INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **47** (2002) 823–838 PII: S0031-9155(02)29686-4

Automatic MR volume registration and its evaluation for the pelvis and prostate

Baowei Fei[1], **Andrew Wheaton**[1], **Zhenghong Lee**[1,2], **Jeffrey L Duerk**[1,2] **and David L Wilson**[1,2]

[1] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106, USA
[2] Department of Radiology, University Hospitals of Cleveland and Case Western Reserve University, Cleveland, OH 44106, USA

E-mail: DLW@po.cwru.edu

Received 16 October 2001
Published 15 February 2002
Online at stacks.iop.org/PMB/47/823

Abstract

A three-dimensional (3D) mutual information registration method was created and used to register MRI volumes of the pelvis and prostate. It had special features to improve robustness. First, it used a multi-resolution approach and performed registration from low to high resolution. Second, it used two similarity measures, correlation coefficient at lower resolutions and mutual information at full resolution, because of their particular advantages. Third, we created a method to avoid local minima by restarting the registration with randomly perturbed parameters. The criterion for restarting was a correlation coefficient below an empirically determined threshold. Experiments determined the accuracy of registration under conditions found in potential applications in prostate cancer diagnosis, staging, treatment and interventional MRI (iMRI) guided therapies. Images were acquired in the diagnostic (supine) and treatment position (supine with legs raised). Images were also acquired as a function of bladder filling and the time interval between imaging sessions. Overall studies on three patients and three healthy volunteers, when both volumes in a pair were obtained in the diagnostic position under comparable conditions, bony landmarks and prostate 3D centroids were aligned within 1.6 ± 0.2 mm and 1.4 ± 0.2 mm, respectively, values only slightly larger than a voxel. Analysis suggests that actual errors are smaller because of the uncertainty in landmark localization and prostate segmentation. Between the diagnostic and treatment positions, bony landmarks continued to register well, but prostate centroids moved towards the posterior 2.8–3.4 mm. Manual cropping to remove voxels in the legs was necessary to register these images. In conclusion, automatic, rigid body registration is probably sufficiently accurate for many applications in prostate cancer. For potential iMRI-guided treatments, the small prostate displacement between the diagnostic and treatment positions

can probably be avoided by acquiring volumes in similar positions and by reducing bladder and rectal volumes.

1. Introduction

We are investigating three-dimensional (3D) image registration to be used in applications of prostate cancer diagnosis, staging and therapy. In particular, we are interested in applications related to the minimally invasive interventional MRI (iMRI) guided treatment of prostate cancer. Our group currently uses iMRI on a low-field open magnet system to guide radiofrequency (RF) thermal ablation of abdominal cancer (Lewin *et al* 1998), and we are investigating this method for prostate cancer treatment. A unique feature of iMRI-guided thermal ablation is that therapy can be monitored either by acquiring images of the thermally induced lesion or by measuring temperature. In addition, MR imaging of the prostate is desirable because it more accurately delineates the prostate than does CT (Milosevic *et al* 1998), which can overestimate the prostate volume (Roach *et al* 1996), and ultrasound, which has a tendency to underestimate the extent of lesions (Boni *et al* 1995).

Several important applications require registration of images of the prostate. First, comparison of registered MR images acquired before and immediately after RF ablation can be used to determine whether a tumour is adequately treated. This is particularly helpful in instances where the edematous response to treatment can be confused with a highly perfused tumour. Second, registration of serial examinations can be used to follow regression/progression of tumour. Third, registration of functional, biochemical images such as single photon emission computed tomography (SPECT), positron emission tomography (PET) and MR spectroscopy with anatomical MR images is useful for detecting and localizing cancer. Fourth, incorporating the functional, biochemical images into the iMRI paradigm will aid image-guided treatments. Fifth, on a low-field magnetic system during iMRI treatments where fast imaging is important, it might be highly desirable to register high-quality MR image from a conventional MR scanner to the live-time iMRI images (Fei *et al* 2001). In this study, we investigate registration of high-resolution MR volumes. Multi-modality image registration results were reported elsewhere (Fei *et al* 2001, Lee *et al* 2000, 2001a).

Many reports describe methods and evaluations for registration in the brain (Hill *et al* 2001); far fewer describe results for the pelvis or prostate. For example, manual registration has been used where an operator cues on segmented vascular structures (Hamilton *et al* 1999) or other anatomical landmarks (Balter *et al* 1995, Liehn *et al* 1992, Narayana *et al* 1997). Others have used automated 3D schemes that match contours of bones and sometimes other structures that are extracted using manual or interactive segmentation (Antolak *et al* 1998, Herk *et al* 1998, Remeijer *et al* 2000). Manual segmentation has also been used to create surfaces for automatic registration (Roeske *et al* 1995, Scott *et al* 1994). All of these methods require either segmentation or visual identification of structures. Voxel-based methods, particularly those based upon mutual information, are robust, require no segmentation that can be prone to error, are suitable for multi-modality registration, are highly accurate for brain registration (Maes *et al* 1997), and are suitable for abdominal registration (Carrillo *et al* 2000). There are no reports of using such methods for pelvis and/or prostate registration. For registration of brain and other organs, registration accuracy has been assessed using fiducial markers (Maurer *et al* 1997, Wang *et al* 1996) and anatomical landmarks (Fitzpatrick *et al* 1998, Peters *et al* 2000, Wilson *et al* 1998).

There are challenges for registration in the pelvis and prostate that might reduce the effectiveness of automatic voxel-based registration. First, the abdomen has irregular boundaries, unlike the head to which registration has been most often applied. Second, the normal prostate is a small organ, which when healthy, measures only about 38.0 mm in its widest dimension transversely across the base (Gray 1977). Third, there are potential factors such as different patient positions, and rectal and bladder filling (Herk *et al* 1995) that can stress registration. In addition, it is more difficult to evaluate pelvic and/or prostate registration because no external markers are available.

In the present study, we perform experiments to determine the potential accuracy of registering prostate MR images using a modified mutual information algorithm that uses rigid-body transformations. High-quality, 3D MR image volumes from a commercially available 1.5 T system are used to determine the best possible results. We examine conditions found in potential applications described previously. We develop and use a variety of assessment methods that include measuring displacements of bony landmarks and of the segmented prostate. One goal is to obtain baseline accuracy measurements for planning future applications of registration in prostate cancer management.

2. Methods

2.1. Data acquisition

All MRI volumes were acquired using a 1.5 T Siemens MRI system (Magnetom Symphony, Siemens Medical Systems, Erlangen, Germany). An 8-element phased array body coil was used to ensure coverage of the prostate with a uniform sensitivity. Typically, two anterior and two posterior elements were enabled for signal acquisition. We used two different MR sequences. First, we used a 3D RF spoiled gradient echo steady-state pulse sequence (FLASH) with TR/TE/flip parameters of 12/5.0/60 which give $256 \times 256 \times 128$ voxels over a $330 \times 330 \times 256$ mm field of view (FOV) to yield $1.29 \times 1.29 \times 2.0$ mm voxels oriented to give the highest resolution for transverse slices. The acquisition time was 5.63 min. This sequence was good for pelvic imaging but was not ideal for the prostate. Second, we used a 3D rapid gradient echo sequence (PSIF) designed to acquire the spin–echo component of the steady-state response, rather than the free induction decay. The spin–echo component was formed immediately prior to the RF pulse and it was shifted towards the prior RF pulse through appropriate gradient waveform design. The sequence with 9.4/5.0/60 (TR/TE/flip) yielded $160 \times 256 \times 128$ voxels over a $219 \times 350 \times 192$ mm rectangular FOV and $1.37 \times 1.37 \times 1.5$ mm voxels oriented to give the highest resolution for transverse slices. There was over sampling at 31% in the slice direction to reduce aliasing artifacts. The acquisition time was 4.25 min. Most often, we used the second sequence, which gave excellent image contrast for the prostate and its surroundings.

2.2. Image volumes for registration

We acquired 3D MRI volume images from three prostate cancer patients and three normal volunteers under four conditions simulating anticipated situations in diagnostic and treatment applications. They are *diagnostic position*, *treatment position*, *empty bladder* and *diagnosis 1 week*. In the diagnostic position, the subject lay supine throughout MR scanning. The *reference* volume was always obtained in the diagnostic position. In the treatment position, the subject was supine and his legs were supported at 30°–60° relative to the horizontal position and separated in a 'V' with an angle of 60°–90° between the legs. This is similar

to the lithotomy position used in some prostate therapies and it should provide access for needle insertion in brachytherapy or RF thermal ablation. In some experiments, the subject micturated to create an empty bladder prior to imaging. For each subject, volumes were typically obtained within an imaging session of 1–2 h. We imaged one subject (V2) a week before the standard imaging session and we refer to this volume as diagnosis 1 week. Between volume acquisitions, subjects got off the MRI table, stretched and walked around to ensure that they would assume a different position on the table. The coil array was centred on the prostate. All images of a subject were acquired using the same pulse sequence and acquisition parameters so as to ensure very similar grey values. In total, we registered 22 volume pairs consisting of one pair for each patient, six pairs for each volunteer and one additional pair for volunteer V2.

There are several preprocessing details. Isotropic voxels are created using 3D linear interpolation or higher order interpolation methods (Carrillo *et al* 2000). From the top and bottom of the volume, we optionally crop transverse slices that are over 35 mm away from the prostate rim. Cropping is done to remove slices having reduced brightness due to sensitivity fall off from the receiver coils, artifacts from a small field of view, displacement of the legs in the treatment position, and/or bladder deformation.

2.3. Similarity measurements

We used two similarity measures, mutual information (MI) and correlation coefficient (CC), in our registration. Suppose one volume R is the *reference*, and the other F is *floating*. Their mutual information MI(R,F) is given below (Maes *et al* 1997):

$$\mathrm{MI}(R, F) = \sum_{r,f} p_{RF}(r, f) \log \frac{p_{RF}(r, f)}{p_R(r) \cdot p_F(f)}$$

The joint probability $p_{RF}(r, f)$ and the marginal probabilities $p_R(r)$ of the reference image and $p_F(f)$ of the floating image, can be estimated from the normalized joint and marginal intensity histogram, respectively. The correlation coefficient CC(R, F) is given below (Press *et al* 1993):

$$\mathrm{CC}(R, F) = \frac{\sum (R(r) - \bar{R}(r))(F(f) - \bar{F}(f))}{\sqrt{\sum (R(r) - \bar{R}(r))^2 \sum (F(f) - \bar{F}(f))^2}}$$

Here $\bar{R}(r)$, $\bar{F}(f)$ denote the average intensities of the reference and floating volumes and the summation includes all voxels within the overlap of both volumes.

2.4. Registration algorithm with special features

The registration algorithm includes special features to improve the robustness for MR pelvic images. We use a multi-resolution approach and perform registration from low to high resolution. At low resolution, we resample both images at $1/4$ or $1/2$ number of voxels along each linear dimension, respectively. Iterative optimization of the similarity is used to vary the six rigid-body transformation parameters (three translations and three angles). We use the correlation coefficient at the two lower resolutions because it gives fewer local maximums and because it can be calculated faster than MI. We use MI at full resolution because the peaked similarity function gives a more precise solution than CC.

We created a method to avoid local minima by restarting the registration with randomly perturbed parameters obtained from a uniform distribution about the very first initial guess at each resolution. The distribution was centred on the initial guess because we wanted to use the

best solution from the lower resolution. The algorithm restarts until the absolute correlation coefficient between the reference and registered volumes is above a threshold or the maximum number of restarts is reached. The perturbation range is $\pm 5^{\circ}$ and ± 5 voxels corresponding to ± 27.3, ± 13.7 or ± 6.8 mm for resolutions $1/4$, $1/2$, or full voxels, respectively. Absolute CC is used for the restart test rather than MI because CC has a well-defined range between 0 and 1, because CC provides an independent check of the MI result, and because, as described later, CC has fewer problems with local and incorrect global maximums for registrations at low resolution far from the optimum value.

We record all important results following an optimization cycle including the CC and/or MI values, the number of restarts and the transformation parameters. At the end of processing at a lower resolution, we always select the transformation parameters having the maximum CC value. We then scale the translation parameters appropriately and assign the new parameters to be initial values at the next higher resolution. At the highest resolution, we select the final transformation parameters to be those with the maximum MI value.

There are several implementation details. We used rigid-body transformation (three translations and three angles) and trilinear interpolation. For optimization, we use the downhill simplex method of Nelder and Mead (1965) and the Powell method (1962), but we prefer the former method as described later. Optimization of alignment ends either when the maximum number of MI calculations is reached (typically 500) or the fractional change in MI is smaller than a tolerance (typically 0.001). For the 22 volume pairs reported here, the maximum number of calculations was reached once and this was only at the lowest resolution. Our very first initial guess at the lowest resolution is all zeros for the three displacements and three angles. Based on our experience, we set the CC thresholds at 0.65, 0.70 and 0.75, and the maximum numbers of restarts at 20, 10 and 5, from low to high resolutions, respectively.

3. Evaluation of registration

3.1. Registration accuracy based on bony landmarks

We evaluated registration of the pelvis by measuring the displacement of bony landmarks following registration. We used six easily found bony landmarks consisting of two great sciatic notches, two lesser sciatic notches, the pubic symphysis, and the coccyx, some of which are illustrated in figure 1. Previously, sciatic notches and the pubic symphysis were used to register CT and MRI images for prostate conformal radiation (Kagawa *et al* 1997). To measure landmark displacements, we used *RegViz*, a program written in IDL (Interactive Data Language, Research System Inc., USA) and created in our laboratory for visualizing and analysing registered image volumes. We navigated transverse, coronal and sagittal MR images slice-by-slice to search the landmarks. The same unique features such as corners and intersections were identified with a cursor on magnified images. A single person repeated this six times over a few weeks and results were averaged to give a 3D location for each landmark. A radiologist confirmed the landmark selection. Following registration, we calculated the root-mean-squared (RMS) distance over the six landmarks (Wang *et al* 1996).

Although this method provides an independent means for evaluating skeletal registration accuracy, there is error in localizing the bony landmarks. To determine the effect of localization error, we performed least-squares point-to-point registration (Maurer *et al* 1997) and compared results to MI registration. The rationale is that if we could identify point landmarks without error on the bony pelvis, point-to-point registration would be perfect. Hence, any displacement left after registration is introduced by localization error. We determined the optimal transformation for matching the six corresponding landmarks. Points were

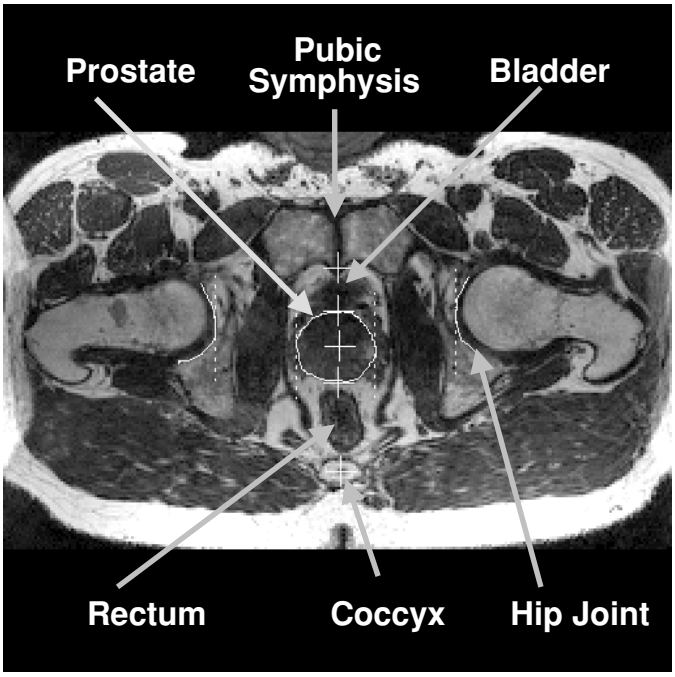

Figure 1. MR prostate image with labelled features used to analyse registration error. This transverse image is from the reference volume of prostate cancer patient P3. The prostate boundary was manually segmented near the image centre. The four vertical dash lines from left to right indicate the rim of the right acetabular socket, the right and left rims of the prostate and the rim of the left acetabular socket, respectively. The five crosses from bottom to top indicate the coccyx, the prostate posterior rim, the 2D centroid automatically calculated from the segmented prostate area, the anterior rim and the pubic symphysis. The image also shows the bottom of the bladder, the rectum, the pubic symphysis and hip joints.

transformed and distances between corresponding points were determined. RMS values were computed and compared to the RMS values from MI registration.

3.2. Registration consistency

We calculated the registration consistency as proposed by Freeborough *et al* (1996). For each of the three volunteers, we used three volumes: reference, diagnosis and empty bladder, all of which were obtained with the subject in the supine position. We call these three volumes A, B and C, respectively. They give three pairs of registrations (A–B, B–C and C–A) and three sets of transformation parameters (T_{ab}, T_{bc}, T_{ca}). Using the transformation parameters, we transformed voxel positions in A to B, and then to C, and then back to A. The distance between the original location and the final position is calculated. Since this is introduced by three transformations, we estimate the error for a single transformation, by multiplying by $3^{-1/2}$ (Freeborough *et al* 1996).

3.3. Voxel displacements

To test the dependency of registration on algorithmic features such as image cropping, one can compare transformation parameters. However, we chose a more meaningful approach

that consisted of finding the average displacement of voxels in a region-of-interest (ROI) (Carrillo *et al* 2000). The 3D distances between transformed voxels were calculated in millimetres and averaged over a cubic ROI just covering the prostate.

3.4. Other evaluation methods including displacement of prostate centroids

We used a variety of other methods to evaluate the registration of the pelvis and prostate. First, we measured potential displacements of the 3D centroid of manually segmented prostates. Second, we used multiple visualization and analysis methods found in RegViz and MIM^{TM} (Zalen LLC, Novelty, OH 44072) such as contour overlap and colour overlay. Third, we calculated the intensity difference between the reference and registered volumes on a voxel-by-voxel basis and computed statistics. Fourth, we calculated the correlation coefficient between corresponding voxels to measure the quality of registration of two MR volumes acquired with identical parameters.

4. Results

4.1. Comparison of mutual information and correlation coefficient

In figure 2, we compare the two similarity measures at different resolutions. Plotted are MI and CC values as a function of translation along the transverse axis where the origin is the optimal transformation. For images at a resolution of $1/4$ voxels along a linear dimension, the CC curves are much smoother than MI, which is noisy and contains many local maximums as shown in figure 2(a). In addition, there is a false global maximum in Figure 2(a) at 18 voxels. At full resolution, figure 2(c) shows that MI is much more peaked than CC, but there is high-frequency noise in the MI curves far from the optimum that gives rise to local maximums that must be avoided. From these figures, we infer that CC is better at low resolution and that MI is better at full resolution when one is close to the optimum value. As described in section 2, our registration algorithm makes use of these features.

4.2. Assessments of pelvic registration

Following registration, we determined displacements between the six bony landmarks. For each subject, there was no consistent displacement of landmarks in one direction versus another. Hence, we measured 3D distances and determined RMS values over the six landmarks. Registration results are plotted in figure 3. The smallest errors are obtained when subjects are in the diagnostic position for both imaging sessions, labelled diagnosis–reference. The average error across the three patients and three volunteers is only 1.6 ± 0.2 mm. Consistently larger errors are obtained when we compare volumes acquired in the treatment position with those in the reference position. Even though the MR acquisition technique used for the patients gave inferior image quality as compared to that for the volunteers, the errors were small.

Additional error analyses are performed on the volunteer images to assess the accuracy of point landmark localization. We used images obtained with the rapid-gradient echo sequence, which have improved the contrast between the prostate and bony landmarks and which give us more confidence in measurements. The isotropic voxels are 1.4 mm on a side, almost as large as the 1.7 ± 0.5 mm error obtained for the volunteer diagnosis–reference data. We assess the error in localizing the bony landmarks by performing point-based registration on 4–6 points per volume pair. The RMS distances after registration averaged across the three volunteers was 1.5 ± 0.2 mm, very nearly the value obtained with MI registration. Hence, the 'error' reported for MI is probably overestimated due to landmark location error. This analysis was

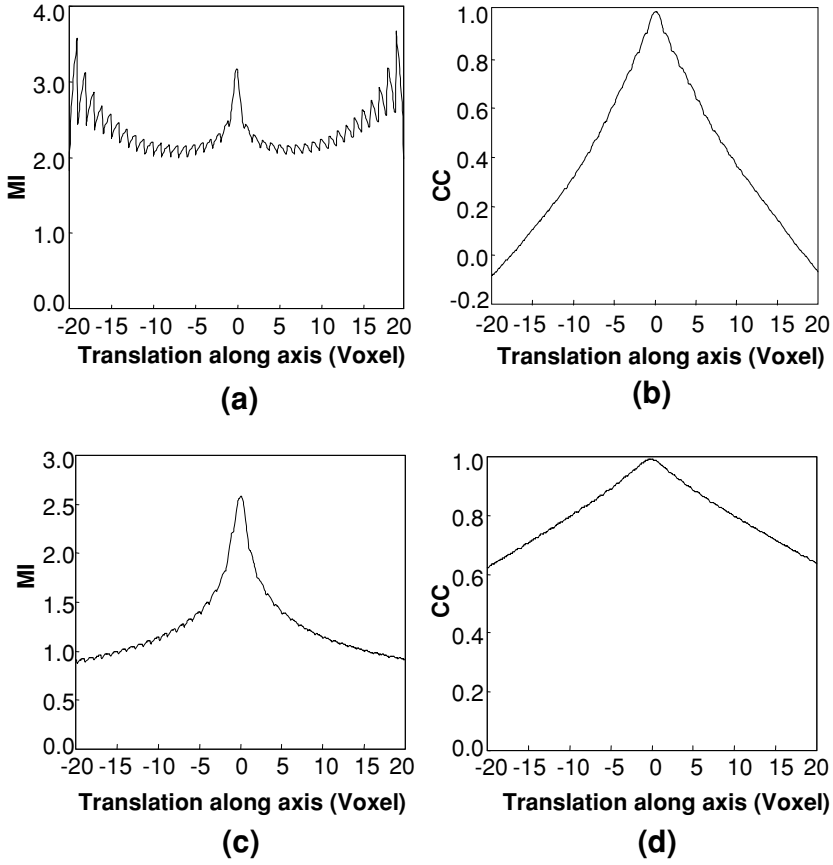

Figure 2. MI and CC similarity functions are plotted to show their relative advantages for registration at different resolutions. Two high-resolution MRI volumes were registered to obtain the optimal parameters. We then computed similarity values as a function of translation along the transverse axis. MI is plotted in (a) and (c); CC is plotted in (b) and (d). Graphs on the top (a) and (b) are at a resolution of 1/4 voxels along a linear dimension, giving a distance between voxel centres of ≈5.5 mm. MI gives a noisy plot having many local maximums and a false global maximum occurs at 18 voxels. Graphs at the bottom are obtained at full resolution. MI has a much sharper peak than CC, which is relatively flat. The voxel size is 1.4 mm. Images are from volunteer V2 in the diagnostic and reference conditions.

prompted by ideas in a previous report (Maurer *et al* 1997) that numerically demonstrated the relationship between point localization uncertainty and point-based registration uncertainty.

Figure 4 shows image intensity differences between reference and registered volumes. The means are quite small with 8 out of 10 registrations giving a mean absolute value <1.5 grey levels, or only 1.7% of typical mean values of 90 grey levels for these 3D MR acquisitions. Again, the only consistent outliers occur when we compare the treatment position to the reference. For the case diagnosis–reference, extremely small image differences are found with V1 giving 0.1 ± 1.6 grey values, a standard deviation that compares favourably with the expectation from image noise alone, or 1.5 grey values. For this volume pair, the subtracted images have very little structure except at the skin surface, indicating excellent registration (not shown). We know that 3D alignment is achieved because all slices across the entire pelvis are well aligned and because rendered images show that the prostate matches well.

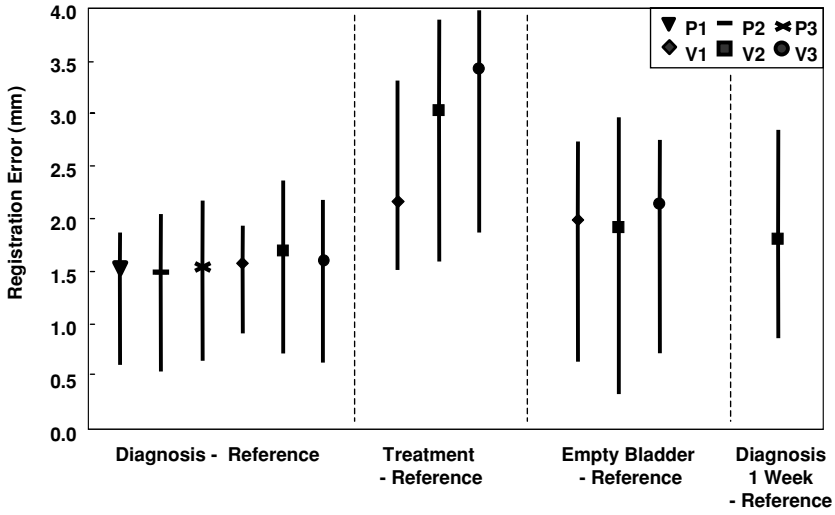

Figure 3. Registration error as determined from bony landmarks. RMS distances as well as maximums and minimums are plotted that show the spread of the data. Conditions along the *x*-axis such as the 'treatment' position are described in section 2. P's and V's refer to patients and volunteers, respectively. Averaging data across all subjects for the best case (diagnosis–reference) gives 1.6 ± 0.2 mm. Averages are 2.9 ± 0.7 mm and 2.0 ± 0.1 mm for treatment–reference and empty bladder–reference respectively.

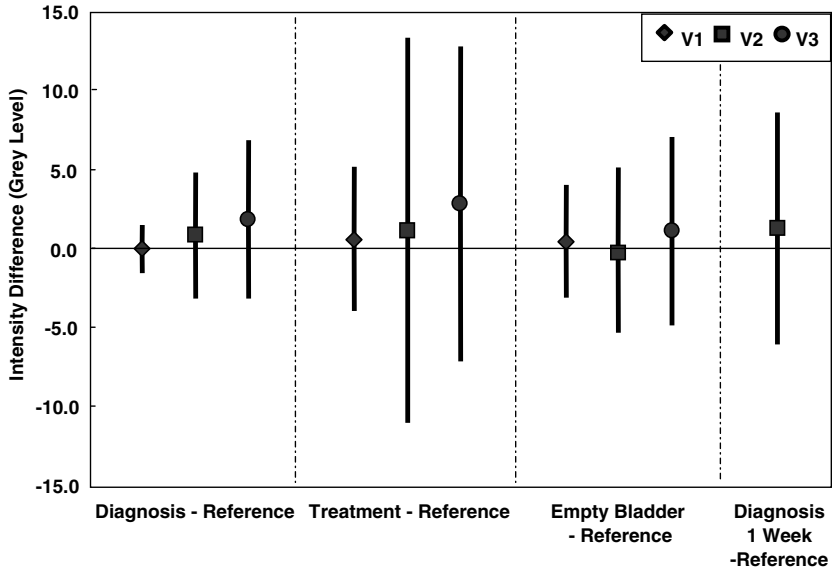

Figure 4. Image intensity difference between registered volume pairs. Means and standard deviations calculated over the entire volumes are plotted. Other details are given in the legend of figure 3. Eight out of ten registrations have a mean absolute grey level difference less than 1.5 grey levels. Average results are 0.9 ± 1.8, 1.5 ± 3.9, and 0.4 ± 2.6 grey levels, or 1.0, 1.7 and 0.4% of typical mean values of 90 grey levels, for diagnosis–reference, treatment–reference and empty bladder–reference respectively.

Registration consistency, as described in section 3, provides yet another means of evaluating the quality of registration. Values were 0.4, 0.8 and 0.7 mm for volunteers V1, V2, and V3, respectively. The average is 0.6 ± 0.2 mm, a value less than half the dimension of a voxel indicating excellent registration consistency.

4.3. Assessment of prostate registration

We determined the quality of prostate registration by visually examining nearly all of the roughly 800 registered image slices using one or more of the methods found in RegViz and MIM^{TM}. A typical example for the case of diagnosis–reference is shown in figure 5 where the boundary overlap is excellent and probably within the manual segmentation error. In some other cases such as treatment–reference, small displacements of the prostate were observed. In a typical volume pair, the prostate is displaced to the posterior direction by $\approx$3.0 mm when the legs are raised. There are no obvious displacements in other directions.

Centroid vector displacements can also be analysed following registration. For the case of diagnosis–reference, centroid displacements are only 1.4 ± 0.2 mm. In the case of treatment–reference, there is a consistent displacement ($\approx$3 mm) in the posterior direction with relatively little change in the two orthogonal directions. In the case of empty bladder–reference, two of the three volume pairs show a displacement in the posterior direction while the other is displaced in the anterior direction. Finally, in the case of a diagnostic volume obtained 1 week before the reference, there was a 4 mm displacement in the caudal direction due mostly to changes in rectal and bladder filling. Because the 3D centroid of the prostate averages over a large region, we believe these measurements to be relatively insensitive to segmentation error. Even so, we consider the uncertainty to be at least 1 mm, and displacements less than this should be disregarded. All significant results above can be visually confirmed.

Prostate volumes were measured for each subject. The typical difference between volumes in an imaging session was $<$1.5%, indicating that segmentation errors were small and that prostate volumes did not change. The average prostate volume for the healthy volunteers was 23.9 ± 3.2 cm^3. Volume measurements are particularly useful for clinicians when assessing the response of prostate cancer treatments such as brachytherapy, chemo- or radiation therapy.

4.4. Effects of image cropping

In figure 6, we plot registration error as determined from bony landmarks with and without the cropping operation described in section 2. For the case treatment-reference, cropping always improved registration accuracy, and for V3, error reduced greatly from 12.6 mm to 3.4 mm. For all other cases, subjects were always in the supine position with legs flat on the table, and there was no consistent effect of cropping. If anything, cropping tends to increase error in these cases, with an increase in five of seven volume pairs. Correlation coefficient always improves with image cropping.

Displacements of bony landmarks might significantly overestimate the change near the prostate. Hence, as described in section 3, we investigated the displacement of voxels in a ROI surrounding the prostate between registrations with and without cropping. For nine of ten analysed volume pairs, the average voxel displacement was $<$0.5 mm indicating that prostate registration is fairly insensitive to cropping. However, for V3 treatment–reference, a much larger voxel displacement of 7.4 mm was obtained indicating that cropping is critical for this volume pair.

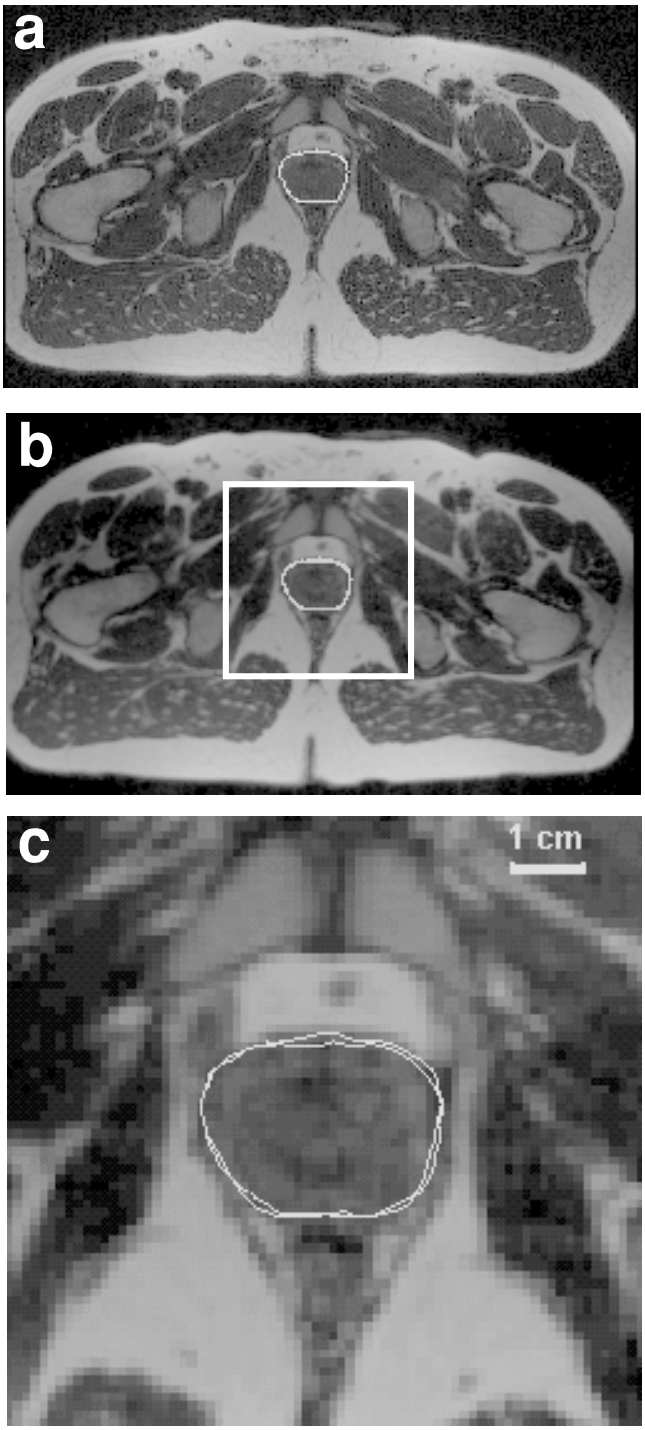

Figure 5. The prostate overlap between reference and registered images. Following registration, the prostate was manually segmented in reference (a) and diagnosis (b) images. The rectangular region in (b) is zoomed in (c) with both boundaries superimposed. Images are from volunteer V2.

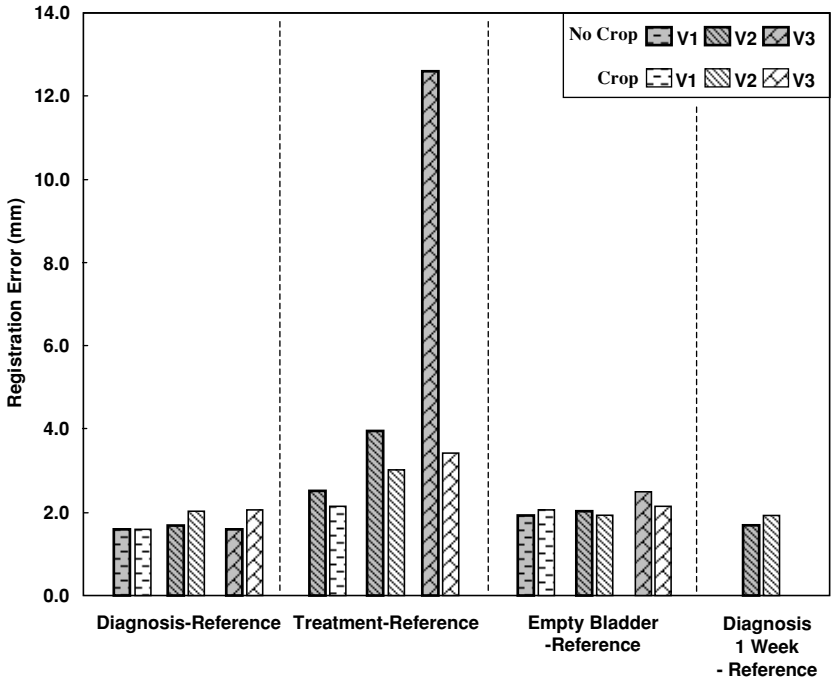

Figure 6. The effect of image cropping on registration accuracy. The light and dark bars are RMS distances between bony landmarks with and without image cropping, respectively, as defined in section 2. Conditions on the *x*-axis are described in section 2.

4.5. Implementation issues

The algorithm was quite robust and always gave very nearly the same transformation parameters (less than 0.01 voxels and 0.01°) for the 22 volume pairs in this study using a wide variety of initial guesses. The restarting and multi-resolution features are important and we report some results for a typical volume pair registration. The multi-resolution approach enabled the program to get close to the final value quickly because of the reduced number of calculations. That is, the time for reformatting at the lowest resolution (1/4) was 9.8 s, which was less than 1/59 times that at the highest resolution, a value nearly equal to 1/64 expected from the change in the number of voxels. The number of restarts was 5, 1 and 1 for resolutions at 1/4, 1/2 and the full number of voxels. Each call to the Simplex optimization resulted in 55 to 94 MI evaluations before the tolerance (0.001) was reached. The simplex optimization method worked about 1.5–2.0 times faster than the Powell method in our implementation. The time for registration using Simplex, typically 10 mines on a Pentium IV, 1.7 GHz CPU, with 1 Gbytes of memory, could probably be greatly improved with optimized C code rather than IDL.

5. Discussion

5.1. Registration accuracy

Our results suggest that MI can be used to accurately register, with an error on the order of a voxel, MR pelvic images obtained under similar conditions. Because it gives an independent,

true 3D measurement, we like to use the method of point bony landmarks to assess accuracy. However, as argued in section 4.2, the true MI registration accuracy might be better than our ability to measure it with point bony landmarks. That is, following point landmark registration, the distance between registered, corresponding landmarks was on the order of that following MI registration. Very possibly, MI is more accurate than point registration using bony landmarks. Additional, independent evidence of excellent MI accuracy comes from the very low error value from the registration consistency measurement (0.6 ± 0.2 mm). Interestingly, this is obtained even though the interpolation artifact present in MI similarity surfaces should reduce the likelihood of sub-voxel accuracy (Pluim *et al* 2000a). Our results for the pelvis with image volumes obtained under the same conditions compare favourably with those for the brain, where MI registers images very accurately giving errors as small as 0.7–0.8 mm for CT-MR (West *et al* 1997).

Visual and quantitative evaluation of prostate organ movement showed good registration even when we acquired images under conditions that greatly stressed the ability to register the images. The small prostate displacements in our study are consistent with earlier reports on respiration-induced prostate movement of $\leqslant 1$ mm for most patients in supine position with 'quiet' respiration (figure 3(b) of the report by Malone *et al* (2000)). The difference between the treatment and diagnostic positions resulted in the most consistent and largest displacement of the prostate. When images were acquired in the diagnostic position 1 week apart, there was significant displacement of the prostate due to a change in rectal filling. This is consistent with previously reported results (Herk *et al* 1995, Tenhaken *et al* 1991), which found rectal filling to be a significant factor in prostate displacement.

There are ways of limiting the small displacements of the prostate. One obvious remedy is to acquire images in the same position. That is, if we want to register an image with that obtained in the treatment position, we should obtain it in the treatment position. Although it is unknown how accurately one must repeat the treatment position, a device to support and constrain the legs is probably required. In addition, there is a dependence of registration error on bladder and rectum content. One solution is use clinical preparations often employed to void the bladder and rectum prior to prostate imaging or therapy. We anticipate that this might even lessen prostate displacements between the diagnostic and treatment positions.

We must consider our results with regard to potential applications such as those described in the section 1. First, registered images acquired before and immediately after treatment can be used to determine whether a tumour is adequately treated. Second, serial examinations can be registered to determine tumour progression or regression. Third, registration of functional images from other modalities such as nuclear medicine or from MR spectroscopy can give molecular markers for prostate cancer (Lee *et al* 2001a, 2001b). Fourth, we want to register high-quality MR images with a few live-time interventional MR images to aid treatment decisions (Fei *et al* 2001). Our results indicate that registering images from the treatment and diagnostic positions can lead to errors and potential steps are described above to limit this error. With images acquired in the same position, our results place a lower limit on registration error of about 1 voxel. Experiments are being conducted to see if this can be achieved with interventional MR images (Fei *et al* 2001).

5.2. Assessment of registration

We are involved in a long-term effort to use registration for detection, assessment and therapy of prostate cancer. Hence, we have developed and used several methods to assess pelvic and prostate registration.

It is highly desirable to have an automatic method for evaluating the quality of a registration so that a poor one can be flagged before it is used clinically. The correlation coefficient would be applicable whenever one uses MR images obtained with identical pulse sequences. It compares favourably with the bony landmark results. Registration consistency provides an additional means to evaluate registration accuracy that does not rely on operator interaction.

Other evaluation methods are applicable for clinical or research applications. RegViz and MIM^{TM} provided visual inspection tools for quick evaluation of the quality of registration and potential prostate displacement. Such methods can be used to verify the quality of registration and possibly account for small displacements in some applications. Boundary overlays provide a good means to evaluate organ deformation as well as displacement. Point anatomical landmarks provide a useful, independent test; but it is time consuming to identify them; and MI might be more accurate than the point landmarks. Centroids are obtained reliably because small segmentation errors are removed by integrating over the entire prostate volume. Centroids provide a good means of quantifying prostate displacements.

5.3. Algorithm with combined similarity measures

Using both CC and MI at different resolutions was an important feature that increased robustness. When only mutual information was used, registrations at low resolution sometimes gave false solutions that misled registration at the next higher resolution. However, CC performed well and gave many fewer local maximums at the lower resolutions (figure 2(a) and (b)). But MI gave a more accurate solution at the full resolution due to the peaked MI surface (figure 2(c) and 2(d)). Our registration algorithm combined advantages from the two similarity measures.

There are probably several reasons why mutual information does not work well at low resolution. First, the similarity curve is noisy with periodic oscillations from the so-called interpolation artifact (Pluim *et al* 2000a) that is accentuated at reduced resolutions (Pluim *et al* 2000b). This results in the many local maximums in figure 2(a) that can trap the optimization. A similar result was reported for brain registration (Lau *et al* 2001, Maes *et al* 1997). Second, when images are of low resolution and there is only a small region of overlap, the mutual information function can even contain incorrect global maximmums (Pluim *et al* 2000b). Such a result was found in figure 2(a) where the global maximum was obtained at very large displacements where the overlap was reduced. This occurs because MI is not only a function of how well the images match in the overlap, but also by how much information is provided by the two images in the overlap (Studholme *et al* 1997).

5.4. Computer implementation

Accuracy is an important issue for automatic registration, but there are others such as robustness, speed, and requirements for operator interaction. With the multi-resolution and restarting features, our modified MI algorithm is quite robust. For a wide range of initial guesses, it worked well for all 22 volume pairs reported here. Three of the volume pairs were from patients and we are confident that routinely acquired clinical images will have sufficient quality for registration. Because good starting values are unimportant, operator interaction is minimal. In one instance, cropping of the legs was important for registering an image volume obtained in the treatment position with that in the diagnostic position. It is not surprising that legs in a very different position have to be cropped. Although this is easy to do manually, we can probably determine an automated method if it is deemed desirable.

The mutual information similarity measure is quite robust. Even though our images are very similar, we had less success with some other measures such as the sum of the squared image difference. An advantage of MI is that it can be used with images from different modalities, a feature that we are starting to use.

6. Conclusion

We have developed an automatic volume registration algorithm with special features for the pelvis and prostate MR volumes. When both volumes in a pair were obtained in the diagnostic position under comparable conditions, our internal registration measures showed accuracy on the order of a voxel. We believe that the MR image registration method is sufficiently accurate and robust for a variety of applications of interest in the pelvis and prostate. We are beginning to explore these applications in clinical procedures and animal experiments.

Acknowledgments

This research was supported by NIH grants R01-CA84433-01 to DLW and R33-CA88144-01 to JLD.

References

Antolak J A, Rosen II, Childress C H, Zagars G K and Pollack A 1998 Prostate target volume variations during a course of radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **42** 661–72

Balter J M, Sandler H M, Lam K, Bree R L, Lichter A S and Ten Haken R K 1995 Measurement of prostate movement over the course of routine radiotherapy using implanted markers *Int. J. Radiat. Oncol. Biol. Phys.* **31** 113–8

Boni R A H, Boner J A, Debatin J F, Trinkler F, Knonagel H, Vonhochstetter A, Helfenstein U and Krestin G P 1995 Optimization of prostate carcinoma staging-comparison of imaging and clinical methods *Clin. Radiol.* **50** 593–600

Carrillo A, Duerk J L, Lewin J S and Wilson D L 2000 Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment *IEEE Trans. Med. Imaging* **19** 175–85

Fei B, Wheaton A, Lee Z, Nagano K, Duerk J L and Wilson D L 2001 Robust registration algorithm for interventional MRI guidance for thermal ablation of prostate cancer *Proc. SPIE* **4319** 53–60

Fitzpatrick J M, Hill D L G, Shyr Y, West J, Studholme C and Maurer C R 1998 Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain *IEEE Trans. Med. Imaging* **17** 571–85

Freeborough P A, Woods R P and Fox N C 1996 Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders *J. Comput. Assist. Tomogr.* **20** 1012–22

Gray H 1977 *Anatomy, Descriptive and Surgical (The Classic Collector's Edition)* (New York: Gramercy Books) p 1010

Hamilton R J, Blend M J, Pelizzari C A, Milliken B D and Vijayakumar S 1999 Using vascular structure for CT-SPECT registration in the pelvis *J. Nucl. Med.* **40** 347–51

Herk M V, Bruce A, Kroes A P G, Shouman T, Touw A and Lebesque J V 1995 Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration *Int. J. Radiat. Oncol. Biol. Phys.* **33** 1311–20

Herk M V, de Munck J C, Lebesque J V, Muller S, Rasch C and Touw A 1998 Automatic registration of pelvic computed tomography data and magnetic resonance scans including a full circle method for quantitative accuracy evaluation *Med. Phys.* **25** 2054–67

Hill D L G, Batchelor P G, Holden M and Hawkes D J 2001 Medical image registration *Phys. Med. Biol.* **46** 1–45

Kagawa K, Lee W R, Schultheiss T E, Hunt M A, Shaer A H and Hanks G E 1997 Initial clinical assessment of CT-MRI image fusion software in localization of the prostate for 3D conformal radiation therapy *Int. J. Rad. Oncol. Biol. Phys.* **38** 319–25

Lau Y H, Braun M and Hutton B F 2001 Non-rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio *Phys. Med. Biol.* **46** 1297–1319

Lee Z, Kemper C, Muzic R F, Berridge M S and Wilson D L 2001a Quantitative pulmonary imaging based on PET-CT co-registration with warping *J. Nucl. Med.* **42** 10–11.

Lee Z, Sodee D B, Duerk J L, Nelson A D and Berridge M S 2000 Automatic registration of SPECT-MRI in the pelvis *J. Nucl. Med.* **41** 232

Lee Z, Sodee D B, Faulhaber P F, Lancaster T L, MacLennan G T and Wilson D L 2001b Comparison of SPECT and PET imaging for prostate cancer with histological correlation *J. Nucl. Med.* **42** 1226

Lewin J S, Connell C F, Duerk J L, Chung Y C, Clampitt M E, Spisak J, Gazelle G S and Haaga J R 1998 Interactive MRI-guided radiofrequency interstitial thermal ablation of abdominal tumors: clinical trial for evaluation of safety and feasibility *J. Magn. Reson. Imaging* **8** 40–7

Liehn J C, Loboguerrero A, Perault X and Demange L 1992 Superimposition of computed-tomography and single photon emission tomography immunoscintigraphic images in the pelvis—validation in patients with colorectal or ovarian carcinoma recurrence *Eur. J. Nucl. Med.* **19** 186–94

Maes F, Collignon A, Vandermeulen D, Marchal G and Suetens P 1997 Multimodality image registration by maximization of mutual information *IEEE Trans. Med. Imaging* **16** 187–98

Malone S, Crook J M, Kendal W S and Szanto J 2000 Respiratory-induced prostate motion: quantification and characterization *Int. J. Radiat. Oncol. Biol. Phys.* **48** 105–9

Maurer C R, Fitzpatrick J M, Wang M Y, Galloway R L, Maciunas R J and Allen G S 1997 Registration of head volume images using implantable fiducial markers *IEEE Trans. Med. Imaging* **16** 447–62

Milosevic M, Voruganti S, Blend R, Alasti H, Warde P, McLean M, Catton P, Catton C and Gospodarowicz M 1998 Magnetic resonance imaging (MRI) for localization of the prostatic apex: comparison to computed tomography (CT) and urethrography *Radiother. Oncol.* **47** 277–84

Narayana V, Roberson P L, Winfield R J and McLaughlin P W 1997 Impact of ultrasound and computed tomography prostate volume registration on evaluation of permanent prostate implants *Int. J. Radiat. Oncol. Biol. Phys.* **39** 341–6

Nelder J and Mead R A 1965 A simplex method for function minimization *Comput. J.* **7** 308–13

Peters A R, Muller S H, de Munck J C and van Herk M 2000 The accuracy of image registration for the brain and the nasopharynx using external anatomical landmarks *Phys. Med. Biol.* **45** 2403–16

Pluim J P W, Maintz J B A and Viergever M A 2000a Interpolation artefacts in mutual information-based image registration *Comput. Vis. Image Understanding* **77** 211–32

Pluim J P W, Maintz J B A and Viergever M A 2000b Image registration by maximization of combined mutual information and gradient information *IEEE Trans. Med. Imaging* **19** 809–14

Powell M J D 1962 An iterative method for finding stationary values of a function of several variables *Comput. J.* **5** 147–51

Press W H, Flannery B P, Teukolsky S A and Vellerling W T 1993 *Numerical Recipes in C: The Art of Scientific Computing* 2nd edn (New York: Cambridge University Press)

Remeijer P, Geerlof E, Ploeger L, Gilhuijs K, van Herk M and Lebesque J V 2000 3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients *Int J. Radiat. Oncol. Biol. Phys.* **46** 1281–90

Roach M, FaillaceAkazawa P, Malfatti C, Holland J and Hricak H 1996 Prostate volumes defined by magnetic resonance imaging and computerized tomographic scans for three-dimensional conformal radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **35** 1011–8

Roeske J C, Forman J D, Mesina C F, He T, Pelizzari C A, Fontenla E, Vijayakumar S and Chen G T Y 1995 Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **33** 1321–9

Scott A M *et al* 1994 Clinical validation of SPECT and CT/MRI image registration in radiolabeled monoclonal-antibody studies of colorectal carcinoma *J. Nucl. Med.* **35** 1976–84

Studholme C, Hill D L G and Hawkes D J 1997 Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures *Med. Phys.* **24** 25–35

Tenhaken R K, Forman J D, Heimburger D K, Gerhardsson A, Mcshan D L, Pereztamayo C, Schoeppel S L and Lichter A S 1991 Treatment planning issues related to prostate movement in response to differential filling of the rectum and bladder *Int. J. Radiat. Oncol. Biol. Phys.* **20** 1317–24

Wang M Y, Maurer C R, Fitzpatrick J M and Maciunas R J 1996 An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head *IEEE Trans. Biomed. Eng.* **43** 627–37

West J *et al* 1997 Comparison and evaluation of retrospective intermodality brain image registration techniques *J. Comput. Assist. Tomogr.* **21** 554–66

Wilson D L, Carrillo A, Zheng L, Genc A, Duerk J L and Lewin J S 1998 Evaluation of 3D image registration as applied to MR-guided thermal treatment of liver cancer *J. Magn. Reson. Imaging* **8** 77–84

Fei BW, Wheaton A, Lee Z, Duerk JL, Wilson DL. Automatic MR volume registration and its evaluation for the pelvis and prostate. Physics in Medicine and Biology 2002;47:823-838.