ELSEVIER

PII: S0301-5629(01)00431-8

● *Original Contribution*

A FAST CALIBRATION METHOD FOR 3-D TRACKING OF ULTRASOUND IMAGES USING A SPATIAL LOCALIZER

Niko Pagoulatos,* David R. Haynor[†] and Yongmin Kim*

*Image Computing Systems Laboratory, Departments of Bioengineering and Electrical Engineering, University of Washington; and [†]Department of Radiology, University of Washington Medical Center, Seattle, Washington, USA

Abstract—We have developed a fast calibration method for computing the position and orientation of 2-D ultrasound (US) images in 3-D space where a position sensor is mounted on the US probe. This calibration is required in the fields of 3-D ultrasound and registration of ultrasound with other imaging modalities. Most of the existing calibration methods require a complex and tedious experimental procedure. Our method is simple and it is based on a custom-built phantom. Thirty N-fiducials (markers in the shape of the letter "N") embedded in the phantom provide the basis for our calibration procedure. We calibrated a 3.5-MHz sector phased-array probe with a magnetic position sensor, and we studied the accuracy and precision of our method. A typical calibration procedure requires approximately 2 min. We conclude that we can achieve accurate and precise calibration using a single US image, provided that a large number (approximately ten) of N-fiducials are captured within the US image, enabling a representative sampling of the imaging plane. (E-mail: ykim@u.washington.edu) © 2001 World Federation for Ultrasound in Medicine & Biology.

Key Words: Calibration, Position sensor, 3-D Ultrasound, Multimodality registration.

INTRODUCTION

Three-dimensional tracking of 2-D ultrasound (US) images involves the determination of the position and orientation of each 2-D image with respect to a given 3-D coordinate system. Ultrasound imaging applications that require such a tracking can be divided into two main categories: 3-D US and multimodality registration. In 3-D US, the reference coordinate system is an arbitrarily selected coordinate system associated with the reconstruction volume (Edwards et al. 1998; Prager et al. 1998); whereas in multimodality registration, the reference coordinate system derives from an externally acquired 3-D tomographic data set [e.g., magnetic resonance (MR) or X-ray computed tomography (CT)] (Comeau et al. 1998; Hata et al. 1997; Peria et al. 1995).

Accurate US image localization is critical for both applications. For example, one application for 3-D US is the accurate measurement of organ volumes, which is impossible without the accurate localization of the individual 2-D US images used to form the 3-D data set (Edwards et al. 1998; Prager et al. 1998). Similarly,

cross-modality registration of US with CT and MR data is a promising method for monitoring anatomy shifts and deformations during surgical procedures (Comeau et al. 1998). In this context, inaccurate tracking of US images could mislead the surgeon about the anatomy (Pagoulatos et al. 1999a).

Methods for US image tracking have primarily been based on position sensors. Different position sensors have been used (e.g., mechanical, acoustical, optical and magnetic) (Pagoulatos et al. 1998). No matter what type of sensor is used, a calibration procedure has to be performed before US image tracking is feasible. A number of different calibration procedures have been presented in the literature. Many require imaging the same point in space from different positions and orientations of the US probe (Detmer et al. 1994; Leotta et al. 1997). These procedures usually require a set of 30–40 images.

Depending on the phantom, identification of the target point can be tedious and time-consuming. We designed a phantom to significantly facilitate the identification of the target (Pagoulatos et al. 1998), but approximately 30 images still had to be acquired.

Prager et al. (1998) presented a method that was based on automatic identification of line segments rather than points. However, a complex protocol and a bulky

---

Address correspondence to: Dr. Y. Kim, University of Washington, Depts. of Bioengineering and Electrical Engineering, Box 352500, Seattle, WA 98195 USA. E-mail: ykim@u.washington.edu
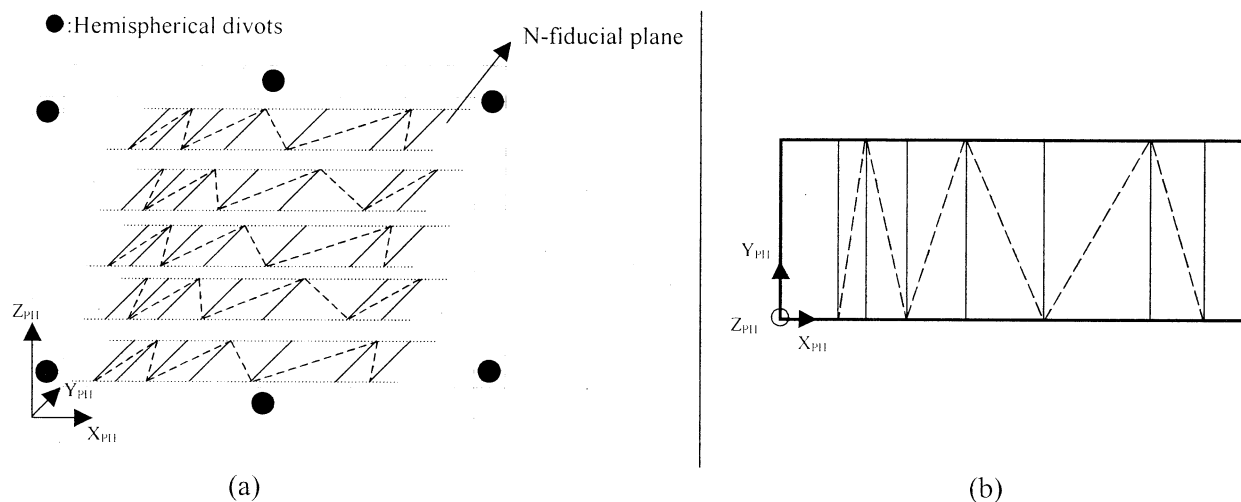
Fig. 1. (a) A drawing of the phantom with the five rows of N-fiducials. Each N-fiducial plane is parallel to the x-y plane of the phantom coordinate system. Each row contains six N-fiducials formed by the nylon strings. In the five rows shown, solid lines denote the parallel strings (perpendicular to the phantom walls) and dashed lines denote the oblique strings of the N-fiducials. Each oblique string corresponds to a single N-fiducial. The width and direction of the N-fiducials is different so that they can be easily distinguished in the US image. (b) A top view of a single N-fiducial plane showing the arrangement of the six N-fiducials within the same x-y plane.

mechanism were involved in their procedure. We believe that a simple calibration method would not only benefit researchers working in the field, but it could also facilitate the use of 3-D US and US-MR (or US-CT) registration in clinical sites.

In this paper, we present a fast and accurate calibration method that only requires freehand acquisition of a single US image of a custom-designed phantom. Preliminary results of this method have been presented elsewhere (Pagoulatos et al. 1999b). The method is applicable to any other position tracking system. We have also studied the effects of the US scanner's depth setting on the calibration results.

MATERIALS AND METHODS

Acquisition system

Our acquisition system consists of (1) an US scanner (Sonoline Elegra, Siemens Medical Systems) equipped with a 3.5-MHz sector phased-array probe (which we used throughout our study), (2) a DC magnetic position sensor (Flock of Birds, Ascension Technologies) and (3) a Pentium II 450-MHz computer equipped with a video digitizer (Matrox, Meteor-II) that receives the video output of the US scanner. Ultrasound images are digitized and stored in real time in the computer along with the corresponding position sensor measurements. The position sensor consists of a stationary transmitter to provide the reference coordinate system, and a probe-mounted receiver. During an US scan, the position and orientation of each US image with respect to the transmitter can be computed knowing two geometric transformations: (1) the rigid-body transformation between the receiver and the transmitter, which is the output of the position sensor, and (2) the transformation between the US image and the receiver coordinate systems.

Determination of this latter transformation is the objective of US probe calibration; it depends on the position and orientation of the receiver with respect to the US probe face and on the relationship between the image and the probe face.

Phantom

We designed a special calibration phantom. The phantom consists of a Plexiglas container (21 cm × 21 cm × 24 cm) with 18 hemispherical divots, each with a diameter of 6 mm, drilled in its outer walls, and a number of nylon strings, 1 mm in diameter, in its interior. The strings are immersed in distilled water and are placed so as to form a set of 30 fiducials, each with the shape of the letter "N," which we call N-fiducials. The N-fiducials are grouped in five planes, each containing six N-fiducials, which are parallel to the x-y plane of the phantom as shown in Fig. 1(a). Thus, each of the five z values is associated with six N-fiducials. A top-view of one plane with its six N-fiducials is shown in Fig. 1(b). In an US image oriented approximately parallel to the z axis of the phantom, an N-fiducial appears as a set of three points
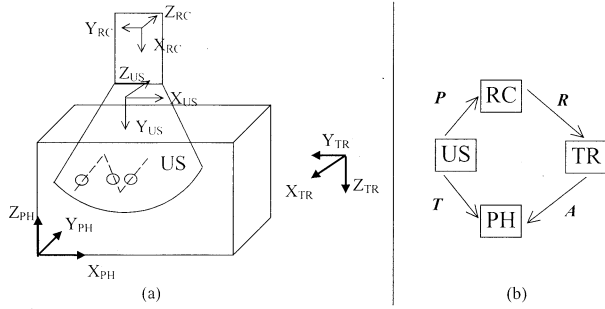
Fig. 2. (a) The geometric configuration of the four coordinate systems involved in the calibration procedure. The intersection of the US plane with one N-fiducial is illustrated as a set of three ellipses. (b) Coordinate transformations (as described in the text) between various coordinate systems. Note the two ways of converting US to phantom coordinates (i.e., T and $A \cdot R \cdot P$).

(ellipses), as illustrated in Fig. 2(a), which shows the phantom as imaged through its upper face. A large number of N-fiducials will be seen in any US image; the exact number depends on the imaging depth and field of view selected in the US scanner. At an imaging depth of 9 cm, 5–7 N-fiducials will typically appear in a single image, whereas at an imaging depth of 16 cm, a single US image can contain 10–15 N-fiducials (US images of the phantom are shown in the Results section). The 1-mm nylon strings produce very little shadowing, so the N-fiducials can be clearly seen at large US image depths. A photograph of the phantom is shown in Fig. 3(a).

Calibration method

Four coordinate systems are relevant to the description of our calibration method:

PH—The coordinate system associated with the phantom itself. The hemispherical divots and N-fiducials all have known locations with respect to PH.

US—The coordinate system associated with the US image. The origin and x and y axes are provided by the US scanner. The z axis is chosen so that a right-handed Euclidean coordinate system is formed.

TR—The coordinate system associated with the stationary transmitter.

RC—The coordinate system associated with the receiver, which is mounted on the US probe.

These coordinate systems are illustrated in Fig. 2(a), along with the various transformations between them, denoted by A, P, R and T illustrated in Fig. 2(b). From Fig. 2(b), we see

$$T = A \cdot R \cdot P \qquad (1)$$

where P is the transformation from the US image-based coordinate system to the receiver coordinate system, R is the transformation from the receiver to the transmitter coordinate system (the output of the position sensing system), A is the transformation from the transmitter to the phantom coordinate system and T is the transformation from the US image to the phantom coordinate system. Both A and T can be determined, given an adequate number of homologous points in the appropriate coordinate systems. The objective of the calibration is to determine P. Once this is done, the location, in the transmitter-based system, of a point in the US image, is calculated by setting $z_{US} = 0$ and applying the transformation $R \cdot P$.

From eqn. (1), P is given by
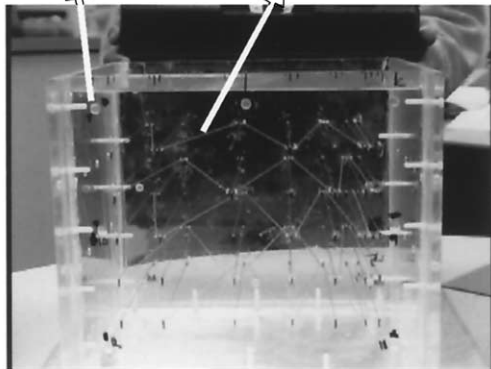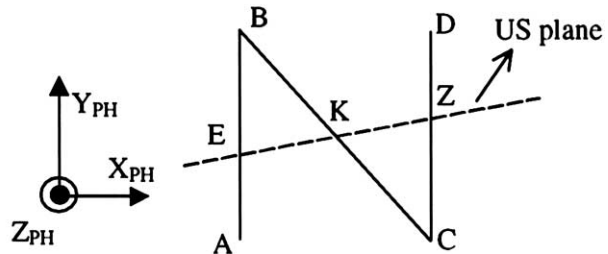
Fig. 3. (a) Photograph of the phantom showing the hemispherical divots on its walls and the nylon strings embedded in it. (b) A top view of an US plane intersecting a single N-fiducial. The dashed line represents the US imaging plane.

$$P = (A \cdot R)^{-1} \cdot T \qquad (2)$$

If we represent P in homogenous coordinates

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

the elements of the first, second and third columns of the matrix P are the projections of the x, y, and z axis unit vectors of the US image into the receiver coordinate system. For example, element P_{32} represents the projection of the y axis unit vector of the US image along the z axis of the receiver. Elements P_{14}, P_{24}, and P_{34} represent the x, y and z coordinate values of the US image origin in the receiver coordinate system.

   To calculate the rigid-body transformations A and T, we used a closed-form solution for matching pairs of homologous points given by Arun et al. (1987). This method has also been used in stereotactic neurosurgery (Maurer et al. 1997) for registration between various image and physical coordinate systems. In the following paragraphs, we describe the determination of the coordinates of the various homologous points in the appropriate coordinate systems.

   For the computation of transformation A, the corresponding points are provided by the centers of the hemispherical divots in the outer walls of the phantom. The coordinates of the hemispherical divots in the phantom space PH are known from the phantom design whereas their coordinates in the transmitter space TR are determined by sequentially touching them with the tip (a stainless steel bead with the same 6-mm diameter as the hemispherical divots) of a stylus attached to the receiver. A detailed description of this procedure has been presented elsewhere (Pagoulatos et al. 1999a) where it was used to register 3-D MR images with the transmitter coordinate system. Typically, six hemispherical divots are used. The best registration between phantom and transmitter coordinate systems is achieved by selecting the six divots approximately uniformly distributed around the scanned region of the phantom (Maurer et al. 1997).

   For the transformation T (US → PH), homologous points are provided by the N-fiducials. The extraction of homologous points in two coordinate systems using N-fiducials was introduced by Brown (1979) for stereotactic frames used in neurosurgery. In an US image approximately transverse to a fiducial, the "N" will be visualized as a set of three ellipses whose centers are collinear (Fig. 2a). Because each N-fiducial is parallel to the x-y plane of the phantom coordinate system, all three ellipse

centers must have the same z coordinate in the phantom space. For the middle ellipse, the x and y phantom space coordinates of its centers can be determined based on the similar triangles BEK and KZC in Fig. 3(b). The phantom coordinates of the points of the intersection of the US imaging plane with an N-fiducial can then be determined by

$$x_K^{PH} = x_B^{PH} + a \cdot (x_C^{PH} - x_B^{PH}),$$
$$y_K^{PH} = y_B^{PH} + a \cdot (y_C^{PH} - y_B^{PH}) \qquad (3)$$

The ratio $a = KE/EZ$ is measured from the locations of the three ellipses in the US image, and the coordinates of the vertices in the phantom space PH are known based on the phantom design. For each N-fiducial, therefore, the middle ellipse provides a pair of homologous points with known coordinates in the US ($z_{US} = 0$ for all points in the US image) and phantom coordinate systems.

   In summary, determination of the calibration matrix P with our method can be divided in the following steps: (1) mount the receiver on the stylus and touch six divots, (2) mount the receiver on the US transducer and acquire an US image and (3) manually identify the intersections of the N-fiducials in the US image. All three steps are facilitated by the graphical user interface we have developed in Microsoft Visual C++.

   Evaluation of any measurement process requires determination of that process's precision and accuracy. Precision is measured to assess the repeatability of our method (i.e., the variation of the calibration matrix when our calibration is performed for a number of times). On the other hand, accuracy is measured to assess how close the measured and true calibration matrices are. Accuracy of the calibration matrix is clinically relevant because it affects the geometric accuracy of the 3-D US data sets. In the following paragraphs, we describe our measurements of the precision and accuracy of the proposed calibration method.

Precision

   A straightforward manner to measure the precision of our method is to perform a number of calibrations and compute descriptive statistics (mean, standard deviation and range) for each calibration matrix element. Ideally, the calibration matrix would remain the same throughout the experiments. Leotta et al. (1997) reported the range of the elements P_{14}, P_{24}, P_{34} for a total of three calibrations. A more realistic measure of the precision of a calibration is obtained by mapping a fixed point p^* in the US image to the receiver coordinate system using the matrices derived from several applications of the calibration process, and studying the variability of those points in the receiver coordinate system (Prager et al.

1998). This measure integrates the variability of all the elements of the calibration matrix rather than the variability of each matrix element separately, but may depend on the image point chosen. In this paper, to measure the variability of the mapped locations of $p*$ in the receiver coordinate system, we have used the RMS point localization error (square root of the sum of squared RMS errors in each x, y and z direction of the receiver coordinate system), which we denote as RMS_{PREC}. Because the variability of the mapped points depends on the location of $p*$ within the US image, we have computed RMS_{PREC} for various different locations across the US image.

Accuracy

Direct measures of the accuracy of the calibration method are not as straightforward as measures of precision because the true calibration matrix is not known. Detmer et al. (1994) measured accuracy indirectly by imaging a fixed target point for different positions and orientations of the US probe with respect to the transmitter. The target point is identified in each of the acquired US images, and its coordinates are transformed to the transmitter coordinate system. The RMS point localization error, derived from the points mapped in the transmitter coordinate system, provides a measure of the relative accuracy of the calibration matrix. This measure is based on the principle that an error in one of the elements of the calibration matrix will result in scattered locations of the fixed target point in the transmitter coordinate system. Therefore, by measuring the precision with which we can localize a fixed imaging target in laboratory (transmitter) coordinates, we indirectly measure the accuracy of the calibration matrix. We denote this point localization error as RMS_{ACC}. A simple example to illustrate this principle is shown in Fig. 4(a). In this figure, it is assumed that the x coordinate of the origin of the US image in the receiver coordinate (refer to Fig. 2a) system (element P_{14} of the calibration matrix) contains an error e. Consequently, for each US probe position (A, B, C and D in Fig. 4a), the measured location (corresponding empty dashed circle shown in Fig. 4a) of the fixed target point in the transmitter coordinate system will differ from the true location of the target (black-filled circle) by an offset of e; therefore, a cloud of points is derived in the transmitter system instead of a single point, as would be the case if there was no error in the calibration matrix. From Fig. 4(a), we can see that the RMS error of the cloud of points in the transmitter coordinate system depends not only on the error e, but also on the positions and orientations (A, B, C, D, . . .) of the US probe used to image the target point. Therefore, when reporting the RMS error using this method, the scanning pattern applied has to be reported as well
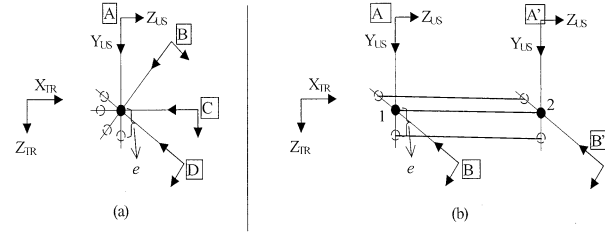

Fig. 4. (a) The principle of measuring the calibration matrix accuracy by measuring a single point target from different positions of the US probe is illustrated. See text for further details. (b) Two target points are imaged with different positions of the US probe and their distance is computed and compared to their true distance. If points 1 and 2 are measured for US probe positions A and A′, respectively, the measured and true distances will be the same, thus error e will be cancelled. To eliminate this effect, multiple probe orientations must be used. For simplicity, both (a) and (b) above are shown in two dimensions and the receiver coordinate system is omitted.

(Leotta et al. 1997; Prager et al. 1998). The maximum extent of the error may not be identified, given that physical considerations limit the possible positions and orientations of the US probe.

Another approach to indirect measurement of the accuracy of the calibration matrix is to measure the distance D_M between two given points by imaging them with the US probe and mapping their coordinates to the transmitter coordinate system. In this case, the true distance D_T between the two points can be found by some other means of measurement (e.g., micrometer), and therefore the descriptive statistics of the difference between the true distance and the measured value (e.g., D_M-D_T) can be used as a measure of the accuracy of the calibration matrix. However, there still exists a dependence of the measured error on the positions and orientations of the US probe selected for imaging the given points. In fact, as it is described in Fig. 4(b) and the corresponding caption, when measuring distances, errors in the calibration matrix can be cancelled out (Detmer et al. 1994). Therefore, it is important to image each point under consideration for a number of different positions and orientations of the US probe to minimize potential cancellation of errors. The method described in this paragraph is essentially a variation of the method described in the previous paragraph, where the true distance D_T of the points under consideration is some nonzero distance (for the method of the previous paragraph, $D_T = 0$).

In addition to the two aforementioned methods for indirectly assessing the accuracy of the calibration matrix, we have employed a third one, which uses the N-fiducials embedded in our phantom. Once the calibration matrix is determined and the phantom is registered
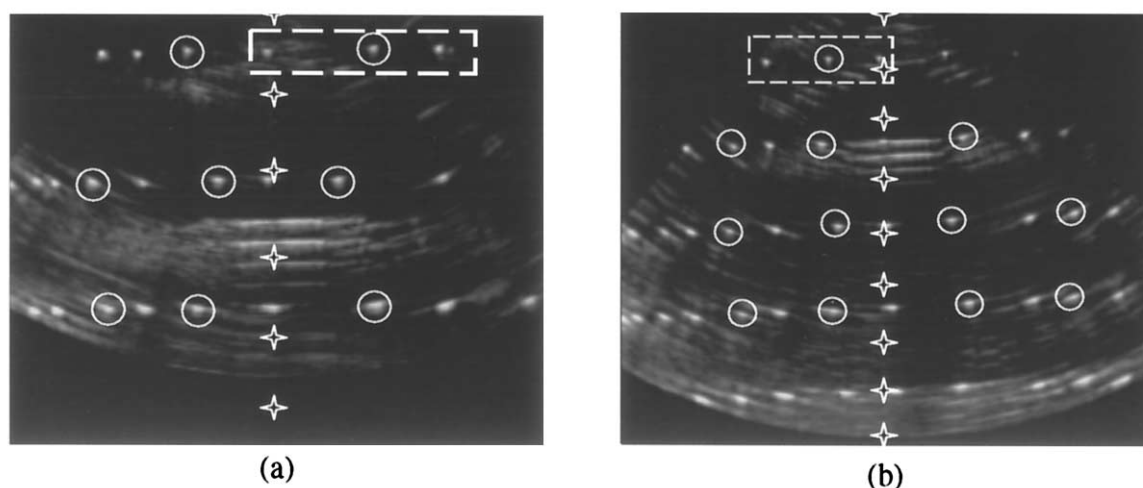
Fig. 5. Typical images of our phantom at imaging depths of (a) 9 cm, and (b) 16 cm. The dashed rectangles show an example of the three ellipses produced from the intersection of the US image with one N-fiducial. The circles show the middle ellipses of various N-fiducials used for calibration. Starred points illustrate the points in the US plane used to compute precision.

with the transmitter coordinate system, we can image the phantom again and determine the coordinates of the middle ellipse of each N-fiducial in the phantom-based system using either of the two methods depicted in Fig. 2(b). By measuring the distance Δ between these two points (derived from the two separate methods) in the phantom coordinate system, we can indirectly compute the accuracy of the calibration matrix. In this case, just one image is enough to estimate the errors in the calibration matrix because based on the N-fiducials we can compute the relation of the US image with respect to the phantom (T) without using the calibration matrix P. Therefore (because only one US image is enough to measure Δ), one can see that the error Δ does not depend on the range of positions and orientations of the US probe when imaging the target, but does depend on the location of the N-fiducial on the US image. Note that both RMS_{ACC} and D_M-D_T depend on the range of the position and orientation of the US probe as it is illustrated in Fig. 4(a) and Fig. 4(b), respectively. The limitations of this method are accuracy of the phantom's construction and of the identification of the center ellipses produced by the N-fiducials in the US image.

In the following sections, we present the results of our experiments evaluating the proposed calibration method. Our goals are to: (1) determine the precision and accuracy of our method, using the metrics previously described, and compare with the results previously reported by other researchers and (2) study how the imaging depth setting affects our calibration method. To experimentally compute the precision and accuracy measures (e.g., RMS_{PREC}, RMS_{ACC}, D_M-D_T and Δ), we

essentially need to obtain an estimate of these measures based on a sample population of calibration matrices. In this paper, we used a sample of 20 calibration matrices as was also used by Prager et al. (1998).

RESULTS

Precision

After registering the phantom with the transmitter coordinate system (using the hemispherical divots in the phantom walls), 20 US images of the phantom were acquired for each imaging depth of 9 and 16 cm. At each imaging depth, we imaged sections of the phantom that would provide a large number (typically 5–7 at 9-cm depth and 10–15 at 16-cm depth) of N-fiducials throughout the extent of the US image as shown in Fig. 5.

For each acquired image, the US plane was approximately perpendicular to the nylon strings that form the N-fiducials (e.g., parallel to the y-z plane of the phantom coordinate system as illustrated in Fig. 2a); this configuration allowed for a clear visualization of the ellipses produced from the intersection of the N-fiducials with the US plane. To further improve the visualization of the ellipses, we used multiple focal zones (five zones at 9-cm depth with a frame rate of five frames/s and seven zones at 16-cm depth with a frame rate of three frames/s), resulting in the best possible lateral resolution. Ultrasound images were digitized statically, thus there was no lag between image and position sensor. For each image, the corresponding N-fiducials were identified and the rigid-body calibration matrix was computed based on eqn (2). In our calculations, we corrected for the differ-

Table 1. Statistics of the rigid-body calibration matrix elements for two different imaging depth settings based on 20 repeated measurements

	Mean	SD	Range (max-min)
P_9			
P_{11}	0.074	0.004	0.015
P_{21}	−0.995	0.002	0.007
P_{31}	−0.060	0.034	0.120
P_{12}	0.996	0.001	0.004
P_{22}	0.076	0.003	0.011
P_{32}	−0.023	0.039	0.138
P_{14}	4.988	0.015	0.051
P_{24}	0.353	0.036	0.112
P_{34}	3.267	0.218	0.645
P_{16}			
P_{11}	0.065	0.002	0.009
P_{21}	−0.998	0.001	0.003
P_{31}	0.002	0.025	0.118
P_{12}	0.997	0.001	0.005
P_{22}	0.065	0.002	0.006
P_{32}	−0.050	0.019	0.092
P_{14}	4.960	0.013	0.049
P_{24}	0.311	0.029	0.104
P_{34}	3.277	0.121	0.426

P_9 represents the calibration matrix derived at 9-cm depth and P_{16} represents the calibration matrix derived at 16-cm depth. The third column of the calibration elements is not shown because it is simply the cross-product of the first two columns, and it is not directly computed from the calibration method. All matrix elements are dimensionless except the elements P_{14}, P_{24}, P_{34}, which are expressed in cm.

ence between the speed of sound in distilled water (1498 m/s) where our experiments were performed and the value assumed in the US scanner (1540 m/s).

The descriptive statistics of each calibration matrix element for the two imaging depths are shown in Table 1. The mean values of the transformation *P* elements shown in Table 1 express the geometric relation between the receiver and the probe face, whereas the SD and the range measure the variation of the transformation elements.

The smaller the SD and the range, the less the variation and the better the precision of the calibration method. We computed the precision RMS_{PREC} of the calibration matrices derived at each of the two imaging depth settings for the starred points in the US plane as shown in Fig. 5; Fig. 6 shows the results.

Accuracy

The relative accuracy of the calibration matrix was assessed by imaging a fixed target point in the transmitter coordinate system for different positions and orientations of the US probe. The target point was the intersection of two nylon wires immersed in distilled water. The target was always imaged from the top part of the phantom, and we performed two different scanning protocols (Leotta et
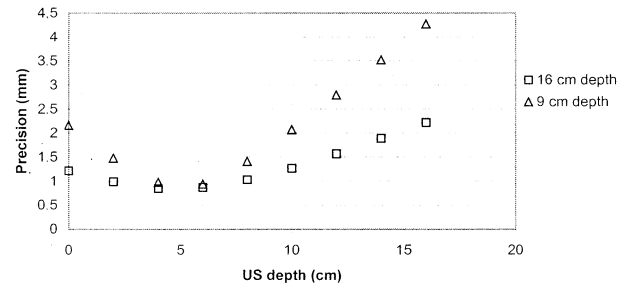

Fig. 6. Precision in the US plane based on experiments, using calibrations performed at two different (9-cm and 16-cm) imaging depths. All the points have the same x US coordinate value.

al. 1997): (1) in scanning protocol I (SP_I), the target was imaged through a single acoustic window and a single orientation of the US probe with respect to the transmitter and (2) in scanning protocol II (SP_{II}), the target was imaged through various acoustic windows and orientations of the US probe. Figure 7 illustrates a top view of the phantom with the cross-wire and the transmitter coordinate system; for the sake of description, the phantom is separated in four partitions. During SP_I, the US probe was placed in partition 1 and was moved along the x transmitter direction, maintaining approximately the same orientation of the US probe and the transmitter; whereas during SP_{II}, the probe was placed in all four partitions, resulting in multiple orientations of the US probe with respect to the transmitter and in large movements of the probe in both x and y transmitter directions. These protocols resulted in a wide variety of locations of the targets within the US image. For SP_I, the target depth in the US image ranged from approximately 6 to 7 cm whereas for SP_{II} it ranged from 5 to 9 cm.

The target was imaged following the two scanning protocols for each of the 9-cm and 16-cm imaging depth settings, resulting in four data sets (two at each imaging
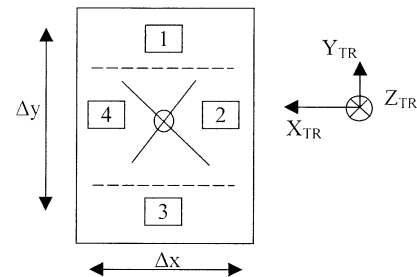

Fig. 7. The phantom containing the point target (circled) and the four partitions used to describe the range of motion of the US probe. The phantom is oriented as shown with respect to the transmitter coordinate system.

Table 2. Accuracy of P_9 and P_{16} calibration matrices using three different accuracy measures

| | | RMS$_{ACC}$ | | | | | | | |
| | | SP$_I$ | | SP$_{II}$ | | D$_M$-D$_T$ | | Δ | |
		P_9	P_{16}	P_9	P_{16}	P_9	P_{16}	P_9	P_{16}
Depth	Mean	1.54	0.71	2.89	2.41	0.23	−0.10	2.48	2.30
9 cm	SD	0.69	0.26	0.56	0.33	2.89	1.45	1.81	1.20
	Min	0.53	0.51	1.58	1.91	−8.73	−5.34	0.24	0.49
	Max	2.90	1.45	3.72	3.01	8.98	5.79	9.70	6.73
Depth	Mean	2.13	1.19	3.14	2.46	0.28	−0.01	3.20	2.76
16 cm	SD	0.64	0.30	0.26	0.20	2.87	1.29	2.14	1.18
	Min	1.13	0.76	2.71	2.22	−10.39	−6.00	0.11	0.33
	Max	3.44	2.10	3.62	3.10	10.68	5.00	12.38	8.32

All numbers are in mm.

depth setting). SP$_I$ consisted of 30 images of the target, whereas SP$_{II}$ consisted of 50 images, resulting in two sets of 30 images and two sets of 50 images. For each of the 40 calibration matrices (20 derived from each of the two imaging depths), RMS$_{ACC}$ was computed for each of the four data sets. The descriptive statistics of RMS$_{ACC}$ for all the combinations of (1) scanning protocols, (2) imaging depth used for calibration and (3) imaging depth used for target are shown in Table 2. The smaller the mean value of the RMS$_{ACC}$, the more accurate the localization of the single point, thus the more accurate the calibration matrix. The SD, min and max can be used to characterize the distribution of the 20 RMS$_{ACC}$ values. Note that 20 RMS$_{ACC}$ values correspond to (1) a single scanning protocol, (2) a single imaging depth for calibration and (3) a single imaging depth for imaging the target.

Subsequently, we computed the difference between true and measured distances, D$_M$-D$_T$, by imaging a grid of 1.5-mm diameter stainless steel beads immersed in distilled water. Using six beads resulted in 15 (6 × 5/2) interbead distances, varying from 1 to 9 cm. For each bead, eight images were acquired following a scanning pattern similar to SP$_{II}$ (beads appeared in depths of 6–8 cm in the US images). Using each of the 40 calibration matrices, we transformed the coordinates of the beads from the US to the transmitter coordinate system and computed their distances. The true distances (measured with a micrometer) were subtracted from the corresponding measured distances. Therefore, we obtained 960 measurements (8 × 8 measurements per distance because for each bead eight images were acquired, for 15 distances) for each calibration matrix and 19,200 (20 × 960) measurements for each set of 20 calibration matrices derived at 9 or 16 cm. The beads were imaged using both 9- and 16-cm depth setting in the US scanner. The descriptive statistics of the difference D$_M$-D$_T$ for each imaging depth used for calibration and imaging of the beads are shown in Table 2. The smaller the difference D$_M$-D$_T$, the more accurate our calibration matrix (ideally, D$_M$-D$_T$ should be zero). The mean, SD, min and max values of D$_M$-D$_T$ can be used to characterize how the values D$_M$-D$_T$ are distributed around zero.

Finally, we computed the third metric for estimating the accuracy of the calibration matrix (*i.e.*, Δ). To measure the accuracy of the computed calibration matrices, the phantom was registered again to the transmitter coordinate system, and a set of ten images for each imaging depth of 9 cm and 16 cm were acquired (the geometric configuration and US scanner settings were the same as with the images acquired for calibration). For each N-fiducial that appeared in various locations of each US image, the coordinates of the middle ellipse were transformed from the US to the phantom coordinate system using the two independent methods illustrated in Fig. 2(b), and the distance Δ between the two corresponding points was computed. The distance of the N-fiducials from the US probe face was larger when the imaging depth setting was set at 16 cm because we could image deeper sections of our phantom (compared to the imaging depth setting of 9 cm). A total of 40 points (approximately four points for each of ten images) for each depth were used. The points selected were located around the center portion of each US image where usually the anatomy of interest lies. This was repeated for all 20 calibration matrices derived at each imaging depth resulting in 800 (20 × 40) points per imaging depth. The descriptive statistics of the distance Δ for each imaging depth (800 measurements of Δ per depth) are shown in Table 2. The smaller the distance Δ, the closer the agreement between the coordinates derived using the two independent methods illustrated in Fig. 2(b), thus the more accurate the calibration matrix. The mean, SD, min and max can be used to characterize the distribution of Δ.

DISCUSSION

The determination of the calibration matrix can be seen as a plane-fitting procedure (Leotta et al. 1997) in which the control points are provided from various imaging targets (the middle ellipses of the N-fiducials in our method). Therefore, we expect that the US plane will be well defined around the neighborhood of the control points (middle ellipses of the N-fiducials). Our results support this argument because, for large imaging depths, the lack of control points for the P_9 calibration matrices (rigid-body calibration matrices derived at 9-cm imaging depth setting) results in a decreased precision compared to the P_{16} matrices (rigid-body calibration matrices derived at 16-cm imaging depth setting) as seen in Fig. 6 and Table 1. In Fig. 6, note how the difference in precision is larger when measured for points with larger depths. In addition, the matrices P_{16} are more accurate than matrices P_9 as can be concluded from the computed accuracy measures RMS_{ACC}, D_M-D_T and Δ shown in Table 2. Note that all three (RMS_{ACC}, D_M-D_T and Δ) corresponding accuracy measures for P_9 and P_{16} matrices are statistically different at a significance level of 5%. We conclude that the calibration matrices derived at 16-cm imaging depth are more precise, accurate and preferable, particularly for deeper locations in the US image, than those derived at 9-cm imaging depth. Similar conclusions were reported by Leotta et al. (1997). From a clinical point of view, these results suggest that the calibration protocol should include fiducials at least as deep as the deepest point possible in subsequent clinical use.

The difference in the accuracy and precision between the P_9 and P_{16} calibration matrices is due to the fact that for the given phantom presented in this paper, more control points were included in the images of 16-cm imaging depth than the images of 9-cm imaging depth. However, accurate and precise calibrations can still be achieved using small imaging depth settings (e.g., when high frequency probes are used), provided that a large number of N-fiducials are included in one US image. This can be achieved by either constructing a separate phantom or providing multiple sections with different density of N-fiducials in a single phantom.

When calibrating using the previously reported methods, a major issue is whether or not the motion of the US probe, during the acquisition of the US calibration images, covers all the degrees of freedom for proper determination of the calibration matrix elements (Prager et al. 1998). In addition, the features used for calibration (line or point targets) have to be centered with respect to the thickness (elevational resolution) of the US image. This centering is very important, especially in the near-field zone where the US beam thickness is large. There-

fore, covering all the adequate degrees of freedom with the US probe motion, centering the imaged targets to the US beam, acquiring a large (30–40) number of images, and the various practical limitations imposed in the movement of the probe when imaging real phantoms result in complex calibration experimental protocols. We have addressed these problems by developing a new and simple calibration procedure, which requires approximately 2 min to be completed.

The fundamental difference between the new calibration method we have presented in this paper and the calibration methods previously reported is that the coordinates of certain features (middle ellipses of the N-fiducials in the US image) can be found with respect to the reference transmitter-based coordinate system before calibration of the US probe with the position sensor has been performed. This results in the establishment of a set of homologous points with known coordinates in both the US image and the reference coordinate system, enabling the direct computation of the calibration matrix. Because of the phantom construction, a large number of homologous points throughout the US plane are acquired with the acquisition of only a single US image provided that a sufficiently large imaging depth setting is used. The large number (10–15 at 16-cm depth) of homologous points provides a representative sampling of the US plane, thus a precise and accurate calibration matrix is obtained. Because we are imaging straight lines that are approximately perpendicular to the scanning plane, centering with respect to the US beam thickness is not an issue. The center of the ellipses provided by the intersection of the US plane with the lines should always represent the center of the US plane. Therefore, the acquisition of one US image and the easy identification (where no centering has to be performed) of the features (N-fiducials) used for calibration simplifies the experimental protocol and leads to a fast and simple calibration procedure. Finally, our phantom can be used for indirectly validating the accuracy of the calibration matrices through the Δ measure, which does not depend on the positions and orientations of the US probe used to image the targets (refer to previous discussion in accuracy subsection of the material and methods section). Those phantoms presented in the literature could only provide measurements of RMS_{ACC} and D_M-D_T, which, as we previously explained, both depend on the positions and orientations of the US probe used to image the targets.

To understand the sources of error of our calibration procedure, we need to consider eqn (2) and the coordinate transformations involved in the computation of the calibration matrix P. Based on the methods described for providing each transformation in eqn (2), the sources of error of our calibration are:

- Phantom construction because it affects both the localization of the hemispherical divots and N-fiducials in the phantom coordinate space.
- Identification of the N-fiducial ellipses in the US images because it affects the coordinates of the middle ellipse of each N-fiducial in the US and phantom coordinate space.
- Calibration of the stylus because it affects the determination of the hemispherical divot coordinates in the position sensor reference coordinate system.
- Position sensor system because it affects the calibration of the stylus, the accurate position and orientation of the US plane (transformation R in eqn 2) used for the calibration and the coordinates of the hemispherical divots in the position sensor reference coordinate system.

As we previously mentioned, the accuracy measures of the calibration matrix are the ones that are clinically relevant. For example, measures RMS_{ACC} and D_M-D_T can be considered as the minimum expected error found in 3-D US reconstructions using similar scanning patterns to SP_I and SP_{II} (Leotta et al. 1997; Prager et al. 1998). The errors are expected to increase in clinical scans where tissue motion and US beam refraction will cause additional misalignment of the acquired US images (Prager et al. 1998). In this case, image-based methods can be used to improve the reconstruction accuracy of the acquired US images (Pagoulatos et al. 2000).

In this work, we were not concerned with the specific properties of the DC magnetic position sensor. The calibration method we present here can be also used with other position sensors (e.g., optical, acoustical, articulated and robotic arms). We believe that calibration of US probes tracked using optical position sensors and robotic arms will become more important in the future where US imaging is likely to be used increasingly in image-guided surgical procedures. The issues involved with using a magnetic position sensor in US probe calibration were extensively presented by Leotta et al. (1997), and we also presented similar results (Pagoulatos et al. 1998) where we used a gel-based sealed version of the point target calibration phantom presented by Leotta et al. (1997). The systematic errors associated with the Flock of Birds magnetic position sensor (Leotta et al. 1997) may affect the accuracy of our calibration when used in clinical scanning protocols involving multiple positions and orientations of the receiver with respect to the transmitter. This is due to the fact that only one configuration between receiver and transmitter is included in our calibration matrix, thus systematic errors cannot be cancelled out through averaging. Our method can be readily modified to include more than one US image so that multiple receiver–transmitter configurations are accounted for the computation of the calibration matrix. In this study, we minimized the effect of systematic errors by operating the magnetic sensor at its optimum range (receiver–transmitter distance at approximately 30–50 cm) (Leotta et al. 1997).

Although direct comparison of calibration results is difficult due to various differences in the methods and the systems used, our accuracy and precision results for P_{16} are comparable to those previously reported from other researchers (Detmer et al. 1994; Leotta et al. 1997; Prager et al. 1998). It is worthwhile noting that our method requires knowledge of the pixel dimensions of the US scanner and the speed of sound in the medium where calibration is performed. Both of these variables are not required by the method presented by Prager et al. (1998). We have also presented a calibration method where an affine transformation (instead of a rigid-body) was used to represent the calibration matrix, and thus knowledge of the US pixel dimension and the speed of sound in the calibration medium was not required (Pagoulatos et al. 1998).

The current version of our phantom has dimensions of $21 \times 21 \times 24$ cm. However, we can employ the same method of calibration and validation as described in this paper by using a much thinner phantom (e.g., $21 \times 21 \times 6$ cm). That would make the phantom less bulky and heavy, which makes it more convenient to use in a clinical environment. Furthermore, we can use a gel-based mixture to fill the phantom to provide better images of the N-fiducials. In this case, a machine vision method for automatic extraction and identification of the ellipses (rather than the manual extraction that we currently use) would further simplify the use of the presented calibration method.

SUMMARY

We have developed a fast and accurate calibration method for determining the coordinate transformation between the US image and the sensor mounted on the US probe. The method is based on a custom-designed phantom that enables the acquisition of only one US image for successful calibration. Typically, a calibration procedure would require approximately 2 min to be accomplished. Based on this method, we studied the effect of US imaging depth setting in the calibration procedure. We performed calibrations at two different imaging depths of 9 and 16 cm. We concluded that when using our new method, 16-cm depth provides both more precise and accurate calibrations. This is due to the fact that, at larger imaging depth settings, we have a larger US field of view and thus more features can be contained in the US image; therefore, a more representative sampling of the US plane is possible. Clinical utility can be en-

hanced with the use of thinner phantoms and automated methods for identifying the fiducials in the calibration image.

REFERENCES

Arun KS, Huang TS, Blostein SD. Least-squares fitting of two 3-D point sets. IEEE Trans Pattern Anal Mach Intel 1987;9:698–700.

Brown RA. A stereotactic head frame for use with CT body scanners. Invest Radiol 1979;14:300–304.

Comeau RM, Fenster A, Peters TM. Intraoperative ultrasound in interactive image-guided neurosurgery. Radiographics 1998;18:1019–1027.

Detmer PR, Bashein G, Hodges T, et al. 3-D ultrasonic image feature localization based on magnetic scanhead tracking: In vitro calibration and validation. Ultrasound Med Biol 1994;20:923–936.

Edwards WS, Deforge C, Kim Y. Interactive three-dimensional ultrasound using a programmable multimedia processor. Int J Imaging Syst Technol 1998;9:442–454.

Hata N, Dohi T, Iseki H, Takakura K. Development of a frameless and armless stereotactic neuronavigation system with ultrasonographic registration. Neurosurgery 1997;41:608–613.

Leotta DF, Detmer PR, Martin RW. Performance of a miniature magnetic position sensor for three-dimensional ultrasound imaging. Ultrasound Med Biol 1997;23:597–609.

Maurer CR, Jr, Fitzpatrick JM, Wang MY, et al. Registration of head volume images using implantable fiducial markers. IEEE Trans Med Imaging 1997;16:447–462.

Pagoulatos N, Edwards WS, Haynor DR, Kim Y. Calibration and validation of freehand 3-D ultrasound systems based on DC magnetic tracking. Proc SPIE Med Imaging 1998;3335:59–71.

Pagoulatos N, Edwards WS, Haynor DR, Kim Y. Interactive 3-D registration of ultrasound and magnetic resonance images based on a magnetic position sensor. IEEE Trans Inf Technol Biomed 1999a; 3:278–288.

Pagoulatos N, Haynor DR, Kim Y. Fast calibration for 3-D ultrasound imaging and multimodality image registration. Proc BMES/EMBS Conf 1999b:1065.

Pagoulatos N, Haynor DR, Kim Y. Image-based registration of ultrasound and magnetic resonance images: A preliminary study. Proc SPIE Med Imaging 2000; 3976:156–164.

Peria O, Chevalier L, Francois-Joubert A, et al. Using a 3-D position sensor for registration of SPECT and US images of the kidney. Proc Computer Vision, Virtual Reality Robotics Surg 1995:23–29.

Prager RW, Rohling RN, Gee AH, Berman L. Rapid calibration for 3-D freehand ultrasound. Ultrasound Med Biol 1998;24:855–869.