. The difficulties have several root causes. Variable out-of-plane orientation of the seeds causes variability in acoustic appearance. Calcifications and other highly reflective objects present in the field of view and multiple reflections create false positive (FP) appearances. Seeds may cluster and/or shadow one another, creating false negative (FN) appearances and even when meticulously hand-segmented, up to 25% of true seeds may remain hidden in ultrasound. Recently, Orio et al. reported that in a study of 25 patients, only 20%-25% of all implanted seeds were visually identifiable in TRUS, and in some cases even less. As false positives and false negatives are concurrent, it is often impossible to differentiate acoustically similar true and false positives. Despite the unfavorable odds, significant efforts have been invested in direct localization of seeds from TRUS, owing to the particular clinical significance of the issue.

Semimanual segmentation on postimplant TRUS using prior knowledge of the position of the seeds from the preoperative model has been suggested by Ref. 12. The visible needle tracks are localized and matched with the preoperative implant model. The visible echoes are then identified as seeds or spacers from the known sequence of the preoperative model. Seeds that do not produce a visible echo (false...
negatives) are interpolated from the model. This approach requires a great deal of manual intervention and subjective judgment leading to unreliable performance for nonstranded “lose seed” implants. Ding et al.\textsuperscript{13,14} also used the needle location information when seeds were being deposited from the needle. Unfortunately, as the prostate often deforms or swells and seeds shift during the procedure, seeds positions may be altered. Thus, this approach is unsuitable for producing dosimetry which reflects the end result. Interestingly, segmentation of seeds in fluoroscopy has met similar obstacles. Prior works used variants of morphological shape-based approaches,\textsuperscript{4,15,16} but could not reliably differentiate true seeds appearances from partial and multiple appearances, especially when applied to images of lower resolution. Template matching and optimization based techniques have been suggested in the literature for resolving seed multiples\textsuperscript{17,18} but they tend to be computationally intensive and not quite feasible intraoperatively. In particular, the problem of image resolution seems to have been underappreciated.

In exploring more complex TRUS imaging schemes, McAlavey et al.\textsuperscript{19} suggested an oscillating external magnetic field to vibrate brachytherapy seeds that are magnetized or ferromagnetic. Doppler ultrasound can then detect the vibration of the brachytherapy seeds and differentiate them from other echo targets. The unique Doppler signature associated with the modified brachytherapy seeds in the oscillating field allows them to be readily distinguished from other bright scatterers in the prostate. Recently, Mitri et al.\textsuperscript{20} suggested vibroacoustography to vibrate and then detect the seeds, and demonstrated the concept under idealized circumstances in soft silicon gel phantoms. Transurethral ultrasound has been proposed\textsuperscript{21,22} to enhance the identification of both seeds and anatomy. This is an ambitious long-term approach that would impart a wholesale change in established clinical hardware and workflow.

I.B. Registration of ultrasound to fluoroscopy

The published history of C-arm fluoroscopy in brachytherapy originates\textsuperscript{23} when it was first used as a solo guidance modality. Shortly after TRUS emerged as a primary guidance modality, fluoroscopy became a secondary tool for gross visual observation. Mobile C-arms are ubiquitious in contemporary prostate brachytherapy, with over 60% of the practitioners using it for qualitative implant analysis during the procedure\textsuperscript{24} in a setup shown in Fig. 1. Recently, accurate reconstruction of seeds from fluoroscopy has become possible,\textsuperscript{25–27} but since it cannot show the prostate and soft tissues (Fig. 2), fluoroscopy alone cannot be used for dosimetry. At a few specialized centers, CT imaging is available during brachytherapy for implant reconstruction.\textsuperscript{28} In all, as seeds can be reconstructed from fluoroscopy or CT, the remaining problem is registration of TRUS with the reconstructed seeds.

To this end, Zhang et al.\textsuperscript{29} suggested affixing x-ray fiducials onto the TRUS probe. French et al.\textsuperscript{30} suggested using the TRUS probe as a registration fiducial. Jain et al.\textsuperscript{25} proposed mounting a fiducial structure above the abdomen, spatially calibrated with the template. In all the above, the registration took place through static fiducials and cannot account for motion of the anatomy between the TRUS and fluoroscopy sessions. When the TRUS probe is retracted from the rectum for fluoroscopy imaging, the prostate relaxes posteriorly. To counteract this problem, Su et al.\textsuperscript{31} suggested using a point-based registration between seeds directly reconstructed between TRUS and fluoroscopy.\textsuperscript{26,31} This requires exact localization of the seeds in TRUS which, as mentioned earlier, has not been possible with clinically sufficient accuracy and robustness.

Generally, intensity-based registration tends not to work between ultrasound and x-ray modalities because in ultrasound, anatomical structures are embedded in noisy and low contrast environment with little distinctive information about material density measured by x ray. But implanting the prostate with seeds changes this situation advantageously. While TRUS images of an implanted prostate are hampered by noise and false positive and negative appearances, seeds carry enough distinctive information for intensity-based registration to hone in on the correct transformation. Still, the apparent straightforwardness of our method should not belie the investment of effort needed to make a workable clinical tool, despite the availability of technical components. In this paper, we report intensity-based registration between TRUS and seeds reconstructed from intraoperative fluoroscopy and CT. We present a practical implementation, experimental analysis on ground-truth phantom, and initial results on patient data.

II. MATERIALS AND METHODS

II.A. General overview

The registration scheme (Fig. 3) follows several steps. (1) Reconstruct the implant from fluoroscopy or CT and obtain a three-dimensional cloud of the seeds. The reconstructed seeds can be represented as a “3D point set” or as a “binary
Central to our approach is filtering the TRUS images, without explicit segmentation of the seeds. We allow false positives (artifacts masquerading as seeds) and false negatives (shadowed seeds) to remain in the data sets during registration. A variety of filters were tested to enhance the quality TRUS images for registration, with the purpose of reducing noise and enhancing the features of true seeds. Again, we stress that we did not attempt explicit segmentation of the seeds and we allowed false positive appearances and other artifacts to remain in the TRUS. We explain the use of three primary filters (noise reduction, phase congruency, and beam width filters) and different combinations of these, also listed in Fig. 4. No filtering (U.S.0) was used as baseline for comparison.

II.B. Ultrasound and CT filtering

Central to our approach is filtering the TRUS images, without explicit segmentation of the seeds. We allow false positives (artifacts masquerading as seeds) and false negatives (shadowed seeds) to remain in the data sets during registration. A variety of filters were tested to enhance the quality TRUS images for registration, with the purpose of reducing noise and enhancing the features of true seeds. Again, we stress that we did not attempt explicit segmentation of the seeds and we allowed false positive appearances and other artifacts to remain in the TRUS.

II.B.1. Noise reduction filter (U.S.1)

Based on the intuition that brighter areas in the image are more likely to contain seeds, successive thresholding is applied on the U.S. images. First, an average intensity is calculated for pixels within the TRUS image. This is called \( \overline{I}_1 \). Then, pixels having value less than \( \overline{I}_1 \) are colored black. The average intensity of the remaining pixels in \( I_1 \) is calculated afterwards and the procedure is repeated. Equations (1) and (2) show the concept of successive thresholding where \( n \) is the number of pixels within the image and \( n_1 \) is the number of remaining pixels after the first thresholding.

\[
\overline{I} = \frac{\sum_{i,j} I(i,j)}{n}, \quad I(i,j) = \begin{cases} 
\text{TRUS}(i,j) & \overline{I} < \text{TRUS}(i,j) \\
0 & \text{otherwise}
\end{cases}, 
\]

\[
\overline{I}_2 = \frac{\sum_{i,j} I_1(i,j)}{n_1}, \quad I_2(i,j) = \begin{cases} 
I_1(i,j) & \overline{I}_2 < I_1(i,j) \\
0 & \text{otherwise}
\end{cases},
\]

\[
\overline{I}_3 = \frac{\sum_{i,j} I_2(i,j)}{n_2}.
\]

After the second thresholding, again the average of the remaining pixels is calculated (\( \overline{I}_3 \)) by defining \( n_2 \) as the number of remaining pixels in \( I_2 \). Remaining pixels are divided to separate regions and average intensity value for each region is calculated. Regions with average intensity less than 0.5\( \overline{I}_3 \), or less than 0.75\( \overline{I}_3 \), and an area less than 100 pixels are removed. The two threshold values for \( \overline{I}_3 \) were chosen so as to eliminate regions with high intensity and area smaller than the acceptable size for a brachytherapy seed.

II.B.2. Phase congruency filter (U.S.2)

Previously, Hacihaliloglu et al. have shown that phase congruency is an effective tool for detecting the true bone edge location in U.S. images. Based on this idea, we adapted phase congruency for processing the U.S. images to enhance the features of true seeds, i.e., to suppress artifacts and false positive appearances. The typical approach for edge detection is to think of edges as being points of high intensity gradient. An alternative approach is to think of features in the frequency domain. The image intensity profile can be thought of as a periodic signal and hence can be presented with its Fourier series. The phase congruency theory assumes that features exist at points in an image where the Fourier components are in phase.

Kovesi proposed a method for measuring how in phase Fourier components are in any point of the signal. The concept can be easily extended to two-dimensional signals by...
applying the one-dimensional analysis over several orientations and combining the results. Along each orientation $\sigma$, a spread filter is defined determining the area covered by that orientation. Also, by introducing the scale parameter $n$, Fourier components can be analyzed independently and then be combined together in the final calculation of the phase congruency.

Hence we applied a uniform threshold for all images. For extracting the seedlike regions from one B-mode image, calculating the phase congruency of pixels provides useful information: The more symmetrical the phase of a region is, the more likely it is a seed. The number of orientations and scales were found empirically (Fig. 5, left). We used six orientations and four scales. For the phase congruency filter implementation, we used the MATLAB algorithm provided by Kovesi. 

### II.B.3. Beam width filter (U.S.3)

The beam width filter accounts for the fact that the ultrasound beam has finite thickness and single focus in the elevation plane. The thickness of the beam varies along the lateral plane, and based on the setting one is using, different number of focal points along the lateral plane can be achieved (i.e., two for our experiment).

For finding the lateral beam profile, we take U.S. images of a nylon wire in different depths. The resulting image of the wire when the probe is held perpendicular to it is a spot. U.S. images of the wire are captured at different depths and in each depth, the width of the spot is considered as the width of the U.S. beam in the lateral direction. Wherever the lateral profile has narrower width (meaning the accuracy is higher), gets more weight. Another experiment is performed for obtaining the elevation profile of the U.S. beam. In this experiment, an inclined plane which is made from rubber is used for evaluating the elevation beam width. If the U.S. probe is held 45° relative to this plane, the image of the plane would be a horizontal line, whose thickness is equivalent to the elevation beam width. Therefore, images of the inclined plane are captured in different depths to find the elevation beam profile. In each depth, the thickness of the line is reported as the elevation beam width for that depth. The exact protocol can be found in Ref. 35.

The two lateral and elevation filters previously obtained can be considered each as a probability map, and since they are independent, the overall probability is obtained by multiplying them. This filter is applied on the TRUS images in order to compensate for the uncertainties caused by the U.S. beam width either in lateral or elevation directions. Pixels located near the focal point receive more weight and thus have more effect on the final result (Fig. 5, right).

### II.B.4. Combination filters (U.S.4–U.S.7)

U.S.4 implements a Bayesian combination of noise reduction, phase congruency, and beam width filter is an image with intensity between 0 and 255. Since the brighter a pixel is, the higher is the probability of that pixel being a seed, results can be rescaled to [0,1] and hence, represent a probability map. Therefore, the image intensities were normalized and then Bayesian formula [Eq. (4)] was used for each pixel. Basically, in this formula, each probability map independently estimates the “probability seedness” of pixels. We applied the Bayesian model for combining different probabilities according to Scepanovic et al.

\[
P(x,y) = \frac{P_1(x,y) \times P_2(x,y)}{P_1(x,y) \times P_2(x,y) + (1 - P_1(x,y))(1 - P_2(x,y))}.
\]

(4)

U.S.5 is a serial filter of noise reduction followed by phase congruency. U.S.6 is another serial filter made up by noise reduction followed by beam width filtering. Finally, U.S.7 is a series of three elementary filters, namely, noise reduction, phase congruency, and beam width filters.

### II.B.5. CT filtering

As seeds are prominent in CT, we only clipped a region of interest and used window-level scaling to create an 8-bit CT volume. Window-level scaling is a method for mapping a range of intensity to a different scale. A window is set by a lower and an upper threshold. Below the lower threshold, all intensities are mapped to black; above the upper threshold all intensities are mapped to white, and intensities inside the window are mapped linearly to 8-bit gray scale. Generally, CT data has intensity values between 2000 and 4095. In our case, true seeds (intensities above 1700) were mapped to the 8-bit gray scale. Then, a suitable threshold of 100 was applied to suppress all remaining artifacts.

### II.C. Intensity-based registration of volume-to-volume (V2V)

Intensity-based registration techniques do not require prior segmentation, a distinctive feature offering manifold advantages over segmentation-based registration that usually requires some level of manual intervention and is hampered by segmentation errors. Intensity-based registration uses the entire volumes and thus both false positive and false negative...
appearances can average out. Generally speaking, V2V is designed to handle scenarios where seeds are reconstructed into a continuous gray-scale volume, such as true CT (Ref. 28) or cone-beam computed tomography volume. V2V should also allow for a gray-scale volume to be binarized, where usually each seed is represented by more than one white voxels. Such volumes are resulted from fluoroscopy reconstructions where seeds are considered to have length and direction, such as in Ref. 38. Since voxel intensity is used, an interpolator is used to evaluate moving volume intensities at nonpixel positions after applying the spatial transformation in each cycle of iteration. For this very reason, we must consider the continuous TRUS as the moving volume and binarized seed cloud as fixed volume. For the first registration, the transformation parameters are initialized with an initial guess, discussed later. Then, a similarity measure evaluates the degree of matching between the transformed moving volume and the fixed volume. The two most widely used similarity measures are mutual information and normalized correlation. Our experimental registration scheme is based on the ITK-Insight Segmentation and Registration Toolkit.

II.C.1. Metric

We implemented Mattes mutual information (MMI) that estimates a probability density function (PDF) uniformly distributed over the intensity interval. The calculations are based on the method of Mattes et al., where the probability density distributions are estimated using Parzen histograms. Compared to standard mutual information, we initialize a random number generator that selects the sample of voxels used for estimating the volume histograms and the joint histogram. Further, the advantage of MMI is that it uses only two parameters: The number of samples used to estimate the PDF and the number of histogram bins used to calculate the entropy. The number of histogram bins was empirically set to 50 and the number of samples used was set to 10% of the total number of voxels making up the fixed volume. For comparative purposes, we also tested NC. Voxel values are taken from the fixed volume; their positions are mapped to the moving volume and result in general in nonpixel position on it. Values at these nongrid positions of the moving volume are interpolated using a user-selected interpolator. The correlation is normalized by the autocorrelations that samples the select of voxels used for estimating the volume histograms and the joint histogram. Further, the advantage of MMI is that it uses only two parameters: The number of samples used to estimate the PDF and the number of histograms used to calculate the entropy. The number of histograms was empirically set to 50. Let \( \text{Img1} \) and \( \text{Img2} \) be the fixed and moving volume, respectively. NC computes the voxelwise cross-correlation and normalizes it by the square root of the autocorrelation. Using a simple 2D image example,

\[
\text{NC} = -1 \times \frac{\sum_{i=1}^{N} \text{Img1}_i \cdot \text{Img2}_i}{\sqrt{\sum_{i=1}^{N} \text{Img1}_i^2 \cdot \sum_{i=1}^{N} \text{Img2}_i^2}},
\]

where \( \text{Img1}_i \) and \( \text{Img2}_i \) are the \( i \)th pixels in the two images and \( N \) is the number of pixels considered. The \(-1\) factor is used to make the metric optimal when its minimum is reached. The optimal value of the metric is then \(-1\). A misalignment between the images results in small measurement values. This metric produces a cost function with sharp peaks and well-defined minima. Naturally, the above metrics work identically for volumes. The number of spatial samples used here was empirically set to 50.

II.C.2. Transformation

Several authors have investigated nonrigid deformation techniques using prostate data. For instance, Karnik et al. conclude that the results from rigid and nonrigid registrations were not statistically significantly different \( p > 0.05 \) in their transrectal prostate biopsies. Xu et al. conclude that there was no apparent impact of deformable registration which indicates that prostate deformation is less of a factor than organ displacement during the needle placement process. Lastly, Misra et al. show that boundary conditions surrounding the organ dominate the deformation more than the constitutive behavior of the tissue itself. Interestingly, after TRUS imaging, the probe is retracted from the rectum to avoid blocking seeds during fluoroscopy imaging. This causes the prostate to relax posteriorly, but usually without significant deformation.

Practice guidelines specifically require slight rectal pressure to mitigate deformation by probe translation during TRUS imaging. In this light, rigid registration should suffice. We implemented a 3D Euler transform of six parameters, three for the Euler angles that represent the rotation, and three for the translational components. We specify the center of rotation as the gravity center of the moving volume. Clinically, we always have accurate and consistent initial guess for the registration. Standard patient positioning allows for estimating the main symmetry axes of the prostate, and alignment of the gravity centers of the TRUS volume and the seed cloud yields accurate guess for translation.

II.C.3. Optimizer

One of the simplest and yet most powerful evolution strategies is the “one plus one evolution strategy,” denoted by \((1+1)\)-ES. The step size adaptation can be performed according to the following rule: If less than 20% of the generations are successful, then decrease the step size for the next generation; if more than 20% are successful, and then increase the step size in order to accelerate convergence. This adaptation is done for every \( N - LR \) generations, where \( N \) is the number of parameters to be optimized and \( LR \) is a constant which here is equal to 1. All parameter values were set default with the maximum number of iterations set to 100 and the minimum value for the Frobenius norm of the covariance matrix set to 0.0015. If the norm is smaller than this value, the optimization process will stop even before it hits the maximum iteration. The registration parameters were assigned initial weights corresponding to a 1° to 1 mm ratio for rotations and translations, respectively.

Lastly, as TRUS and fluoroscopy are both likely to be spotted with false positives, the registration could be trapped in local minima. To counteract this problem, for all experiments, we restart the registration 15 times with slightly changing the initial pose (i.e., ±10% of last optimized pa-
rameter value) and then take as a final solution the smallest root mean square results of the 15 convergence optimizations.

II.D. Intensity-based registration of points-to-volume (P2V)

An alternative to volume-to-volume registration is an algorithm that maps a discrete 3D point set to the moving TRUS volume. Once again, the goal of the registration method is to find the set of parameters of the transformation that optimizes a specified metric. Similarly to V2V earlier, we consider the 3D point cloud as fixed volume and the TRUS as moving volume and we search for a transformation that maps the moving volume to the fixed one with maximum similarity. To each point in the cloud, we assign a voxel value of 255 and thus create a black volume sparsely spotted with single white voxels. Each seed is represented by exactly one white voxel. Two metrics were tested for P2V: Normalized correlation and MS metric, the latter computing the square difference between the average moving volume intensity for an inside point and the average moving volume intensity for an outside point. The optimal value of the metric is zero. Poor matches between the images result in large values of the metric. This metric is simple to compute and has a relatively large capture radius. The number of pixels considered, $N$, was set to the image size.

$$MS = \frac{1}{N} \sum_{i=1}^{N} (\text{Img}_1 - \text{Img}_2)^2.$$  \hspace{1cm} (6)

An inside point is defined as points for which the corresponding value (in the point set node) is negative. An outside point is defined as points for which the corresponding value (in the point set node) is positive. The transform and optimizer are kept the same as described in the preceding section. Generally, P2V is designed for schemes where each seeds is reconstructed into a single point, such as in Refs. 25 and 26, and seeds are represented as a list of points.

III. EXPERIMENTS AND RESULTS

III.A. Ground-truth phantom

The objective here is to construct ground-truth seed locations relative to TRUS. A commercial prostate brachytherapy phantom (CIRS Inc., VA) was implanted with 48 nonradioactive seeds according to a clinically realistic implant plan (Fig. 6).

We selected CT imaging for ground-truth measurement for several reasons: High geometrical accuracy, availability and ease of use. Reconstructed seeds are faithfully represented by thresholded binarized CT volume, regardless to which x-ray modality they originate from, let that be CT, limited view C-arm fluoroscopy or cone-beam computed tomography. Six small BB fiducials were mounted on the phantom box, three on the left, and three on the right side. The fiducials were placed so that the prostate lay in the center of gravity of the fiducials, guaranteeing maximum target registration accuracy inside the prostate. We acquired CT images of the phantom with $0.3 \times 0.3 \text{ mm}^2$ in-plane resolution and 0.6 mm slice thickness.

We carefully segmented the six fiducials attached to the walls of the phantom. The CT fiducials were also localized with a calibrated pointer (Traxtal Inc., Toronto, Ontario) and Certus optical tracker (NDI, Waterloo, Ontario) with respect to the DRB coordinate on the phantom. Finally, we transformed the positions of seeds segmented in CT images to the TRUS coordinate system, thereby defining the ground truth for the registration. In order to maximize registration and tracking accuracy, the fiducials and tracking bodies were arranged so that their centroids fell close to the center of the prostate. This process yielded ground-truth registration between the TRUS and CT and the target registration error (TRE) was 0.2 mm.

The TRUS probe was tracked optically with Certus relative to the DRB. The TRUS image space and DRB space were calibrated with the system reported in Ref. 35. Using tracked TRUS acquisition, we scanned the phantom systematically with translational motion and continuous image capture. The in-plane pixel size was $0.14 \times 0.13 \text{ mm}^2$. Interframe spacing was about 0.5 mm. Following filtering, an 8-bit image volume was compounded from the 2D slices. The TRUS reconstruction relies on a splatting technique for high-quality interpolation, where each pixel of a B-mode image is smeared into a $N \times M \times O$ kernel, which is then either compounded or “alpha-blended” into the 3D image reconstruction volume at the appropriate $(x, y, z)$ location. The 2D B-mode images are splatted one by one as they are captured to provide real-time reconstruction. Compounding requires the use of an accumulation buffer which is the same size (i.e., same number of voxels) as the reconstruction volume. For any U.S. image, a tracking file is stored con-
Translation and registration error and standard deviations were calculated as the difference between true and estimated seed positions, obtained from the ground truth and the registration. We applied random misalignment of maximum 15° rotation. This capture range, especially from the optical tracker to the phantom, was larger than we need in actual clinical practice. Again, we note that the angular capture range tested was larger than we need in actual clinical practice.) All filters with both MMI and NC metrics gave a mean seed registration error less than 1 mm, which is well below the clinically acceptable limit. The NC metric outperformed the MMI metric for all filters, likely because NC uses the entire image to perform registration whereas the MMI randomly selects samples during the optimization phase. Noise reduction combined with phase congruency filter (U.S.5) gave the best TRE (0.75 mm) for MMI metric. Noise reduction filter alone (U.S.1) gave the best TRE (0.62 mm) with NC metric. Somewhat surprisingly perhaps, no filtering (U.S.0) yielded excellent TRE with both metrics, 0.65 mm with MMI, and 0.77 mm with NC. Figure 7 shows images after registration of TRUS and binary CT.

The effects of false positives and false negatives in TRUS were investigated by randomly masking true seeds and adding false seeds to the 3D volume created from the ground-truth CT data. To start with, there were many more false positives than false negatives in the phantom (Fig. 6). To simulate false positive and false negative appearances in TRUS, we added and subtracted, respectively, up to 15 seeds.

### III.A.1. Registration evaluation

Using the gold standard transformation, the two data sets are first aligned. To test the two registration algorithms, the volume selected as the moving image (i.e., TRUS data) is perturbed from the aligned position by applying a random transformation within the range selected for the perturbation. The registration then tries to bring the perturbed volume back to the aligned position. The registration error is evaluated in terms of TRE, which is defined as the distance between corresponding points other than the fiducial points after registration. In this paper, the values for mean registration error and standard deviations were calculated as the difference between true and estimated seed positions, obtained from the ground truth and the registration, respectively.

Capture range is defined as the range within which the algorithm is more likely to converge to the correct optimum. We applied random misalignment of maximum ±5 mm translation and ±15° rotation. This capture range, especially for rotation, is larger than the error of the initial guess one can achieve clinically. The robustness of registration is defined here as the ability of the optimization to get close to the global minimum on all trials. Robust performance in the presence of many false positive seed appearances will be of key importance in clinical use. For good measure, we have tested for both false positive and false negative during experimentation. In conclusion, we performed the registration 25 times for each TRUS filter.

### III.B. V2V registration results

Table I shows results from the runs that yielded TRE below the clinically acceptable threshold of 2 mm. (The diameter of an implant needle and a seed is about 2 and 1 mm, respectively. Seed localization error below 2 mm is considered as clinically adequate.)

While the algorithm converged in each run, 27% yielded TRE values above 2 mm. As customary in such cases, we adjusted the initial guess and restarted the registration. (Again, we note that the angular capture range tested was larger than we need in actual clinical practice.) All filters with both MMI and NC metrics gave a mean seed registration error less than 1 mm, which is well below the clinically acceptable limit. The NC metric outperformed the MMI metric for all filters, likely because NC uses the entire image (all samples) to perform registration whereas the MMI randomly selects samples during the optimization phase. Noise reduction combined with phase congruency filter (U.S.5) gave the best TRE (0.75 mm) for MMI metric. Noise reduction filter alone (U.S.1) gave the best TRE (0.62 mm) with NC metric. Somewhat surprisingly perhaps, no filtering (U.S.0) yielded excellent TRE with both metrics, 0.65 mm with MMI, and 0.77 mm with NC. Figure 7 shows images after registration of TRUS and binary CT.

The effects of false positives and false negatives in TRUS were investigated by randomly masking true seeds and adding false seeds to the 3D volume created from the ground-truth CT data. To start with, there were many more false positives than false negatives in the phantom (Fig. 6). To simulate false positive and false negative appearances in TRUS, we added and subtracted, respectively, up to 15 seeds.

### Table I. Mean and standard deviation of TRE in phantom trials for V2V registration (all units in mm).

<table>
<thead>
<tr>
<th>Filter</th>
<th>NC</th>
<th>MMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.0</td>
<td>0.61 ± 0.17</td>
<td>0.71 ± 0.15</td>
</tr>
<tr>
<td>U.S.1</td>
<td>0.54 ± 0.11</td>
<td>0.82 ± 0.06</td>
</tr>
<tr>
<td>U.S.2</td>
<td>0.74 ± 0.26</td>
<td>0.87 ± 0.12</td>
</tr>
<tr>
<td>U.S.3</td>
<td>0.79 ± 0.12</td>
<td>0.88 ± 0.13</td>
</tr>
<tr>
<td>U.S.4</td>
<td>0.77 ± 0.25</td>
<td>0.91 ± 0.17</td>
</tr>
<tr>
<td>U.S.5</td>
<td>0.65 ± 0.15</td>
<td>0.69 ± 0.14</td>
</tr>
<tr>
<td>U.S.6</td>
<td>0.71 ± 0.10</td>
<td>0.91 ± 0.18</td>
</tr>
<tr>
<td>U.S.7</td>
<td>0.69 ± 0.12</td>
<td>1.01 ± 0.15</td>
</tr>
</tbody>
</table>

### Table II. FP and FN evaluation by randomly adding/masking out 15 seeds for the V2V registration algorithm; mean and standard deviation of TRE in phantom trials for the least performing filter (all units in mm).

<table>
<thead>
<tr>
<th>Filter</th>
<th>NC</th>
<th>MMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP</td>
<td>U.S.7</td>
<td>U.S.3</td>
</tr>
<tr>
<td></td>
<td>1.12 ± 0.34</td>
<td>0.91 ± 0.18</td>
</tr>
<tr>
<td>FN</td>
<td>U.S.4</td>
<td>U.S.3</td>
</tr>
<tr>
<td></td>
<td>0.96 ± 0.40</td>
<td>1.13 ± 0.32</td>
</tr>
</tbody>
</table>
tering the TRUS data was beneficial, as six of the seven Medical Physics, Vol. 37, No. 6, June 2010

computer. average of 150 s on an Intel Core2, 2.4 GHz dual-core computer.

trials for the least performing filter for the P2V algorithm; mean and standard deviation of TRE in phantom trials for the least performing filter (all units in mm).

<table>
<thead>
<tr>
<th>Filter</th>
<th>NC</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.0</td>
<td>1.19 ± 1.99</td>
<td>0.40 ± 0.17</td>
</tr>
<tr>
<td>U.S.1</td>
<td>0.58 ± 1.44</td>
<td>0.70 ± 0.22</td>
</tr>
<tr>
<td>U.S.2</td>
<td>0.88 ± 1.68</td>
<td>0.49 ± 0.12</td>
</tr>
<tr>
<td>U.S.3</td>
<td>1.89 ± 2.23</td>
<td>0.56 ± 0.12</td>
</tr>
<tr>
<td>U.S.4</td>
<td>0.67 ± 1.67</td>
<td>0.38 ± 0.19</td>
</tr>
<tr>
<td>U.S.5</td>
<td>0.66 ± 1.75</td>
<td>0.65 ± 0.24</td>
</tr>
<tr>
<td>U.S.6</td>
<td>0.54 ± 1.36</td>
<td>0.88 ± 0.17</td>
</tr>
<tr>
<td>U.S.7</td>
<td>0.76 ± 1.51</td>
<td>0.83 ± 0.21</td>
</tr>
</tbody>
</table>

(31%) present in the CT data. This false positive and negative rate is significantly higher than observed in human patient images.

In the false positive analysis, all simulations yielded acceptable TRE values that were less than the 2 mm clinical threshold. Table II shows the least accurate results occurred when removing 15 seeds from the CT data; the U.S.7 filter using the MMI metric (TRE of 1.26 ± 0.40 mm) and the U.S.3 filter using NC metric (TRE of 0.99 ± 0.18 mm). For comparison, the baseline nonfiltering scheme (U.S.0 filter) yielded the following results: TRE of 0.97 ± 0.41 mm using MMI and TRE of 0.88 ± 0.31 mm using NC metric.

In the false negative analysis, the least accurate results were observed using the U.S.4 filter in the MMI metric (TRE of 0.996 ± 0.40 mm) and the U.S.3 filter using NC metric (TRE of 1.13 ± 0.32 mm). For comparison, the baseline nonfiltering scheme (U.S.0 filter) yielded the following results: TRE of 0.92 ± 0.34 mm using MMI and TRE of 1.02 ± 0.18 mm using NC metric.

For the unoptimized MATLAB prototype, V2V took an average of 150 s on an Intel Core2, 2.4 GHz dual-core computer.

III.C. P2V registration results

For the point-to-volume registration technique, results for the NC metric were less accurate when compared to the above volume-to-volume analysis. As seen in Table III, filtering the TRUS data was beneficial, as six of the seven filters yielded more precise TRE values than the unfiltered TRUS. In particular, U.S.6 filter showed the best value of 0.67 ± 1.67 mm. The MS metric showed stronger statistical significance for average and standard deviation. Still, all filters yielded a final TRE below 1.0 mm.

In the false positive and false negative analyses, both the MS metric yielded excellent TRE below 2 mm. Table IV shows the least accurate results were achieved with the U.S.3 filter using the NC metric and U.S.5 filter using TRE of 0.87 ± 0.11 mm and TRE of 0.81 ± 0.14 mm for the false positive and false negative analysis, respectively. With the NC metric, the results deteriorated as the least precise values were achieved with the U.S.3 and U.S.2 filters yielding TRE of 1.97 ± 1.34 mm and TRE of 1.18 ± 1.12 mm for the false positive and false negative analysis, respectively.

The unoptimized MATLAB code for P2V executed in 90 s, in average, on an Intel Core2, 2.4 GHz dual-core computer.

III.D. Initial clinical patient data results

Clinical patient data was collected under ethics board approval. Here we present results on the first, and so far our only patient data set available. Seeds were reconstructed from four C-arm fluoroscopy images as a cloud of seeds using the method of Jain et al.,9 compounded into a binary CT volume. We then performed P2V registration with both NC and MS metrics, with results shown in Fig. 8 for the MS metric. A quick visual inspection reveals that there is no fluoroscopy seed without a white blotch in TRUS, suggesting that we do not have false negatives. As it was expected, there are some white blotches in TRUS without corresponding seed from fluoroscopy, suggesting the presence of multiple false positives. At the same time, the expert clinician had great difficulties in segmenting the dense series of TRUS slices with many seeds extending across consecutive images. For this particular patient, C-arm fluoroscopy reconstructed all the 81 implanted seeds. In the TRUS data, the expert clinician identified only 41 seeds with confidence, while 40 seeds could not be distinguished from noise and artifacts. Based on the 41 seeds positively segmented in TRUS, the average TRE was 2.86 mm, with a standard deviation of 1.26 mm.

IV. DISCUSSION AND CONCLUSION

To our knowledge, there has been no study to quantify the effects of TRUS and fluoroscopy registration error on dosi-
metric quality, i.e., to analyze the effect of each translation and rotation on dose volume histograms. Until such results become available in the literature, we use the accuracy threshold applied for the seed localization error, which is about 2.0 mm.

Intensity-based registration between TRUS and CT/fluoroscopy reconstructions of prostate implants was found to be excellent in phantom studies, in terms of TRE, capture range, and robustness to both false positive and negative seed appearances in TRUS.

As TRE and robustness seems statistically similar for the MMI and NC metrics in the V2V scheme, we favor NC for simplicity and speed. Notably, the MS metric on extremely sparse seed volumes in the P2V tests outperformed the NC and MMI metrics on richer volumes in the V2V scheme. This clearly suggests that recovering the seed centroids is sufficient for registration and there is no need for a costly and complicated recovery of seed length and direction, unless that is deemed useful for dosimetry purposes.

Surprisingly perhaps, filtering TRUS did not yield tangible advantages. According to Table I and the right column (MS metric) of Table III, unfiltered TRUS (baseline U.S.0) outperformed all but one filters. In any case, the differences were both statistically and clinically insignificant. Compliant filters, including the dummy U.S.0, of which was achieved, capture range was larger than required in standard clinical workflow, robustness was excellent to false positive and false negative seed appearances, and temporal performance was adequate. Early experience with limited clinical data was promising. A robust clinical validation methodology, such as implanted dual-modality fiducials, is necessary for prospective evaluation in patients.

ACKNOWLEDGMENTS

The authors are grateful to Siddharth Vikal, MSc, for his valuable advice and assistance throughout this project and to Xiao Xiao Ma, BSc, for medical illustration and Dr. David Gobbi for the ground-truth phantom protocol and experimentation. Dr. Pascal Fallavollita was supported by a postdoctoral fellowship of the Ontario Ministry of Research. This research was also funded by the NSERC Idea to Innovation PJ 364006–07 and by Queen’s University.

1A. Karim-Aghaloo and P. Abolmaesumi were with Queen’s University during this research.
6Z. Karim-Aghaloo and P. Abolmaesumi were with Queen’s University during this research.
7Electronic mail: pascal@cs.queensu.ca


