Available online at www.sciencedirect.com
SCIENCE DIRECT
Medical Image Analysis 10 (2006) 385–395
MEDICAL IMAGE ANALYSIS
www.elsevier.com/locate/media

Cadaver validation of intensity-based ultrasound to CT registration

G.P. Penney [a,b,*], D.C. Barratt [a,b], C.S.K. Chan [b], M. Slomczykowski [c], T.J. Carter [a,b], P.J. Edwards [b,d], D.J. Hawkes [a,b]

[a] *Centre for Medical Image Computing, University College London, 2nd Floor Malet Place Engineering Building, Malet Place, Off Torrington Place, London, WC1E 6BT, UK*
[b] *Computational Imaging Science Group, Guy's, King's and St. Thomas' Schools of Medicine, King's College London, SE1 9RT, UK*
[c] *i-Orthopaedics, Depuy a Johnson & Johnson Company, Leeds, UK*
[d] *Department of Surgical Oncology and Technology, Imperial College London, London, UK*

Received 2 June 2005; received in revised form 7 November 2005; accepted 12 January 2006
Available online 7 March 2006

Abstract

A method is presented for the rigid registration of tracked B-mode ultrasound images to a CT volume of a femur and pelvis. This registration can allow tracked surgical instruments to be aligned with the CT image or an associated preoperative plan. Our method is fully automatic and requires no manual segmentation of either the ultrasound images or the CT volume. The parameter which is directly related to the speed of sound through tissue has also been included in the registration optimisation process. Experiments have been carried out on six cadaveric femurs and three cadaveric pelves. Registration results were compared with a "gold standard" registration acquired using bone implanted fiducial markers. Results show the registration method to be accurate, on average, to 1.6 mm root-mean-square target registration error.

Keywords: Ultrasound to CT registration; Image-guided surgery

1. Introduction

Recent years have seen the emergence of image-guidance systems for orthopaedic surgery (Amiot and Poulin, 2004; Barger et al., 1998; DiGioia et al., 1998; Jaramaz et al., 1999; Nabeyama et al., 2004). All of these systems require a registration between physical and image space. Most systems achieve this by identifying point landmarks (either fiducial markers or anatomical positions) or by delineating surfaces using a tracked pointer. In order to achieve accurate registrations, bony landmarks or fiducial markers attached to bone should be located. Implanting fiducials or exposing additional bone surface for registration is invasive and may cause additional pain (Honl et al., 2003; Nogler et al., 2001) and increase the risk of infection, whereas limiting surface information to regions exposed in standard procedures may adversely affect registration accuracy (Heger et al., 2005). Therefore, there is a trade off between the invasiveness and the accuracy of the registration process. The drive to develop less invasive registration methods will increase with the adoption of minimally invasive surgical techniques (Berger, 2003; DiGioia et al., 2003). This is very important for computer assisted orthopaedic surgery (CAOS) systems, as one of the main arguments for their use is that they have the potential to greatly reduce the invasiveness of procedures (DiGioia and Nolte, 1998).

A number of authors have proposed using imaging devices to acquire information on the position of bony structures in the operating room. The use of X-ray or fluoroscopy images has been put forward with promising results (Guéziec et al., 1998; Livyatan et al., 2003). Proposals to use ultrasound (US) for registration in image-guided surgery go back approximately a decade (Barbe et al., 1993; Ault and Siegel, 1994; Lavallée et al., 1995). Tracked

* Corresponding author.
E-mail address: g.penney@cs.ucl.ac.uk (G.P. Penney).

1361-8415/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2006.01.003

US has advantages over X-ray imaging as it can acquire 3D information, rather than 2D projections, is typically cheaper and more portable than X-ray equipment, and is non-ionising. However, there are a number of image artefacts associated with US: speckle noise, saturation of the reflected echo at the bone-tissue boundary and variation of the speed of sound in different tissues can all make precise location of the bone surface difficult.

Methods to match B-mode US images to bone have been reported by a number of authors. The main application areas have involved registrations on vertebrae (Barbe et al., 1993; Lavallée et al., 1995; Muratore et al., 2002; Brendel et al., 2002; Ionescu et al., 1999; Kowal et al., 2003), pelves (Amin et al., 2003; Ionescu et al., 1999; Tonetti et al., 2001) and long bones (Ault and Siegel, 1994; Brendel et al., 2003; Jaramaz et al., 2003). The work presented in this paper differs from previously published work in three main ways:

(1) We have used an intensity-based algorithm, therefore no segmentation is required in either the US or CT volume. All previously published US to CT bone registration algorithms have required segmentation of the CT volume. This does not affect the clinical utility of such approaches, as the segmentation can be carried out prior to the procedure, and automated segmentation methods exist (Kang et al., 2003). However, segmentation errors can occur, particularly at joint interfaces, which may affect registration accuracy. Some algorithms also require segmentation of the US images. This is more problematic to the clinical process as, due to time constraints during a procedure, the segmentation must be carried out quickly. Manual segmentation is therefore not feasible, and accurate automated segmentation of US images is a challenging problem, although some techniques have recently been published (Daanen et al., 2004). Another incentive for developing an intensity-based technique is that we believe it has the potential to produce more accurate results than techniques which require segmentation of one or both modalities. The improved accuracy of intensity-based compared to surface-based registration techniques has been shown for CT to MR registration of head volumes (West et al., 1999).
(2) The algorithm presented here also optimises one of the probe calibration parameters: the parameter which is directly related to the average speed of sound within the US imaged medium.
(3) Our validation strategy uses human cadavers, results are compared to an independently calculated "gold standard" registration based on bone implanted fiducial markers, and we use clinically realistic starting positions. Much of the previous work has either used dry or plastic bones in water baths (Barbe et al., 1993; Ault and Siegel, 1994; Lavallée and Szeliski, 1995; Muratore et al., 2002), where the lack of soft tissue structures and easy access to all of the bone surface greatly simplifies the registration process. The use of human cadavers enables us to acquire a set of US images which should closely represent a clinical dataset in two ways: firstly in terms of the individual image characteristics, e.g., speckle, distortion, speed of sound variations; and secondly in terms of in which regions of human femur and pelvis is it possible to obtain clear images of the bone surface using US. Previous studies have used a number of validation methods: quoting residual errors (Lavallée and Szeliski, 1995), visual inspection (Lavallée and Szeliski, 1995; Brendel et al., 2002), positioning surgical instruments (Barbe et al., 1993), comparison with an image-guided surgery system (Amin et al., 2003; Kowal et al., 2003) and the use of fiducial markers (Muratore et al., 2002; Jaramaz et al., 2003). However, fiducial markers have only been used previously with either dry bones in water baths (Muratore et al., 2002) or animal cadavers (Jaramaz et al., 2003). An ultrasound-based computer-assisted surgery system has been used clinically for the positioning of iliosacral screws on four patients (Tonetti et al., 2001). Postoperative CT volumes were used to verify screw positioning, and hence also provided a validation of registration accuracy. Promising results were obtained, however the registration algorithm required manual segmentation of the US images, a process which they describe as being "very delicate" and time consuming.

2. Method

2.1. Overview of system

The registration algorithm used in this paper is an extension of an algorithm previously described for registering US images to magnetic resonance images of the liver (Penney et al., 2004). An overview of the registration system is given in Fig. 1. An optical localiser (Optotrak 3020, Northern Digital Inc., Ont., Canada) was used to track an US probe and a dynamic reference object (DRO) which is rigidly attached to the bone. Our aim is to calculate the registration transformation, $\mathbf{T}_{\text{reg}}$, which transforms physical positions in the coordinate system of the DRO into voxel positions within a preoperative CT volume.

To calculate $\mathbf{T}_{\text{reg}}$, 3D freehand US images are acquired of the bone surface. These images are then registered to the preoperative CT volume to obtain transformation, $\mathbf{T}$, which maps pixel positions, $\mathbf{x}_{\text{US}}$, in the ith US image to voxel positions, $\mathbf{x}_{\text{CT}}$, within the CT volume. Transformation $\mathbf{T}$ is computed from four separate transformations:

$$\mathbf{T} = \mathbf{T}_{\text{reg}}(\mathbf{T}^{i}_{\text{DRO}})^{-1}\mathbf{T}^{i}_{\text{probe}}\mathbf{T}_{\text{cal}} \quad (1)$$

The calibration matrix, $\mathbf{T}_{\text{cal}}$, which transforms positions from the US image to positions relative to the infrared-

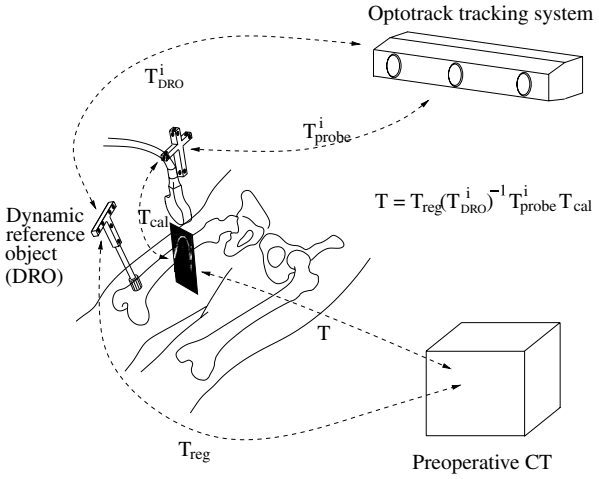

Fig. 1. Overview of system and transformations.

emitting diodes (IREDs) attached to the probe was calculated by scanning a pin-head in a water–glycerol solution (Barratt et al., 2001). This calibration consists of eight parameters, six rigid-body parameters and two scaling parameters. The transformation $\mathbf{T}^i_{DRO}$ transforms positions relative to the IREDs on the dynamic reference object (DRO) to positions relative to the cameras on the Optotrak localiser. Similarly, the transformation $\mathbf{T}^i_{probe}$ transforms positions relative to the coordinate system defined by IRED positions on the US probe to the Optotrak coordinate system.

The algorithm described in this paper calculates the transformation $\mathbf{T}$ by altering the six rigid-body parameters which define transformation $\mathbf{T}_{reg}$ and the probe calibration scaling parameter in the vertical (or y) US image direction, s_y, in order to optimise the value of a similarity measure between the US and CT images. After optimisation the rigid-body parameters (or the equivalent transformation, i.e., $\mathbf{T}_{reg}$) can be used to determine the position and orientation of any tracked and calibrated object (such as a surgical drill) in the CT scan, and so relate the intraoperative positions of such instruments to a preoperative plan described in the CT image.

Following the conversion of CT and US images into probability images, described in Section 2.2, the similarity measure for the algorithm is calculated as follows. The current estimate of the rigid-body parameters and s_y are used to reslice the CT probability image in the plane of each US slice. The pixel values in these reformatted slices and in the US probability images are then compared using the normalised cross-correlation similarity measure (also known as linear correlation, or *Pearson's r* (Press et al., 1992)).

Our optimisation method differs slightly from that described in our previous publication (Penney et al., 2004), in that a multi-resolution approach is used and s_y is also optimised. A three stage optimisation approach is used: low-resolution, high-resolution and then high-resolution including s_y optimisation.

The algorithm proceeds by altering each of the rigid-body parameters (and s_y in the last stage) by a given step size, and calculating the value of the similarity measure. The algorithm then updates the current optimum value of the registration parameters by moving in the direction which produced the greatest improvement in the similarity measure. When no improvement can be made by altering any of the parameters the step size is reduced by a factor of two. When a minimum stepsize is reached, the algorithm moves on to the next stage. Table 1 summarises the parameters used in each stage of the optimisation process. The US image resolution was not altered at each stage, but it was reduced by a factor of four prior to registration, by Gaussian blurring and subsampling to give pixel sizes of 0.486×0.486 mm^2.

2.2. Formation of probability images

The US and CT images are converted from intensity images, $I(\mathbf{x})$, into probability images, $P(\mathbf{x})$, where $\mathbf{x}$ represents the image position, and $P(\mathbf{x})$ represents the probability of a pixel or voxel containing a bone-to-soft-tissue interface. The probability images are calculated using a probability density function (PDF), p, which, given a set of image features $\mathbf{F}(\mathbf{x})$, returns an estimate of the probability that position $\mathbf{x}$ is a bone-edge, i.e., $P(\mathbf{x}) = p(\mathbf{F}(\mathbf{x}))$. The image features extracted from the CT and US images are described in the following sections. It is important to note that our aim here is not to produce a perfect segmentation of the bone boundary, but to produce probability images of sufficient quality to allow accurate and robust registrations.

2.2.1. Converting the CT volume

The PDF for the CT image, p_{CT}, was based on two image features, i.e., $\mathbf{F}(\mathbf{x})_{CT} = (f_{CT1}(\mathbf{x}), f_{CT2}(\mathbf{x}))$. Both CT features were calculated in 2D by applying operators to

Table 1
Summary of optimisation parameters

Stage	Blurring σ	Subsampling, voxel size	Translation step sizes	Rotation (°) step sizes	s_y % change
Low-res	4	2	$4 \rightarrow 0.063$	$2 \rightarrow 0.031$	0
High-res 1	1	1	$1 \rightarrow 0.063$	$0.5 \rightarrow 0.031$	0
High-res 2	1	1	$1 \rightarrow 0.063$	$0.5 \rightarrow 0.031$	$1 \rightarrow 0.063$

Except where specified in the table all measurements are in mm. The blurring and subsampling apply to the CT volume and σ denotes the standard deviation of the Gaussian blurring kernel. The $\rightarrow$ symbol separates the maximum and minimum step sizes (and % change for s_y) used by the algorithm at each stage of the registration.

each CT slice in turn. The first CT feature, $f_{CT1}(\mathbf{x})$, was the intensity of a gradient image $G_{2D}(\mathbf{x})$. This gradient image was set to be the magnitude of vertical and horizontal gradient images calculated by convolving each 2D CT slice $I_{2D}(\mathbf{x})$ with 3×3 vertical and horizontal Sobel operators (Sonka et al., 1998). The second feature, $f_{CT2}(\mathbf{x})$, was set equal to the maximum value under a 3×3 mask centred on pixel $\mathbf{x}$. The effect of the first feature is to highlight edges within the CT volume, while the second feature helps distinguish between bone-to-soft-tissue edges and skin-to-air edges. Fig. 2 shows a sagittal cross-section through the probability image of a femur and the corresponding section through the original CT volume.

The PDF p_{CT} was calculated using a set of CT images as training data. Two sets of voxels in the training data were identified: (i) all of the voxels which lie on a bone-to-soft-tissue boundary, S_{edgeCT}; and (ii) all the voxels in the training data S_{CT}. Therefore, $S_{edgeCT} \subset S_{CT}$. The set of voxels, S_{edgeCT}, was found in the following way: Each of the training CT images were taken in turn and the bone was semi-automatically segmented by delineating the bone boundary in each transverse slice using the software package Analyze v. 6.0 (Mayo Foundation, Rochester, MN). A binary image of the bone was formed by setting voxels outside the delineated boundary to zero and voxels within the boundary to one. The set of voxels on the interface between the bone and soft-tissue (S_{edgeCT}) was obtained by applying vertical and horizontal Sobel gradient operators to each slice of the binary volumes. The PDF was then defined as follows, where the function C calculates the cardinality (i.e., the number of elements) of the set:

$$p_{CT}(a,b) = \frac{C(\{\mathbf{x}|\mathbf{x} \in S_{edgeCT}, f_{CT1}(\mathbf{x}) = a, f_{CT2}(\mathbf{x}) = b\})}{C(\{\mathbf{x}|\mathbf{x} \in S_{CT}, f_{CT1}(\mathbf{x}) = a, f_{CT2}(\mathbf{x}) = b\})} \quad (2)$$

2.2.2. Converting the ultrasound slices

The first step in converting the US images is an artifact removal stage. Our method makes use of basic information about US image formation; in particular that most strong reflections and artifacts cause a loss of signal intensity along the direction of the beam. The artifact removal algorithm begins at the bottom of each column of pixels in the US image, and moves upwards towards the transducer face. The image is labelled as artifact until a threshold T_{art} value is reached, see Fig. 3(b). A value of $T_{art} = 40$ was used for all the experiments in this paper, the same value as was used previously for artifact removal in US liver images (Penney et al., 2004). A small region (3 mm) immediately adjacent to the transducer face is also labelled as artifact. This is to remove artifacts caused by imperfect acoustic coupling at the skin surface boundary. Image regions labelled as artifact were not used in any subsequent image processing.

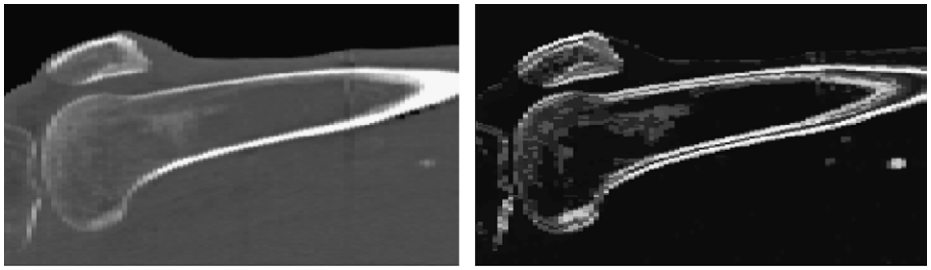

Fig. 2. Sagittal CT section through cadaver femur (a), and the corresponding section through the volume computed by assigning a probability to each voxel in the CT volume (b).

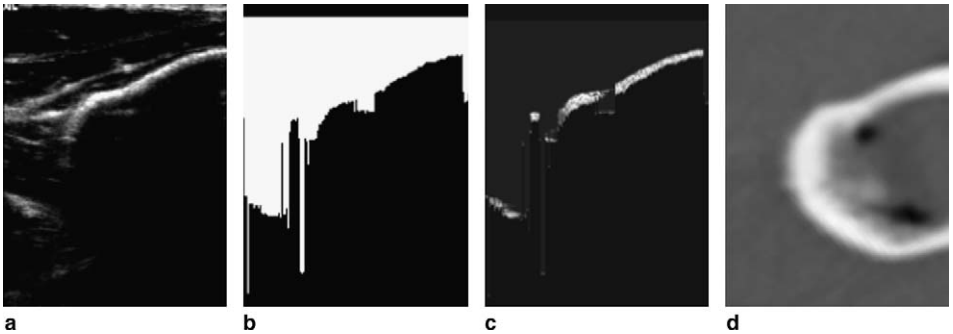

Fig. 3. Overview of the formation of the US probability images: (a) a sample US slice through the femur; (b) the mask produced by the artifact removal stage (only pixels in the white region are included for subsequent image processing and registration); (c) the final probability image; (d) the corresponding reformatted slice through the CT volume (calculated using the "gold standard" transformation) to show the position of bone boundaries. In (c) some non-bone-boundary pixels can be seen to have been allocated a high value in the probability image and vice-versa. However, we would like to reiterate that our aim is not to produce a perfect segmentation, but to produce probability images of sufficient quality to allow accurate and robust registration.

The PDF for the US slices, p_{US}, was also based on two image features, $\mathbf{F}(\mathbf{x})_{US} = (f_{US1}(\mathbf{x}), f_{US2}(\mathbf{x}))$. The first US feature, $f_{US1}(\mathbf{x})$, was the intensity of the US image $I_{US}(\mathbf{x})$. The second US feature, $f_{US2}(\mathbf{x})$, was the number of pixels between $\mathbf{x}$ and the bottom of the image which were not labelled as artifact. Both of these features are able to highlight bone edges due to the large change in acoustic impedance between soft-tissue and bone, which results in high US intensity values at the boundary (detected by $f_{US1}(\mathbf{x})$) with a bone shadow region of very low US intensities (detected by $f_{US2}(\mathbf{x})$). Similar information has been used previously (Amin et al., 2003; Daanen et al., 2004) to allow automatic segmentation (along with registration) of bone from US. Fig. 3 shows an example of an US probability image.

The PDF, p_{US}, was calculated using a set of US slices as training data. These images were initially processed to remove artifacts as explained above. Pixels designated as artifactual were not included in the calculation of p_{US}. The set of pixels, S_{edgeUS}, was then found in the following way: Each of the US training images were taken in turn and a number of points (approximately 15) were manually picked along the centre of the bone edge. The entire bone edge was defined by connecting adjacent points using straight lines. The closest pixel (in each vertical column of pixels) to the line defining the bone edge was found. The complete set of these pixels forms the set S_{edgeUS}. The pixel set, S_{US}, was comprised of all the non-artifact US pixels from the training data, and, as before, $S_{edgeUS} \subset S_{US}$. The probability density function, p_{US}, is then defined as follows, where again the function C calculates the cardinality of the set:

$$p_{US}(a,b) = \frac{C(\{\mathbf{x}|\mathbf{x} \in S_{edgeUS}, f_{US1}(\mathbf{x}) = a, f_{US2}(\mathbf{x}) = b\})}{C(\{\mathbf{x}|\mathbf{x} \in S_{US}, f_{US1}(\mathbf{x}) = a, f_{US2}(\mathbf{x}) = b\})} \tag{3}$$

2.3. Experiments

2.3.1. Data acquisition

Data acquisition for this study was carried out in the Institute of Anatomy, Ludwig-Maximilian University, Munich using 3 complete female cadavers preserved using the method of Thiel (2002). Fiducial markers were implanted into the femur and pelvis of each cadaver (4 in each femur and 5 in each hemi-pelvis). The markers were of a similar design to those used in the Acustar fiducial system (Maurer et al., 1997). They consisted of titanium bone screws to which two types of fiducial marker cap could be attached. During CT scanning an imaging marker cap was attached. This consisted of a chamber filled with CT contrast agent (Urografin 370, Schering AG, Berlin, Germany). Fig. 4(a) shows an axial slice through a femur and a fiducial marker with imaging cap attached. After CT scanning, physical location caps were attached. These caps have a 3 mm divot, the centre of which is coincident with the centroid of the fluid-filled chamber in the contrast filled markers, as illustrated in Fig. 4(b). These markers were used to calculate a "gold standard" registration between image and physical space. A single high-resolution spiral CT scan (Siemens SOMATOM Plus 5) was obtained of the whole pelvis and leg down to below the knee for each cadaver. All the CT images had a 2 mm slice thickness, with in-plane voxel dimensions varying slightly between scans from a minimum of 0.71×0.71 mm^2 to a maximum of 0.79×0.79 mm^2.

Each CT scan was divided into three smaller volumes, each of which contained an entire femur or pelvis plus a small amount of the surrounding anatomy. This was achieved by visually determining the extent of each bone in all three dimensions and then extracting the relevant region from the whole image. The effect of this process was twofold. Firstly, it reduced the amount of data presented to the registration algorithm, and so increased registration speed. Secondly, because the registration algorithm rotates images about their centroid, it placed the centre-of-rotation close to the centre of the bone to be registered. This assists the optimisation scheme enabling it to operate efficiently and to avoid local minima.

Prior to acquiring US images, a dynamic reference object (DRO) was rigidly attached to the bone. All measurements were recorded relative to the co-ordinate system of the DRO. The use of a DRO enabled the position of the cadaver to be changed during US acquisition so that

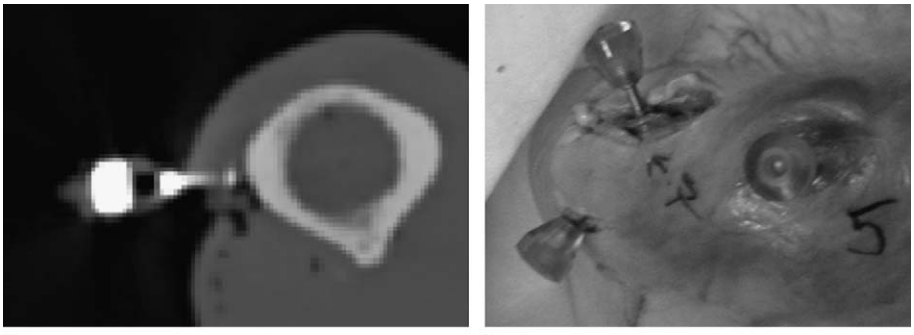

Fig. 4. Bone-implanted fiducial markers used to obtain the "gold standard" registration. (a) CT slice showing contrast-filled fiducial marker cap and titanium bone screw attached to a cadaveric femur. (b) Photograph of three fiducial markers with physical location caps attached to the pelvis. The 3 mm divot can clearly be seen in the fiducial labelled number 5.

images could be obtained in regions that would otherwise be inaccessible with the cadaver remaining in the supine position. The attachment of DROs (DiGioia et al., 1998; Nabeyama et al., 2004) or external fixators (Barger et al., 1998) is standard practice for most image-guided orthopaedic surgery systems.

US images were acquired using a Philips-ATL HDI-5000 scanner and a high frequency linear-array probe (L12-5, 5-12 MHz broadband transducer). Between 168 and 565 tracked US images were acquired on each bone. The aim was to scan using a variety of approaches so that as many anatomical features on each bone as possible were imaged. Typical positions where US images were acquired on femur and pelvis are shown in Fig. 5. Images were acquired using continuous scanning, and no attempt was made to filter out the better quality images from images which showed little or no bone surface. Examples of typical US images from a femur are shown in Fig. 6. The acquisition process took approximately 8 min, although we believe that this time could be substantially reduced if an optimal protocol was established.

2.3.2. Calculation and accuracy of "gold standard" transformations

The "gold standard" transformation, $\mathbf{T}_{GS}$, was obtained by manually picking the centre of the fiducial markers in the CT scan, and locating the markers in physical space using a tracked pointer. These paired positions were matched together using a closed-form solution to the "orthogonal Procrustes" problem (Arun et al., 1987). Bone-implanted fiducial markers have been shown previously to be able to produce "gold standard" transformations of sub-millimeter accuracy (Maurer et al., 1997). The "gold standard" transformation is compared to each registration transformation by calculating a root-mean-square (RMS) target registration error (TRE) (Fitzpatrick et al., 1998) in the following way: A region of interest (ROI) was defined as the set of voxels which contain the bone surface, S_{edgeCT}, as described in Section 2.2.1. The position of the centre of each voxel within the ROI ($\mathbf{x} \in S_{edgeCT}$) was transformed by the "gold standard" transformation ($\mathbf{T}_{GS}$) and by the registration transformation ($\mathbf{T}_{reg}$). The distance between the two resultant positions was then calculated for each voxel. The RMS value, d_{RMS}, of these distances, as given by Eq. (4) where N equals the number of voxels in set S_{edgeCT}, was computed as an overall measure of the target registration accuracy. The maximum TRE value, d_{max}, for each registration was also calculated, as given by Eq. (5). The accuracy values presented in Section 3 are the mean and standard deviation values of d_{RMS} and d_{max} calculated over a number of registrations.

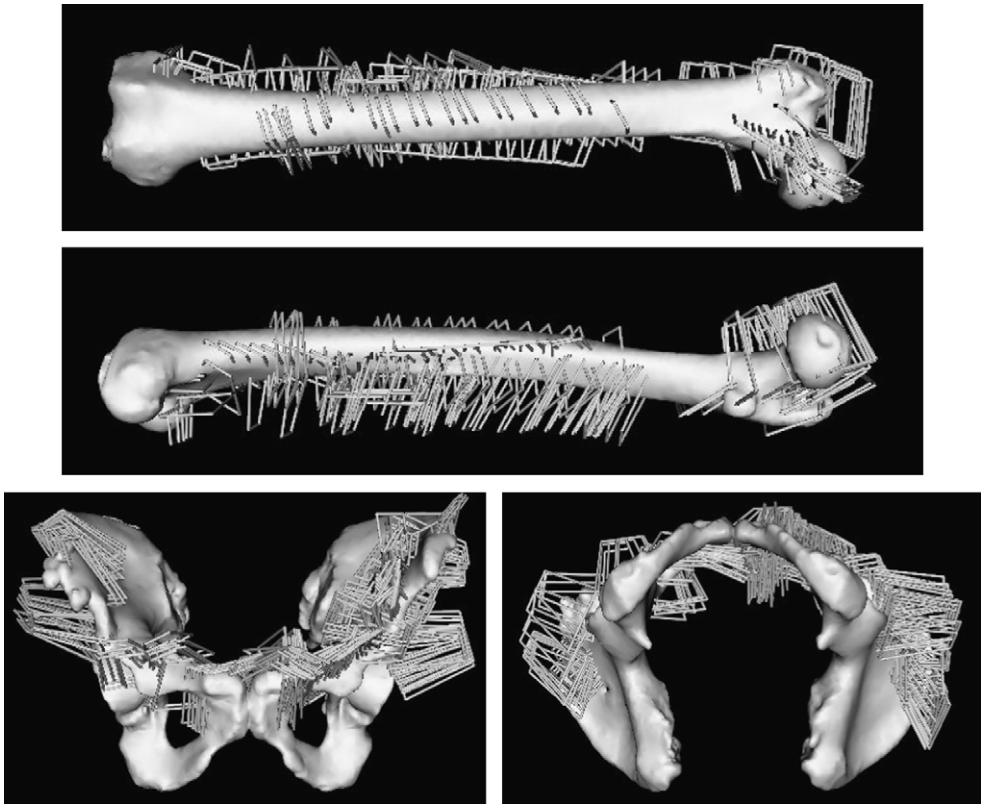

Fig. 5. Anterior (top) and lateral (middle) rendered views of right femur showing outlines of where US images were acquired for cadaver 1. Anterior (bottom left) and inferior (bottom right) rendered views of pelvis showing outlines of where US images were acquired for cadaver 2.

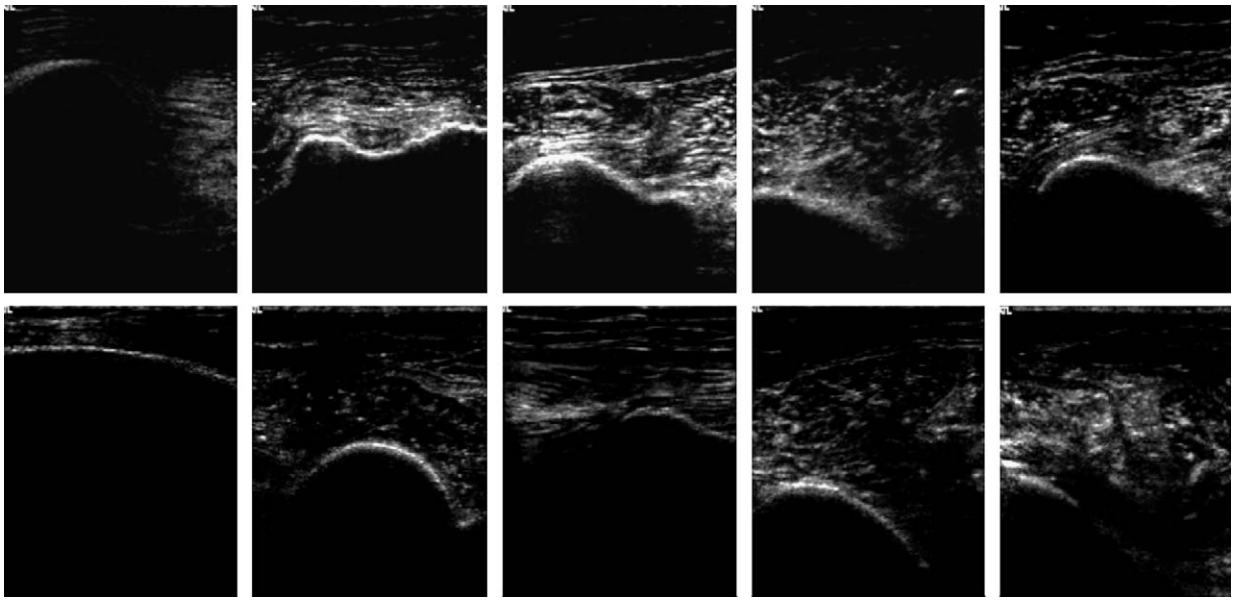

Fig. 6. Samples (every 50th US image) of the US images for cadaver 3 left femur.

$$d_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{\mathbf{x}\in S_{\mathrm{edgeCT}}} \|\mathbf{T}_{\mathrm{GS}}\mathbf{x} - \mathbf{T}_{\mathrm{reg}}\mathbf{x}\|^2} \quad (4)$$

$$d_{\max} = \max_{\mathbf{x}\in S_{\mathrm{edgeCT}}} \|\mathbf{T}_{\mathrm{GS}}\mathbf{x} - \mathbf{T}_{\mathrm{reg}}\mathbf{x}\| \quad (5)$$

Using the equations presented by Fitzpatrick et al. (1998, 2000), it is also possible to estimate the TRE of the "gold standard" registration. These equations define the relationships between the expected values of the fiducial localisation error (FLE), the fiducial residual error (FRE) and the TRE. The FLE is a measure of how accurately a fiducial marker can be localised. The FRE is defined as the residual error between fiducials after point-based registration, which is an easily measured value. Eq. (6) allows an estimate for the expected value of the FLE to be calculated from the weighted average of FRE values over M registrations, each of which used N_m fiducials,

$$\langle \mathrm{FLE}^2 \rangle \approx \frac{1}{M}\sum_{m}^{M} \frac{N_m}{N_m - 2} \times \mathrm{FRE}^2(m) \quad (6)$$

$$\langle \mathrm{TRE}^2 \rangle \approx \frac{1}{N}\left(1 + \frac{1}{3}\sum_{k=1}^{3}\frac{d_k^2}{f_k^2}\right)\langle \mathrm{FLE}^2 \rangle \quad (7)$$

Eq. (7) allows the expected value of TRE to be calculated from $\langle \mathrm{FLE}^2 \rangle$, the number of fiducials (N), the fiducial configuration and the position of the target point; d_k is the distance of the target point to the kth principal axis of the fiducial markers, and f_k is the RMS distance of the fiducials to the same axis.

The results for this study are given in Table 2. Problems arose with the fixation of some fiducials into the bone, especially in the pelvis, where the presence of osteoperotic trabecular bone in some of the cadavers meant that rigidity of some of the fiducial markers was compromised. These fiducials were identified and not included in the calculation of the "gold standard". The relatively large estimated RMS TRE for cadaver 2 right femur is believed to be due to poor fiducial marker positioning (i.e., the markers were approximately in a straight line).

Table 2
RMS TRE of "gold standard" registrations

Cadaver	Bone	No. fiducials	FRE (mm)	RMS TRE (mm)
1	L. Femur	3	0.06	0.96
	R. Femur	3	0.15	0.99
	Pelvis	5	0.44	0.68
2	L. Femur	4	0.47	0.78
	R. Femur	4	0.63	1.14
	Pelvis	8	0.86	0.49
3	L. Femur	4	0.39	0.67
	R. Femur	4	0.73	0.69
	Pelvis	5	0.86	0.55

2.3.3. Clinically realistic starting positions

One method for obtaining a starting estimate for the algorithm in a clinical situation is to pick corresponding skin positions, both in the CT scan and physically using a tracked pointer. As corresponding skin positions cannot be located to a high degree of accuracy, the starting estimates will almost certainly include relatively large errors. Nevertheless, they can provide a sufficiently accurate starting position for the registration algorithm.

We have simulated the above process to obtain 100 starting estimates for each bone. Four positions were chosen for each bone. For each pelvis, skin positions were picked in the CT scan above the anterior superior iliac spine and the anterior pubic ramus on both sides (i.e., left and right) of the cadaver. For the femur, three skin positions were picked: the skin next to the lateral and medial

epicondyle positions and the skin adjacent to the greater trochanter. The femur pivot point, which is calculated by rotating the femur with the DRO attached, was used as the fourth initial registration point. This pivot point is currently used for registration in some image-guided surgery systems (Nabeyama et al., 2004). The skin positions were then transformed into the DRO coordinate system using the "gold standard" transformation, to give a simulated set of physically located positions. For the pivot point, the physically located position of this point was used.

Errors in the CT-picked positions were simulated by randomly choosing a point from the set of voxels on the skin within 20 mm for each skin position, and from the set of voxels within 10 mm for the pivot position. Voxel positions were randomly chosen from each of these sets and registered to the simulated set of physically located positions to give a starting estimate. This was then repeated 100 times to produce 100 starting estimates.

Our maximum error of 20 mm on the location of skin landmarks should be a reasonable estimate, firstly, because intra-examiner pelvis and lower limb skin landmark precision has been reported in the literature with RMS values in the range of 6–21 mm (Della Croce et al., 1999); and, secondly because skin-affixed fiducial markers could be used. These markers would not increase the invasiveness of the procedure, and they have a reported accuracy of around 2 mm (Hajnal et al., 2001), albeit for neurosurgical applications where less skin movement is expected. Our justification for using a maximum error of 10 mm for the pivot point position comes from a separate cadaver experiment (5 cadavers and therefore 10 hip joints) where, using the same method and data acquisition system as used in this study, we calculated the mean error in pivot point position to be 4.8 mm, with a standard deviation of 2.1 mm.

2.3.4. Training data for PDF calculations

Sets of training data are required to calculate the PDFs, p_{US} and p_{CT}. For the experiments described in this paper all the CT images and between 100 and 141 US images from each cadaver were used as training data. Data used to produce the PDFs was kept separate from data used in the registration experiments in the following way: each of the cadavers were taken in turn and PDFs were produced using images from the other two cadavers. These PDFs were then were used to convert the images from the first cadaver. Therefore, for example, the PDFs used to convert the images for cadaver 1 were constructed only using images from cadavers 2 and 3, thus ensuring independence.

3. Results

Table 3 gives mean RMS TRE values (i.e., mean d_{RMS}, from Eq. (4)) averaged over all successful registrations. The mean d_{RMS} value was calculated at three stages during the registration process: after the low-resolution optimisation, after the high-resolution optimisation (high-res 1) and after the scaling parameter, s_y, was included in the optimisation (high-res 2). The factor by which s_y has changed, the failure rate and the number of US slices used for each bone registration are also given. The results show that the mean d_{RMS} value after registration was less than 2.3 mm for all the registrations except for the right femur of Cadaver 1. The results for this case also show the only instance when the mean d_{RMS} value increased when the optimisation included the y scaling parameter; when this parameter was held fixed, the mean d_{RMS} value was 1.7 mm. These registrations used the smallest number of US images (only 168) compared to over 500 for some of the other femur registrations, and we believe that this may have been a factor in the higher TRE calculated from this case.

Table 4 shows the mean of the maximum TRE values (i.e., mean d_{max}, see Eq. (5)) averaged over all successful registrations. Again, these values have been calculated at three positions during the optimisation process. The differences between the results in Tables 3 and 4 are caused predominantly by rotational registration errors. If the error between the "gold standard" and the registration transformations could be corrected solely by a translation (i.e., there is zero rotational or s_y scaling error) then both tables

Table 3
Registration accuracy quantified by the mean RMS TRE (i.e., mean d_{RMS}) for each bone averaged over all successful registrations

Cad	Bone	Mean (and s.d.) d_{RMS} value (mm)				s_y	No. fail	No. US slices
		Initial	Low-res	High-res 1	High-res 2			
1	L. Femur	7.3 (2.2)	3.9 (0.9)	2.2 (0.1)	1.9 (0.1)	1.004	0	226
	R. Femur	7.1 (2.3)	4.5 (0.2)	1.7 (0.1)	3.0 (0.1)	1.026	0	168
	Pelvis	11.0 (3.5)	2.2 (0.1)	1.8 (0.0)	1.8 (0.0)	1.001	0	200
2	L. Femur	7.7 (2.5)	4.8 (0.1)	1.1 (0.0)	0.7 (0.0)	0.992	0	247
	R. Femur	7.9 (2.6)	4.8 (0.1)	1.0 (0.0)	1.0 (0.0)	1.000	0	556
	Pelvis	15.1 (7.0)	3.1 (0.2)	1.7 (0.0)	1.1 (0.0)	0.975	0	317
3	L. Femur	7.1 (2.3)	2.7 (0.8)	1.9 (0.6)	1.5 (0.2)	0.965	0	516
	R. Femur	7.3 (2.5)	4.9 (0.3)	2.2 (0.1)	1.2 (0.0)	0.969	0	565
	Pelvis	11.6 (4.0)	4.9 (0.2)	3.0 (0.1)	2.3 (0.0)	0.973	1	331

The RMS TRE was calculated over the whole bone surface. A registration was defined to be a failure if its d_{RMS} value was more than double the mean d_{RMS} value calculated over all 100 registrations.

Table 4
Mean of the maximum TRE value on bone surface (i.e., mean d_{max} value) averaged over all successful registrations

Cad	Bone	Mean (and s.d.) Max TRE (mm)			
		Initial	Low-res	High-res 1	High-res 2
1	L. Femur	12.1 (3.4)	6.8 (2.5)	4.2 (0.7)	4.0 (0.5)
	R. Femur	11.4 (3.3)	9.1 (0.8)	3.4 (0.4)	6.0 (0.3)
	Pelvis	17.1 (5.5)	3.1 (0.3)	2.4 (0.1)	2.4 (0.1)
2	L. Femur	11.9 (3.7)	6.5 (0.1)	1.9 (0.0)	1.3 (0.1)
	R. Femur	11.8 (3.8)	7.2 (0.2)	2.0 (0.1)	2.0 (0.1)
	Pelvis	24.3 (11.5)	4.0 (0.4)	3.1 (0.1)	2.1 (0.1)
3	L. Femur	11.7 (3.8)	4.0 (1.0)	2.7 (1.0)	1.9 (0.2)
	R. Femur	11.4 (3.5)	7.8 (0.3)	4.2 (0.1)	2.4 (0.1)
	Pelvis	18.1 (6.6)	8.9 (0.4)	5.2 (0.1)	3.8 (0.1)

would contain identical values. Large mean d_{max} values have occurred for the registrations to both femurs for Cadaver 1 (4.0 and 6.0 mm) and for the pelvis registrations using Cadaver 3 (3.8 mm). Visual inspection of where these large TREs occurred in the femurs showed that for the right femur the errors were predominately at the distal end, around the condyles and the maximum error near the femoral head was 2.4 mm. For the left femur, errors as large as 3.8 mm were observed at the femoral head. Analysis of the position of the errors in the Cadaver 3 pelvis registrations showed average TRE values around one acetabulum (around 2.6 mm), though TRE values in the other acetabulum were as high as 3.3 mm.

4. Discussion

We have presented an automatic intensity-based algorithm which can register a CT scan to a set of tracked US images. From clinically plausible starting estimates, mean RMS TREs of 2.3 mm or less were achieved for eight out the nine bones registered, and for over half the bones, the mean RMS TRE was 1.5 mm or less. The algorithm was robust; no failures occurred when registering the femur images, and only one failure occurred for the pelvis registrations. The method is non-invasive and should be compatible with future minimally invasive orthopaedic procedures.

Accuracy requirements for CAOS systems are application specific, and the error measures used to validate such systems should also be designed relative to the particular application (Simon et al., 1995). In some procedures, a successful clinical outcome may be very sensitive to the angle at which the implant is positioned, whereas translational errors may be less critical. In such cases both rotational and translational errors should be quoted, preferably using notation and coordinate systems which are in current clinical use. Part of our future work plan is to investigate such error metrics specifically to validate total hip arthroplasty procedures. In this paper we have used a very general measure for registration accuracy (i.e., mean TRE). This can make it difficult to relate the error values to what is clinically acceptable. However, the use of a general error measure has the advantage that it should allow comparison with other registration methods, and our method can be seen to compare favourably with more conventional methods such as paired-point and surface matching TRE = 3.8 mm (van Hellemondt et al., 2002) and restricted surface matching TRE = 1.6 mm (Bächler et al., 2001).

Very basic image processing filters have been used to produce the features on which the probability images are calculated. These have proven sufficient to allow accurate and robust registrations. However, in the future other, more sophisticated filters may be used which are more accurate and robust in their ability to extract bone-to-soft-tissue edges from US (Daanen et al., 2004) and CT images.

Methods for US to MR and CT intensity-based registration have been proposed to match soft tissue structures of brain images (Roche et al., 2001), and images of the kidney (Leroy et al., 2004). These methods used the correlation ratio (CR) as a similarity measure, and did not require any training data. Our approach relies more heavily on the use of pre-processing image filters to enhance similar features within the images, and then a simple similarity measure is used; whereas, the use of CR (although some pre-processing is carried out) places more emphasis on the ability of the similarity measure to find the relationship between the intensity values in the two modalities.

The inclusion into the optimisation process of the US image scaling parameter, s_y, which is related directly to the average speed of sound, improved the TRE values for over half the registrations, for one dataset by more than 1 mm. Only in one case did the error increase, and this is believed to be due to insufficient numbers of US images.

The current time required by the registration algorithm on a 2.8 MHz Intel Pentium 4 processor was between 2 and 10.5 min. The predominant factor in processing time was the number of US slices used in the registration. A processing time above 2 min is probably too long for most procedures, however, we believe substantial decreases in registration time are possible as much more work can be carried out to optimise the code, and the computation time could be reduced substantially by computing the similarity measure using parallel processing.

No standard protocol was used for the acquisition of the US images. When acquiring US images of the femur it is much easier to acquire data along the femoral shaft rather than regions around the condyles and the proximal femur. However, it is images of the distal and proximal regions of the femur which provide the algorithm with features which are most useful to registration, particularly in relation to constraining the rotation of the femur around its long axis. Care must therefore be taken to ensure that images are acquired from these regions as well as along the femoral shaft. It is also very important to obtain US images from a variety of directions. If this is not the case then the registration problem is ill-posed as altering the s_y parameter will have an almost identical effect as a particular combination of translations. The mean depth of the bone in the US images was 17.7 mm (max of 58.3 mm). Larger depths can be expected in obese patients, which would make imaging the bone boundary more difficult, and may effect registration accuracy. Further work is required to establish a standard protocol for the US acquisition process.

The sampling of freehand 3D US data depends greatly on the skill of the operator. Therefore, it is possible – very likely in our experience – to acquire datasets with very uneven sampling. In the current algorithm this can bias the registration to the regions which have been scanned most frequently. These areas may be the regions which are easiest to scan, as opposed to the regions which are most useful to constrain the registration. Future improvements to the algorithm may be achieved by filtering the US data in order to achieve more regular spatial sampling.

5. Conclusions

We have presented an automatic method to register freehand 3D US images to a CT volume of a pelvis or femur. Our method is based on a strategy in which we preprocess images to produce probability images of corresponding features. This preprocessing step uses information gathered from a set of manually segmented training data. The algorithm also optimises the parameter which defines the speed of sound through tissue. Our method has been compared to a "gold standard" registration based on bone implanted fiducial markers. Registrations have been validated using three cadavers (6 femurs and 3 pelves). Results show that our method is accurate to a mean RMS TRE value of 2.3 mm or less (in all-but-one case), which should be sufficiently accurate for most total hip replacement procedures.

Acknowledgements

We thank the EPSRC (grant number GR/R03525/01) for funding this project. Our thanks also go to Prof. Dr. med. R. Putz of Anatomische Anstalt, Ludwig-Maximillian-Universitat (LMU), Munich for providing the cadavers, the staff in the LMU Radiology department for performing the CT scans, to Jens Krugman of BrainLab AG, Munich for his assistance in organising the experiments and to Alan Black in the Medical Physics workshop at Guy's Hospital for the production of fiducial markers and probe attachments.

References

Amin, D., Kanade, T., DiGioia III, A., Jaramaz, B., 2003. Ultrasound registration of the bone surface for surgical navigation. Comput. Aided Surg. 8, 1–16.

Amiot, L., Poulin, F., 2004. Computed tomography-based navigation for hip, knee, and spine surgery. Clin. Orthop. Relat. Res. 421, 77–86.

Arun, K., Huang, T., Blostein, S., 1987. Least-squares fitting of two 3-D point sets. IEEE Trans. Pattern Anal. Mach. Intell. 9, 698–700.

Ault, T., Siegel, M., 1994. Frameless patient registration using ultrasonic imaging. In: Proceedings of the First International Symposium on Medical Robotics and Computer-assisted surgery. 1, 74–82.

Bächler, R., Bunke, H., Nolte, L., 2001. Restricted surface matching – numerical optimization and technical evaluation. Comput. Aided Surg. 6 (3), 143–152.

Barbe, C., Troccaz, J., Mazier, B., Lavallée, S., 1993. Using 2.5D echography in computer assisted spine surgery. In: Engineering in Medicine and Biology Society (EMBS), Proceedings of the 15th Annual Conference of the IEEE. pp. 160–161.

Barger, W., Bauer, A., Borner, M., 1998. Primary and revision total hip replacement using the ROBODOC′ system. Clin. Orthop. Relat. Res. 354, 82–91.

Barratt, D., Davies, A., Hughes, A., Thom, S., Humphries, K., 2001. Accuracy of an electromagnetic three-dimensional ultrasound system for carotid artery imaging. Ultrasound Med. Biol. 27 (10), 1421–1425.

Berger, R., 2003. Total hip arthroplasty using the minimally invasive two-incision approach. Clin. Orthop. Relat. Res. 417, 232–241.

Brendel, B., Winter, S., Rick, A., Stockheim, M., Ermert, H., 2002. Registration of 3D CT and ultrasound datasets of the spine using bone structures. Comput. Aided Surg. 7, 146–155.

Brendel, B., Winter, S., Rick, A., Stockheim, M., Ermert, H., 2003. Bone registration with 3D CT and ultrasound data sets. In: Computer Assisted Radiology and Surgery 2003. Elsevier Sciences B.V., pp. 426–432.

Daanen, V., Tonetti, J., Troccaz, J., 2004. A fully automated method for the delineation of osseous interface in ultrasound images. In: Barillot, C., Haynor, D., Hellier, P. (Eds.), Medical Imaging Computing and Computer-Assisted Intervention – MICCAI '04. Springer, pp. 549–557.

Della Croce, U., Cappozzo, A., Kerrigan, D., 1999. Pelvis and lower limb anatomical landmark calibration precision and its propagation to bone geometry and joint angles. Med. Biol. Eng. Comput. 37, 155–161.

DiGioia III, A., Nolte, L.-P., 1998. The challenges for CAOS: What is the role of CAOS in orthopaedics?. Comput. Aided Surg. 7 (3) 127–128.

DiGioia III, A., Jaramaz, B., Blackwell, M., Simon, D., Morgan, F., Moody, J., Nikou, C., Colgan, B., Aston, C., Labarca, R., Kischell, E., Kanade, T., 1998. Image guided navigation system to measure intraoperatively acetabular implant alignment. Clin. Orthop. Relat. Res. 355, 8–22.

DiGioia III, A., Plakseychuk, A., Levison, T., Jaramaz, B., 2003. Mini-incision technique for total hip arthroplasty with navigation. J. Arthroplasty 18 (2), 123–128.

Fitzpatrick, J., Hill, D., Maurer Jr., C., 2000. Image registration. In: Handbook of Medical ImagingMedical Image Processing and Analysis, vol. 2. SPIE Press, p. 471 (Chapter 8).

Fitzpatrick, J.M., Hill, D.L.G., Shyr, Y., West, J.B., Studholme, C., Maurer Jr., C.R., 1998. Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain. IEEE Trans. Med. Imaging 17 (4), 571–585.

Guéziec, A., Kazanzides, P., Williamson, B., Taylor, R.H., 1998. Anatomy based registration of CT-scan and X-ray images for guiding a surgical robot. IEEE Trans. Med. Imaging 17 (5), 715–728.

Hajnal, J., Hill, D., Hawkes, D. (Eds.), 2001. Medical Image Registration. CRC Press, pp. 253–278 (Chapter 12).

Heger, S., Portheine, F., Ohnsorge, J., Schkommodau, E., Radermacher, K., 2005. User-interactive registration of bone with A-mode ultrasound. IEEE Eng. Med. Biol. 24 (2), 85–95.

Honl, M., Dierk, O., Gauck, C., Carrero, V., Lampe, F., Dries, S., Quante, M., Schwieger, K., Hille, E., Morlock, M., 2003. Comparison of robotic-assisted and manual implantation of a primary total hip replacement. J. Bone Joint Surg. (Am) 85-A (8).

Ionescu, G., Lavallée, S., Demongeot, J., 1999. Automated registration of ultrasound with CT images: application to computer assisted prostate radiotherapy and orthopaedics. In: Taylor, C., Colchester, A. (Eds.), Medical Imaging Computing and Computer-Assisted Intervention – MICCAI '99, Lecture Notes in Computer Science, 1679. Springer, pp. 768–777.

Jaramaz, B., Nikou, C., Cavalier, B., Blendea, S., Labarca, R., DiGioia III, A., 2003. Experimental validation of ultrasound registration of long bones. In: Langlotz, F., Davies, B., Bauer, A. (Eds.), 3rd Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery. Steinkopff Darmstadt, pp. 160–161.

Jaramaz, B., Nikou, C., Levison, T., DiGioia III, A., LaBaraca, R., 1999. Cup align: computer assisted postoperative radiographic measurement of acetabular components following total hip arthroplasty. In: Medical Imaging Computing and Computer-Assisted Intervention – MICCAI '99, LNCS 1679. Springer, pp. 876–882.

Kang, Y., Engelke, K., Kalender, W., 2003. A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. IEEE Trans. Med. Imaging 22 (5), 586–598.

Kowal, J., Hamdan, R., Heini, P., Nolte, L., Styner, M., 2003. Ultrasound based registration for minimally invasive interventions. In: Langlotz, F., Davies, B., Bauer, A. (Eds.), 3rd Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery. Steinkopff Darmstadt, pp. 196–197.

Lavallée, S., Szeliski, R., 1995. Recovering the position and orientation of free-form objects from image contours using 3-D distance maps. IEEE Trans. Pattern Anal. Mach. Intell 17 (4), 378–390.

Lavallée, S., Troccaz, J., Sautot, P., Mazier, B., Cinquin, P., Merloz, P., Chirossel, J., 1995. Computer-assisted spinal surgery using anatomy-based registration. In: Taylor, R.H., Lavallée, S., Burdea, G.C., Mosges, R. (Eds.), Registration for Computer Integrated Surgery: Methodology, State of the Art. MIT Press, Cambridge, MA, pp. 425–449.

Leroy, A., Mozer, P., Payan, Y., Troccaz, J., 2004. Rigid registration of freehand 3D ultrasound and CT-scan kidney images. In: Medical Imaging Computing and Computer-Assisted Intervention – MICCAI '04, LNCS 3216, pp. 837–844.

Livyatan, H., Yaniv, Z., Joskowicz, L., 2003. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Trans. Med. Imaging 22 (11), 1395–1406.

Maurer Jr., C., Fitzpatrick, J., Wang, M., Galloway Jr., R., Allen, G., 1997. Registration of head volume images using implantable fiducial markers. IEEE Trans. Med. Imaging 16, 447–462.

Muratore, D., Herring, J., Dawant, B., Galloway Jr., R., 2002. Three-dimensional image registration of phantom vertebrae for image-guided surgery: a preliminary study. Comput. Aided Surg. 7, 342–352.

Nabeyama, R., Matsuda, S., Miura, H., Mawatarim, T., Kawano, T., Iwamoto, Y., 2004. The accuracy of image-guided knee replacement based on computed tomography. J. Bone Joint Surg. (Br) 86-B, 366–371.

Nogler, M., Maurer, H., Wimmer, C., Gegenhuber, C., Bach, C., Krismer, M., 2001. Knee pain caused by a fiducial marker in the medial femoral condyle. Acta Orthop. Scand. 72 (5), 477–480.

Penney, G., Blackall, J., Hamady, M., Sabharwal, Y., Adam, A., Hawkes, D., 2004. Registration of freehand 3D ultrasound and magnetic resonance liver images. Med. Image Anal. 8, 81–91.

Press, W., Teukolsky, S., Vetterling, W., Flannery, B., 1992. Numerical Recipes in C, second ed. Cambridge University Press.

Roche, A., Pennec, X., Malandain, G., Ayache, N., 2001. Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. IEEE Trans. Med. Imaging 20 (10), 1038–1049.

Simon, D., O'Toole, R., Blackwell, M., Morgan, F., DiGioia, A., Kanade, T., 1995. Accuracy validation in image-guided orthopaedic surgery. In: Medical Robotics and Computer Assisted Surgery 1995, pp. 185–192.

Sonka, M., Hlavac, V., Boyle, R., 1998. Image Processing, Analysis, and Machine Vision, second ed. PWS Publishing.

Thiel, W., 2002. Ergänzung für die konservierung ganzer leichen nach w. thiel. Ann. Anat. 184, 267–270.

Tonetti, J., Carrat, L., Blendea, S., Merloz, P., Troccaz, J., Lavallée, S., Chirossel, J., 2001. Clinical results of percutaneous pelvis surgery. Computer assisted surgery using ultrasound compared to standard fluoroscopy. Comput. Aided Surg. 6 (4), 204–211.

van Hellemondt, G., de Kleuver, M., Kerckhaert, A., Anderson, P., Langlotz, F., Nolte, L., Pavlov, P., 2002. Computer-assisted pelvic surgery. Clin. Orthop. Relat. Res. 405, 287–293.

West, J., Fitzpatrick, J., Wang, M., Dawant, B., Maurer Jr., C., Kessler, R., Maciunas, R., 1999. Retrospective intermodality registration techniques for images of the head: surface-based verses volume-based. IEEE Trans. Med. Imaging 18 (2), 144–150.