ELSEVIER

doi:10.1016/j.ultrasmedbio.2004.11.001

● *Review*

A REVIEW OF CALIBRATION TECHNIQUES FOR FREEHAND 3-D ULTRASOUND SYSTEMS

LAURENCE MERCIER,* THOMAS LANGØ,† FRANK LINDSETH,† and LOUIS D. COLLINS*

*Montreal Neurological Institute, McGill University, Montreal, QUE, Canada; and †SINTEF Health Research,
Medical Technology, Trondheim, Norway

Abstract—**Three-dimensional (3-D) ultrasound (US) is an emerging new technology with numerous clinical applications. Ultrasound probe calibration is an obligatory step to build 3-D volumes from 2-D images acquired in a freehand US system. The role of calibration is to find the mathematical transformation that converts the 2-D coordinates of pixels in the US image into 3-D coordinates in the frame of reference of a position sensor attached to the US probe. This article is a comprehensive review of what has been published in the field of US probe calibration for 3-D US. The article covers the topics of tracking technologies, US image acquisition, phantom design, speed of sound issues, feature extraction, least-squares minimization, temporal calibration, calibration evaluation techniques and phantom comparisons. The calibration phantoms and methods have also been classified in tables to give a better overview of the existing methods. (E-mail: laurence@bic.mni.mcgill.ca) © 2005 World Federation for Ultrasound in Medicine & Biology.**

Key Words: Ultrasound probe calibration, 3-D ultrasound, Freehand acquisition, Navigation.

INTRODUCTION

Ultrasound (US) is an appealing imaging modality because it is relatively inexpensive, safe, noninvasive, compact, portable and can image in real-time almost any body tissue. For these reasons, US is widely used and even gaining popularity in fields such as intraoperative imaging. Conventional US is a 2-D modality, in contrast to computed tomography (CT), magnetic resonance imaging (MRI) and other modalities that are volumetric. Three-dimensional (3-D) US is an emerging new technology that has many advantages over 2-D imaging: it allows the direct visualization of 3-D anatomy; 2-D slice views can be generated at arbitrary orientations; and volume and other measurements may be obtained more accurately. Measuring the volume of the prostate (Crivianu-Gaita et al. 1997; Hoffmann et al. 2003), monitoring fetal development (Kelly et al. 1994) or evaluating brain shift during neurosurgery (Comeau et al. 2000; Unsgaard et al. 2002) are examples of applications for 3-D US. For a more detailed list of applications, refer to Nelson and Pretorius (1998).

There are four general methods to construct an US volume. They are classified into the following categories: 1. constrained sweeping techniques, 2. 3-D probes, 3. sensorless techniques, and 4. 2-D tracked probe (also known as "freehand") techniques. They can be described as follows:

1. The constrained sweeping systems are characterized by a spatially predefined, constrained sweeping of the entire 2-D probe body that can be accomplished with a motor attached to the probe. Slices are generally either acquired in a wedge (fan-like) pattern, in a series of parallel slices (translation, as for MRI/CT), or with a rotation around a central axis (Fenster and Downey 2000).

2. 3-D US probes usually consist of 2-D arrays that allow explicit imaging in 3-D. These probes are relatively large and expensive in comparison with 2-D probes and their image resolution is not as good as their 2-D counterparts; refer to Light et al. (1998) for more information. Other 3-D probes can be either mechanically or electronically steered within the probe housing. An annular array producing a thin US beam can be accurately controlled by an internal mechanical motor in 2-D, to obtain a 3-D volume with high resolution. 2-D probes can also be electron-

Address correspondence to: Laurence Mercier, M.Eng, McConnell Brain Imaging Center, room WB325, Montreal Neurological Institute, 3801 University, Montreal, QUE H3A 2B4 Canada. E-mail: laurence@bic.mni.mcgill.ca

ically steered within the image plane to increase the field-of-view (FOV), as in Rohling et al. (2003).

3. The sensorless techniques attempt to estimate the 3-D position and orientation of a probe in space. Pennec et al. (2003), for example, proposed a system where a time sequence of 3-D US volumes is registered to play the role of a tracking system. Sensorless tracking can be done by analyzing the speckle in the US images using decorrelation (Tuthill et al. 1998) or linear regression (Prager et al. 2003). However, Li et al. (2002) found that it was impossible to accomplish real freehand scanning using only speckle correlation analysis. Although the Prager et al. (2003) results are encouraging, their sensorless approach is still far from the accuracy obtained with tracked probes.

4. Freehand systems allow image acquisition with unconstrained movement. They generally consist of a sensor (attached to a probe) that is tracked by a device that calculates the sensor's position and orientation at any point in time. This information is used to compute the 3-D coordinates of each pixel of the US images. Locating US images within a tracked coordinate system opens up a new world of possibilities: the images can be registered to a patient and to images from other modalities (Arbel et al. 2001; Brendel et al. 2002; Comeau et al. 2000; Dey et al. 2002; Lindseth et al. 2003b; Lindseth et al. 2003c). All the tracking devices used for freehand systems work in a similar manner: the device tracks the position and orientation (pose) of the sensor on the probe, not the US image plane itself. So, an additional step must be added to compute the transformation (rotation, translation and scaling) between the origin of the sensor mounted on the probe and the image plane itself. The process of finding this transformation is called calibration and is the focus of this article.

The objective of this paper is to review what has been published in the past 10 years in the field of calibration techniques for freehand 3-D US systems. The first section of this article covers the tracking technologies and the second section covers the acquisition of US images. The third section introduces calibration and all aspects of the problem. The fourth section discusses the methods used to test the calibration. Finally, the fifth section summarizes the results obtained by the different research groups. The main contribution of this paper is the comprehensive review and classification of all the different calibration techniques.

TRACKING

There are four common technologies to track medical instruments, 1. mechanical, 2. acoustical, 3. electromagnetic, and 4. optical. The role of a tracking system in the context of 3-D US is to determine the position and orientation of a sensor attached to the US probe. After calibration is performed, every pixel in each 2-D image is mapped in the 3-D coordinate system of the tracking device to reconstruct a geometrically correct volume. In the following paragraphs are brief descriptions of the four tracking technologies. For more details on any of them, refer to Meyer and Biocca (1992) or Cinquin et al. (1995).

Mechanical technologies

Mechanical localizers were the first tracking systems to be used. They are articulated arms with tip positions that can be determined by the angles formed by each joint. The FARO surgical arm (FARO Medical Technologies, Orlando, FL) is an example of this technology. Mechanical arms are accurate, but can only track one object at a time, which can be a limitation during surgery when multiple tools may be tracked simultaneously. Most importantly, they are cumbersome. This has led to the use of other tracking technologies below.

Acoustical technologies

Acoustical position trackers are composed of speakers (emitters) that emit US waves that are detected by microphones (receivers). There are two approaches: time of flight (TOF) systems measure the propagation time of the sound waves from the emitters to the receivers and phase-coherent (PC) systems deal with phase difference to compute relative positions. These systems are affected by variations in temperature, pressure and humidity, all of which affect the propagation speed of sound in air. They also require lines of sight between the speakers and microphones.

Electromagnetic technologies

The idea behind the electromagnetic system is to have a receiver placed on a probe that measures the induced electrical currents when moved within a magnetic field generated by either an alternating current (AC) or direct current (DC) transmitter. The AC and DC devices are both sensitive to some types of metallic objects placed too close to the transmitter or receiver, and to magnetic fields generated by power sources and devices such as cathode-ray tube monitors. Therefore, both types of electromagnetic systems are challenging to use in an environment such as an operating room, where various metallic objects are moved around in the field (Birkfellner et al. 1998). The two metal-related phenomena that influence the performance of electromagnetic tracking systems are ferromagnetism and eddy currents (Kindratenko 2000). Ferromagnetic materials (*e.g.*, iron, steel) affect both AC and DC systems, because they change the homogeneity of the tracker-generated mag-

Table 1. Accuracy of four tracking devices as given by the manufacturers

Models	Accuracy
Polaris passive or active (NDI)	Position: 0.76 mm r.m.s.; Orientation: 0.15° r.m.s. (for a single marker in whole volume)
Optotrack 3020 (NDI)	0.1 mm r.m.s. for x, y coordinates; 0.15 mm r.m.s. for z coordinate (for a single marker at 2.25 m distance)
Fastrack (Polhemus)	Position: 0.76 mm r.m.s.; Orientation: 0.15° r.m.s. (static accuracy)
Flock of Birds (Ascension)	Position: 1.8 mm r.m.s.; Orientation: 0.5° r.m.s. (static accuracy)

netic field, although the DC systems may be more sensitive to these effects (Birkfellner et al. 1998). In contrast, the AC technology is more affected by the presence of good conductors such as copper and aluminum because of distortions caused by eddy currents (Birkfellner et al. 1998). DC systems minimize the eddy-current–related distortions by sampling the field after eddy currents have decayed. Refer to Rousseau and Barillot (2002) for further comparisons of AC and DC localizers.

Optical technologies

The general idea with optical tracking is to use multiple cameras with markers distributed on a rigid structure, where the geometry is specified beforehand. At least three markers are necessary to determine the position and orientation of the rigid body in space. Additional markers allow a better camera visibility of the tracked object and improve the measurement accuracy. In addition, both the visibility of the tracked object and the accuracy of its 3-D position and orientation are highly dependent on the position of the markers. Refer to the paper of West and Maurer (2004) for more details. The markers can be infrared light-emitting diodes (active markers) or infrared light reflectors (passive markers) in the shape of spheres or discs. Both passive (Lindseth et al. 2003c) and active (Treece et al. 2003) markers have been used for calibration and tracking of medical instruments. To make a distinction between the markers on the tools, a different strategy is used for each technology. The passive markers are placed asymmetrically on the rigid body, which leaves no ambiguity as to the orientation of the tool. Therefore, to track multiple objects simultaneously using passive markers, the spatial configuration of markers attached to each object must be both asymmetrical and unique. Each tool with active markers must be wired to a control unit that cyclically activates the diodes. If more than one active tool is used, each will fire in its own time frame, allowing the same geometry to be used multiple times.

There are two types of camera systems, the two-camera models such as the Polaris by NDI (Northern Digital, Toronto, ONT, Canada) and the three-camera models such as the Optotrack by NDI or the FlashPoint 5000 (Boulder Innovation Group, Inc., Boulder, CO) used by Welch et al. (2000).

Comparing the electromagnetic and optical technologies

The most popular tracking devices used for calibration have been the AC electromagnetic models by Polhemus (Polhemus Incorporated, Colchester, USA) (Barry et al. 1997; Carr 1996; Prager et al. 1998b), the DC electromagnetic models by Ascension (Ascension Technology, Burlington, USA) (Berg et al. 1999; Boctor et al. 2003; Leotta 2004; Liu et al. 1998; Pagoulatos et al. 2001; Rousseau et al. 2003b), the Polaris optical tracker by NDI (Bouchet et al. 2001; Gobbi et al. 1999; Lindseth et al. 2003c; Treece et al. 2003) and NDI's Optotrack optical tracker (Blackall et al. 2000; Kowal et al. 2003; Muratore and Galloway 2001; Sato et al. 1998; Zhang et al. 2002). The discussion below will, thus, focus on these technologies in particular. For simplicity, the word "sensor" will be used through the article to identify both the receiver in an electromagnetic system and the rigid body with markers in an optical tracking system.

The choice of one technology over another depends on the type of utilization and the environment in which the tracking system will be used. The accuracy requirements also vary with the clinical application. For example, navigation during brain surgery usually requires higher accuracy compared with routine fetal examinations. Both the optical and electromagnetic systems have their strengths and limitations. One of the first major issues is accuracy. For a general idea, the accuracies, as reported by the manufacturers of the four most used models for tracking medical instruments, are found in Table 1. Caution must be taken when comparing the accuracy of the different models because the statistics given are often not equivalent. The Polaris and Optotrack, for example, are both manufactured by NDI, but no comparable accuracy measures are given to compare the two. For more detailed statistics on accuracy of the Polaris, refer to Wiles et al. (2004).

Chassat and Lavallée (1998) and Schmerber and Chassat (2001) compared four different optical systems in various conditions, the Polaris (active and passive), the Optotrack 3020 and the Flashpoint 5000. They found considerable differences between the passive and active systems, unlike that reported by the manufacturers (Wiles et al. 2004). The authors explain that these differences might be caused by the fact that they did not consider a flag that indicates when the markers are in
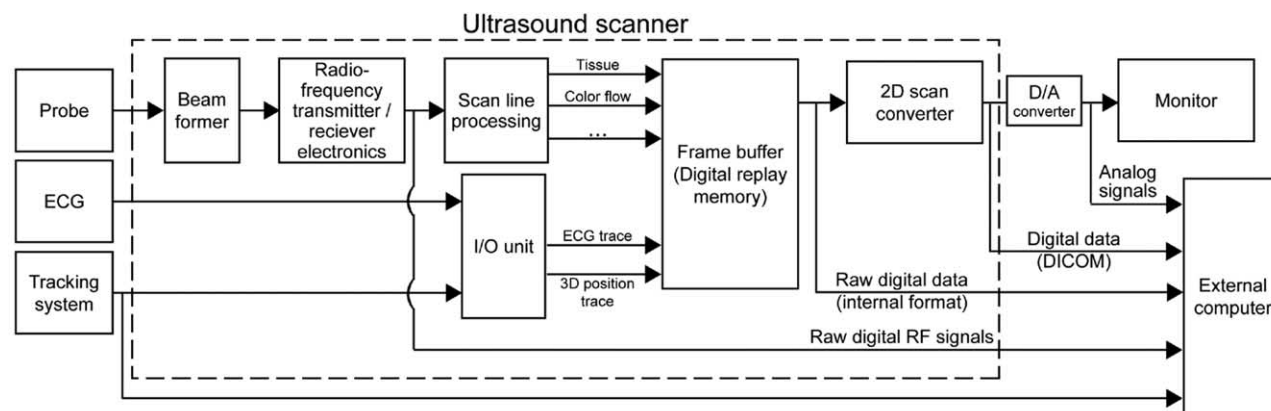
Fig. 1. The schematic diagram shows the data flow involved in a 3-D US setup. There are four data acquisition modes: analog signals, digital data after scan conversion, raw digital data before scan conversion and raw digital RF data, which are data before any scan line processing has been performed. The tracking system is generally connected to a computer, but it can also be routed into the scanner, as for the neuronavigation system SonoWand® (Gronningsaeter et al. 2000).

view, but are not in the optimal measurement volume. They conclude that the Optotrack has the best overall accuracy and robustness of the four systems compared. Khadem et al. (2000) compared the FlashPoint (300 mm, 580 mm and 1 m models) and Polaris (passive and active) camera jitter (precision) and found that, in all systems, the jitter was higher in the axis directed toward the camera. The Optotrack was found to be more accurate than the FARO mechanical arm (Rohling et al. 1995). Rousseau and Barillot (2002) found no significant difference between the FastTrack (Polhemus) and Flock of Birds (Ascension) electromagnetic systems.

Chassat and Lavallée (1998) and Schmerber and Chassat (2001) suggested that the accuracy of optical trackers could be better when used statically (probe held with a clamp) than when hand-held or moved around. During initial tests with an optical tracker, Langø (2000) also noted smaller root mean square (r.m.s.) values when acquiring images with the probe held still at each position, instead of in a continuous movement. Pagoulatos et al. (1998) performed similar tests, but with an electromagnetic tracker, and did not find any significant difference. It is not clear whether it is related to the findings above, but some authors using optical trackers acquired their images with the probe held in a clamp (Langø 2000; Lloret et al. 2002; Muratore and Galloway 2001; N Welch et al. 2000). In the temporal calibration section, we will see that this setup also reduces synchronization problems between the US machine and the tracking system to a minimum. Ionescu (1998) also noted that there was another advantage in taking several images at the same position; by averaging many images, they improved the signal-to-noise ratio (SNR). By taking images

from different positions, Rohling et al. (1997) reduced speckle artefact in their technique, known as spatial compounding.

In addition to accuracy, the price is often a major concern. In general, the more accurate the system, the more expensive it is. Finally, optical systems require clear lines of sight between the markers and the cameras, but the magnetic systems are unaffected by sensor occlusion. It is not a major concern for calibration itself, but it can become an important issue when using the system in an already cluttered operating room where the freedom of movement is limited.

IMAGE ACQUISITION

There are two common solutions to transfer images from an US machine to a computer. The most popular technique is to connect the analog output (*e.g.*, composite video, S-video) of an US machine to a frame-grabbing card on a computer (Comeau et al. 2000; Detmer et al. 1994; Meairs et al. 2000). The second method is to directly acquire digital images from the US machine, often by connecting them through a network cable (Barratt et al. 2001a; Berg et al. 1999; Lindseth et al. 2003c). Barry et al. (1997) opted for a different solution; they recorded the images using an S-videotape that was then digitized. Figure 1 shows a schematic block diagram of an US scanner with four possible methods for transferring the US data to an external computer. One output is analog and the three others are digital. They will be described in more detail after a brief description of the image formation process below.

Image formation process

The beam-former scans a narrow US beam over the image field by either electronic or mechanical steering of the transducer. The reflected signals correspond to the convolution integral between the spatial tissue density/compressibility distribution and the point-spread function (psf) of the imaging system. If we assume that the spatial tissue density distribution includes very high-frequency components (*i.e.*, small and closely located scatterers), then the limiting factor to the frequency contents of the signal will be the spatial frequency response of the psf. The real-time scan line processing unit analyses the backscattered radiofrequency (RF) signal from each US scan line to generate the tissue image, the flow image and the Doppler measurements. For tissue imaging, a compressed amplitude of the backscattered signal as a function of depth is generated after passing it through time-gain compensation and various filters. Afterward, these signals are transferred *via* a standard computer bus to the scan converter and the display unit. The data also go into a memory bank, so that, whenever the scanning is frozen, the last frames of the scan are stored.

Digital acquisition

The most common methods to obtain digital US data are to 1. transfer the signal after the scan line processing, but before scan conversion, or 2. transfer the RF signal before any scan line processing. The data can be transferred to an external computer over a digital connection. The advantages of these methods are that the representation of the information is compact; thus, reducing both transfer time and storage requirements, and there is no need for redigitization of the signal. The latter advantage eliminates the possible loss of information contained in the analog signal. The main disadvantage is that the scanner has to be built with the necessary hardware and software incorporated into it. Also, scan conversion needs to be done on the external computer after transfer to present images without geometric distortion. These are probably the best methods because of the compact data representation and lack of distortion through analog-to-digital conversion. The RF signal can be transferred by sampling it directly and storing it to a dedicated unit as raw digital RF data for further analysis. Sampling of the RF signals provides the largest freedom for postprocessing of the signals. This is useful for testing new algorithms for scan line processing where the phase of the echoes is necessary, for example. Finally, the third possible digital output is the one used by the DICOM output of some US machines. A DICOM file contains both a header and image data. The header contains information on the patient, the scan and the image dimensions.

Some groups have made special agreement with a manufacturer to have access to the real-time raw data. Hence, the digital setups usually work with a specific model, making this option less flexible than the frame-grabbing solution that can basically be used with any US machine. Some new US machine models with more open architectures, such as the Ultrasonix 500 RP (Ultrasonix Medical Corporation, Burnaby, Canada) used by Rohling et al. (2003), may soon eliminate the need for a special agreement with a manufacturer.

Analog acquisition

When using the analog output, the US machine performs a digital-to-analog conversion of the signal, which is then converted back to digital by the frame-grabbing card on the computer. This double (digital-to-analog-to-digital) conversion affects the image quality. Hence, the quality of the original digital images is certainly better than that of the converted data. However, no studies have compared the two configurations to confirm that the differences are important enough to allow a more accurate calibration.

Furthermore, there is a mismatch between the frame rate of the US machine and the frame rate of the analog video standards. The frame rate of an US machine is typically between one to 100 images per s (Rohling et al. 2003). Many settings, such as the depth and the number of focal zones, affect the frame rate. On the other hand, the PAL analog format produces 25 images per s and the NTSC format provides 29.97 images per s. When the video format has a lower frame rate than the US machine, there will be a degradation of the temporal resolution because frames will be dropped. Conversely, if the video format has a higher frame rate than the US machine, some frames will be duplicated and, hence, the same frame could have multiple tracking positions associated to it. Some US machines provide some means of identifying each frame; thus, allowing removal of the identical frames (Meairs et al. 2000).

CALIBRATION

This section defines, in a more detailed and graphic manner, what is involved for calibration. We begin with a short summary of the calibration process and then describe each step in detail. Figure 2 illustrates the coordinate systems used for locating the image plane of an US probe in space. Three coordinate systems are represented in Fig. 2, that of the position sensing device (called "world coordinate system"), that of the sensor mounted on the probe and that of the image plane. In Figs. 2 and 3 and in eqns (1) to (4), the transformations are represented by the letter "*T*" and the subscript should be read from right to left in the same manner as the matrix multiplications are carried out. The same notation
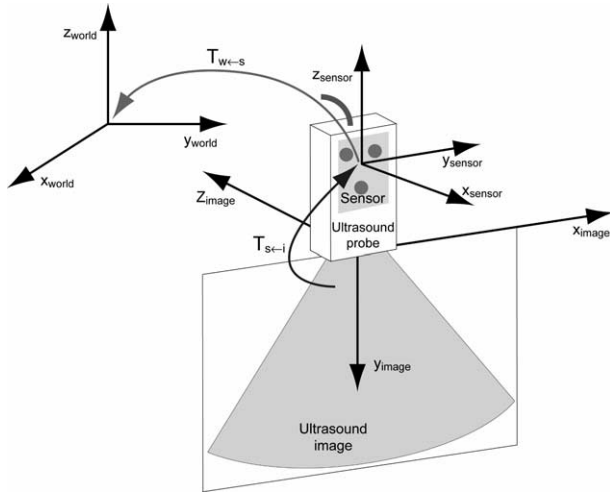
Fig. 2. World, sensor and image coordinate systems. $T_{w \leftarrow s}$ is the transformation relating the two spaces. $T_{s \leftarrow i}$ is the transformation relating image space to sensor space.

was used in Lindseth et al. (2003c). The sensor-to-world transformation $T_{w \leftarrow s}$ is measured by the tracking device. The origin of the world coordinate system is defined by a reference device (optical tracking) or by a transmitter (electromagnetic tracking) and must be fixed in relation to the phantom. To minimize angle effect errors (due to a lever effect at the reference), the reference device or transmitter should be placed as close as possible to the object to be scanned to perform the calibration (Detmer et al. 1994). The image-to-sensor transformation $T_{s \leftarrow i}$ is determined by calibration. Again, the sensor (receiver or rigid body) should be placed as close as possible to the imaging plane, to minimize angle effect errors. In concrete terms, calibration yields eight transformation parameters, three translations, three rotations and two image scaling factors.

A simple method for estimating the translational parameters is to perform external measurements of the probe casing and its attached sensor (Hughes et al. 1996). The problem with this method is the absence of external markers on the probe to identify the origin of the image. The same applies for magnetic sensors whose origin is also embedded in the receiver casing. In addition, the method obviously does not take into account the fact that the image is not necessarily centered on, nor perpendicular to, the probe face (rotation parameters).

A more precise calibration can be obtained by scanning an object known as a phantom with known geometric properties. The idea is to image the phantom and to identify its features on the US images. These features are also located in the physical phantom space. The spatial relationship between the position of the features in the image and the features on the phantom is estimated in the calibration

process. Hence, the coordinate system of the phantom must be included in the calculations (see Fig. 3).

A least squares minimization technique is used to minimize the distance between the sets of features (points or lines) identified in the image and on the phantom and, thereby, to find the unknown calibration parameters. Equation (1) converts a point in the kth image into the phantom coordinate system:

$$\begin{pmatrix} x_k \\ y_k \\ z_k \\ 1 \end{pmatrix} = T_{p \leftarrow w} \cdot T_{w \leftarrow s} \cdot T_{s \leftarrow i} \cdot \begin{pmatrix} s_x \cdot u_k \\ s_y \cdot v_k \\ 0 \\ 1 \end{pmatrix}. \quad (1)$$

For practical reasons, the origin of the image is often placed at the top center of the image. For a curved-array probe, this corresponds to the center of curvature. This point is preferred because its position does not vary when changing the depth setting. Any error in the location of that point will be compensated for in the translation parameters of $T_{s \leftarrow i}$. The point at position (u, v) from that origin is first scaled by s_k and s_y. It is then mapped in sensor space by the rigid transformation $T_{s \leftarrow i}$, then into world space by $T_{w \leftarrow s}$ and, finally, in phantom space by $T_{p \leftarrow w}$. $T_{w \leftarrow s}$ and $T_{p \leftarrow w}$ are rigid body transformations (translation + rotation). $T_{s \leftarrow i}$ is an affine transformation if the scaling factors are included in the calculations.

The next subsections summarize the most important aspects of calibration. The first subsection discusses the different categories of phantom designs. The second introduces the issues involved in the choice of a coupling medium to immerse the phantom. The third covers the strategies to locate the phantom in world space to find the $T_{p \leftarrow w}$ transformation. The fourth presents the methods to
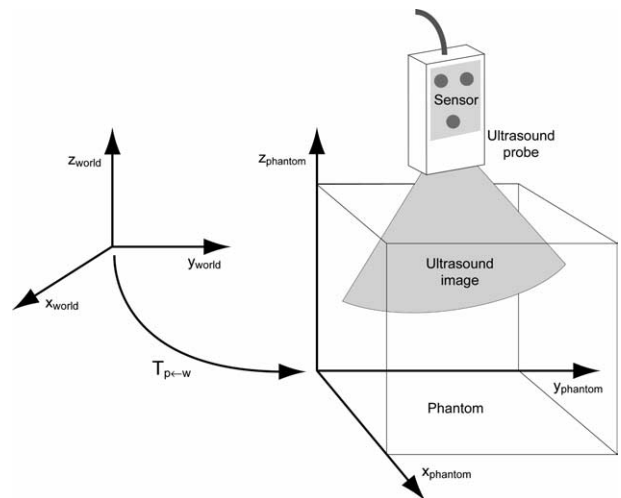

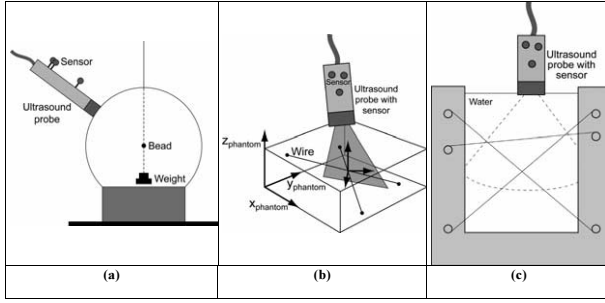Fig. 3. World and phantom coordinate systems. $T_{p \leftarrow w}$ is the transformation relating the two spaces.

Fig. 4. (a) Point target, (b) single cross-wire and (c) multiple cross-wire phantom examples.

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = T_{\mathrm{p \leftarrow w}} \cdot T_{\mathrm{w \leftarrow s}} \cdot T_{\mathrm{s \leftarrow i}} \cdot \begin{pmatrix} s_{\mathrm{x}} \cdot u_{\mathrm{k}} \\ s_{\mathrm{y}} \cdot v_{\mathrm{k}} \\ 0 \\ 1 \end{pmatrix}. \qquad (2)$$

extract the phantom's features on the US images. The fifth describes the methods to resolve the minimization problem. Finally, the sixth subsection classifies and compares all the published spatial calibration techniques found by the authors.

Phantom design

All the phantoms have one characteristic in common: they are built in or placed in a container that is filled with a coupling medium for US imaging. Coupling media will be discussed in the next section. The first detailed papers describing calibration with the help of a phantom were published by Detmer et al. (1994) and Trobaugh et al. (1994). State et al.(1994) also published on the subject in the same year, but with very little detail. This section introduces the different types or categories of phantoms that have been published since then. However, it excludes the phantoms that were solely used for testing because they will be discussed in the testing section.

Single point target and cross-wire phantoms. The first two types of phantom to be used were the point target (State et al. 1994) and cross-wire phantoms (Detmer et al. 1994; Trobaugh et al. 1994b). The single-point target phantom generally is based on imaging a small spherical object such as a bead or a pin head (see Fig. 4a). The single cross-wire phantom is composed of two intersecting wires (see Fig. 4b). The point target or wire crossing are aligned in the US image plane and are imaged from several viewing angles. Usually, the center of the point target or the intersection of the two wires is manually segmented on the image, although some automatic techniques exist. In some papers, these two methods are referred to as single-point methods or point-based methods because they both are based on mapping a single point from image space to phantom space. The center of the point target or the intersection of the two wires is usually considered as being the origin of the phantom coordinate system, yielding the following equation (Prager et al. 1998b):

The accuracy of these methods depends on how well the point-of-interest can be located in the phantom and on the image (Prager et al. 1998b). In all cases except one, the point-of-interest was fixed in the phantom. The exception being Muratore and Galloway (2001) who proposed an original point target method requiring no phantom. They simply imaged the tip of a tracked pointer that could freely move around within the coupling medium while the probe was fixed.

Multiple point targets and cross-wire phantoms. Multiple cross-wire techniques are derived from the single cross-wire phantoms. These phantoms are composed of more than one wire crossing that, again, requires alignment of one or more of the crossings in the image (see Fig. 4c). They are either three colinear points (Trobaugh et al. 1994a) or three coplanar wires forming a triangle (Henry 1997; Ionescu 1998; Péria et al. 1995). Meairs et al. (2000) also used three coplanar wires forming a triangle and added a single cross-wire below. However, the triangle and cross-wire were not coplanar, so only one edge of the triangle and the lower cross-wire were aligned at a time. The US images of these phantoms are, hence, composed of points (wire-cross) and/or lines. Kowal et al. (2003) described a phantom where the US image plane needed to be aligned with four coplanar 1-mm pins, whose tips acted as point targets and, as such, are classified in the multiple-point targets phantom family. Another phantom classified in this last category is that of Leotta (2004). The method is based on aligning the image plane with a planar array of beads attached to strings. Only the location of the strings and a reference bead is known exactly. The other beads are coplanar, but have arbitrary positions. Their purpose is to help in the alignment process. Additional out-of-plane strings also serve as visual guides for a better alignment.

2-D shape alignment phantoms. 2-D shape alignment phantoms are similar to the multiple cross-wire phantoms (see Fig. 5a). The idea, in both cases, is to align points-of-interest of a 2-D object in the US image. In the former case, the object is a solid 2-D geometric form with corners. In the latter, the object is delimited by intersecting wires. On the US images of these phantoms, the outline of the structures is of higher intensity and the corners are usually manually segmented (Sato et al. 1998).

Three-wire phantoms. The three-wire phantoms are made of three orthogonal wires that are scanned sequen-
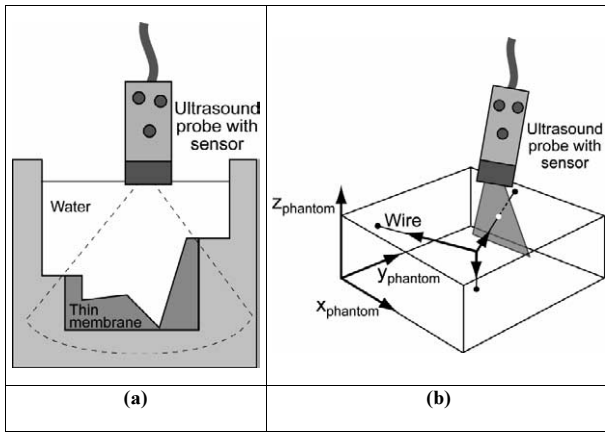
Fig. 5. (a) 2-D shape alignment and (b) three-wire phantom examples.

tially along their length (see Fig. 5b). The idea is to put the origin of the phantom coordinate system at the intersection of the three wires and to assign each wire to one axis. The equation representing any point x_k on the wire associated with the x-axis is (Prager et al. 1998b):

$$\begin{pmatrix} x_k \\ 0 \\ 0 \\ 1 \end{pmatrix} = T_{p \leftarrow w} \cdot T_{w \leftarrow s} \cdot T_{s \leftarrow i} \cdot \begin{pmatrix} s_x \cdot u_k \\ s_y \cdot v_k \\ 0 \\ 1 \end{pmatrix} \quad (3)$$

where u_k, v_k, s_x and s_y are the same as before. This type of phantom does not require alignment with the scan plane, so the scanning procedure is facilitated, although one must keep track of which wire is being scanned.

Z-fiducial phantoms. Z-fiducial (or N-fiducial) phantoms (see Fig. 6) were inspired by the stereotactic head frame described by Brown (1979) to register preoperative CT scans with the patient during neurosurgery. The wires of these phantoms form Z shapes that are intersected by the image plane, as illustrated in Fig. 6. The position of the end-points E1, E2, E3 and E4 (Fig. 6c) are known by construction. The colinear points U1, U2 and U3 are visible on the image (Fig. 6b). Using similar triangles, it is possible to compute the coordinate of point U2 on the phantom. Hence, each z-fiducial produces three visible points on the image, but only one serves as a homologous point between the image and the phantom. The first article published with this technique for US described a phantom with only three z-fiducials (Comeau et al. 1998) that was actually made of small tubes instead of wires. Since then, the number of z-fiducials has gradually increased to 30 (Pagoulatos et al. 2001), increasing the registration accuracy. Lindseth et al. (2003c) proposed a phantom with a pyramidal arrangement of z-fiducials for curved-array probes (see sample US image in Fig. 6d). It also had a higher density of fiducials near the top of the image, so that, even when smaller depth settings were chosen, enough z-fiducials were visible.

Other wire phantoms. There are other techniques using wires that are not z-fiducials and where the wire-crossings do not need to be aligned. These will simply be classified as other wire phantoms. This category includes a triangular pyramid (Liu et al. 1998), a ladder of strings (Beasley et al. 1999), nine orthogonal wire crossings forming a cube called the Diagonal phantom (Lindseth et al. 2003c) and parallel wires in the shape of a cross known as the Hopkins US phantom (Boctor et al. 2003).

Wall phantoms. The wall methods all produce a line on the US image, which is attractive because image information for the line is more redundant, making it easier to segment than points. If a line is partially missing, it can still be easily identified, which is not the case for points. Three techniques are in this family; they are the single-wall, membrane and Cambridge phantoms. The simplest wall method is the single-wall technique
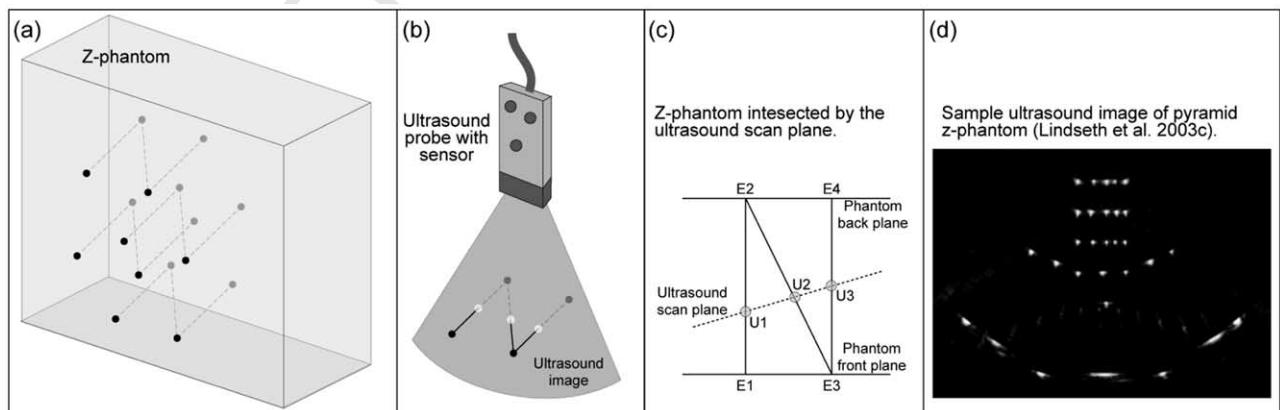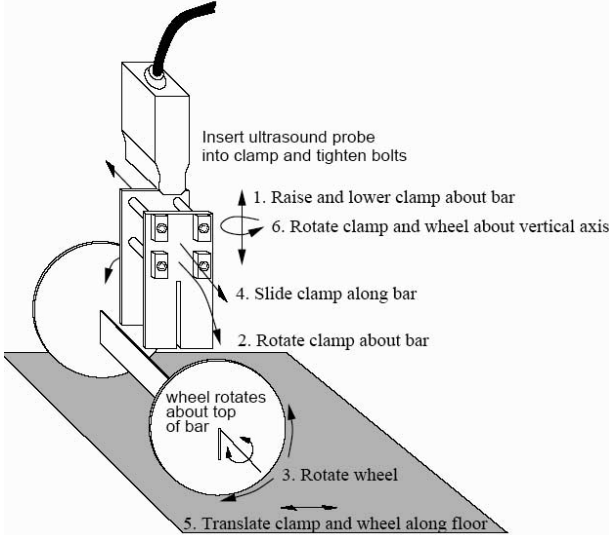

Fig. 6. An example of a Z-phantom.

Fig. 7. Cambridge phantom (Prager et al. 1998b).

(Prager et al. 1998b) which is based on imaging the bottom of a water tank. The membrane technique (Langø 2000) solves the reverberation problems of the first, by imaging a thin membrane instead of the bottom of the tank. Hence, this solution produces thinner lines on the images. Care must be taken, however, to choose a membrane rigid enough to minimize the membrane oscillation caused by the movements of the probe in water (Langø 2000). In both cases, difficulties arise when imaging at an angle far from the normal. In the first case, most beams will be reflected away from the probe because of specular reflection, yielding a lower intensity line. In the second case, the line on the image will loose its sharpness, because of the US beam thickness. In the case of the single-wall phantom, simply roughening the bottom of the tank helps to compensate for the specular reflection problem. Mathematically, the plane is considered to be at z = 0, with the z-axis orthogonal to the plane; hence, the two phantoms above are described by:

$$
\begin{pmatrix} x_k \\ y_k \\ 0 \\ 1 \end{pmatrix} = T_{p \leftarrow w} \cdot T_{w \leftarrow s} \cdot T_{s \leftarrow i} \cdot \begin{pmatrix} s_x \cdot u_k \\ s_y \cdot v_k \\ 0 \\ 1 \end{pmatrix}. \quad (4)
$$

The Cambridge phantom (Prager et al. 1998b) was created to solve the problems mentioned above. The probe is attached in a clamp in such a way that the top of a thin brass bar is always in the center of the beam (see Fig. 7). To ensure this alignment, Prager et al. (1998b) describe a separate technique including another piece of equipment. After alignment, the phantom is immersed in a water bath; the clamp with the probe is placed over the

bar and the bar is scanned from all possible angles, subject to the constraints imposed by the setup. The top edge of the bar acts as a virtual plane, yielding a line in the US image that is sharper and of relatively higher intensity. The wall methods are among the quickest solutions for calibration, due to the possibility of automatic extraction of the lines in the US images. The Cambridge phantom is patented (Prager 1997) and it is possible to buy one from the Cambridge group.

Multimodal registration phantom. Blackall et al. (2000) took a very different approach to the calibration problem. They acquired a set of 2-D tracked US images and a 3-D MRI scan of a gelatin phantom. The calibration parameters were estimated by registering the 2-D US images to the corresponding plane in the MRI volume. In fact, the chosen calibration parameters were those that maximized the similarity between the US images and the 3-D MRI scan. The similarity was evaluated by using normalized mutual information. The method requires no feature extraction and is fully automatic.

Other aspects of phantom design. Phantoms are often imaged from many directions and positions, but the imaging is often restricted to the top of the tank where the only opening is found. Some authors have used special tanks to be able to image the phantom from more viewpoints. Legget et al. (1998), for example, enclosed their phantom in a water-filled plastic ball (see Fig. 4a). Detmer et al. (1994) used a polypropylene bottle, enabling images to be taken from the sides. Boctor et al. (2003) used a tank, from which it was possible to image from the top and also from all four sides through rubber windows. They compared the top-only acquisition with the top and sides acquisition, both comprising the same number of images. Their results suggest that their multisided tank improves accuracy.

Finally, important aspects of the phantom design are the efficiency and facility of data collection. If an alignment-based calibration method is used, the probe can be positioned by hand or by attaching the probe in a holder to ensure stable probe positions during image acquisition. Acquiring images of the phantoms that do not require any alignment, such as the z-fiducial phantom, is a lot simpler as it can be done completely by freehand acquisition.

Speed-of-sound issues

US machines use the propagation speed of sound waves in a medium to compute the distance between a transducer (source) and an acoustic interface from which a sound wave has been reflected (reflector). The simple calculation to find the distance *d* from a source to a reflector is:
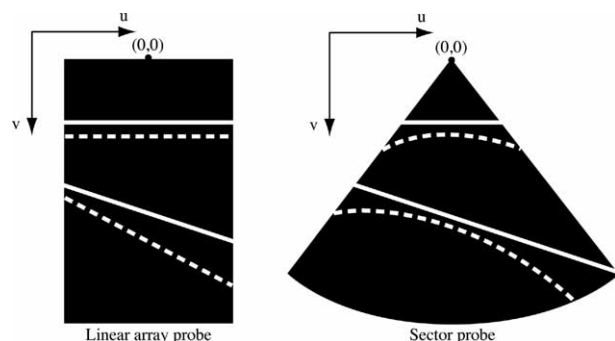
Fig. 8. Distortion (dotted lines) of straight wires immersed in a water bath at room temperature. (—) The correct imaging of the lines when adjusting for the sound speed.

$$d = \text{propagation time } * \text{ speed of sound in medium} / 2. \tag{5}$$

The speed of sound depends on the medium in which the sound waves travel and its temperature. The speed of sound assumed by most US machines is 1540 m/s, which is the average speed of sound in human tissue. If the speed of sound in a medium is different from the one assumed by the US machine, the objects in this medium will appear farther (< 1540 m/s) or closer (> 1540 m/s) and their shapes might appear distorted. Figure 8 illustrates what theoretically happens to a straight wire when imaged in a medium with a lower speed of sound, such as water at room temperature, which has a speed of sound of approximately 1485 m/s (Bilaniuk and Wong 1993). It shows that the distortion is different, depending on the beam pattern of each type of probe. Because sound waves travel along beams, the speed difference linearly scales the computed distances along each beam. When imaging with a linear probe, only the depth parameter, represented by the letter v in Fig. 8, will be affected. The depth will be scaled by the ratio R:

$$R = \text{assumed speed of sound} / \text{actual speed of sound.} \tag{6}$$

When imaging in water at room temperature, that ratio would be:

$$R = 1540 / 1485 \sim 1.04. \tag{7}$$

Hence, an object placed at a depth of 10 cm would, in fact, appear at 10.4 cm on the US image. In the case of a sector probe (that includes both phased and curved-array probes), both the u and the v axes will be nonlinearly scaled. If the image is viewed from a polar coordinate system, the radius would be scaled by the same ratio R as above. See Goldstein (2000) for a review of the impact of sound speed errors for different probe types.

Thus, the distortion produced by the difference in the speed of sound affects the scaling parameters of the calibration either in a linear or nonlinear fashion, depending on the probe used. Note that the rotational and translational components of the calibration matrix are not affected. Because the scaling parameters vary with the speed of sound of the medium in which the calibration is done, they will only be valid for a particular medium at a particular temperature. This means that, if one wishes to build an *in vivo* US volume of some organ ($\approx$ 1540 m/s), it is not ideal directly to use the scaling parameters obtained from a calibration in pure water at room temperature ($\approx$ 1485 m/s). For a comprehensive list of the speeds of sound for *in vivo* and *in vitro* mammalian tissue, see Goss et al. (1978).

Some groups working with raw digital images avoid the speed-of-sound problem by directly modifying the assumed speed of sound in the US machine (Lindseth et al. 2003c). When it is not possible to modify the assumed speed of sound, two solutions can be applied. In the first, the appropriate ratio R is computed for the type of probe used and the medium's measured temperature, as was explained above (Ionescu 1998; Pagoulatos et al. 2001; Trobaugh et al. 1994b). The distances along the beams are then divided by this ratio to obtain a corrected image. This solution is simple for linear probes, but becomes more complicated for other probe designs, as seen in Fig. 8. The second solution is to work directly in a medium with a speed of sound similar to that of human tissue (Barry et al. 1997; Bouchet et al. 2001; Comeau et al. 1998; Pagoulatos et al. 1998; Rousseau et al. 2003b; Treece et al. 2003). Recent studies suggest that the first solution might produce images of lower quality. Imaging in a medium with a speed of sound different from the one assumed would cause negative effects in addition to the shift and distortion presented above. It would first create a broadening of the beam and, hence, a decrease in the lateral resolution (Anderson et al. 2000; Dudley et al. 2002). Second, it would decrease a point target echo amplitude relative to its surrounding speckle background, therefore reducing the contrast (Anderson et al. 2000).

To match the speed of sound in tissue, a simple solution adopted by Boctor et al. (2003) and Treece et al. (2003) is to increase the temperature of water to approximately 50°C (122 °F). Another simple solution is to add glycerol (Gobbi 2003) or ethanol (Rousseau 2003) to water at room temperature. Martin and Spinks (2001) have shown that, by mixing 9.5 $\pm$ 0.25% ethanol with water at 20 $\pm$ 0.75 °C, the resulting speed of sound is 1540 $\pm$ 1.5 m/s. Others have used tissue-mimicking media (Bouchet et al. 2001; Pagoulatos et al. 1998) additionally to match the attenuation of sound in tissue.

Very few authors have published results that deal

directly with speed-of-sound-related distortions. Sato et al. (1998) performed their calibrations in water at 25°C and their validation in water at 25°C and 40°C. The reconstruction at 40 °C appeared to be slightly worse. In his thesis, Rousseau (2003) compared water at room temperature (w) and an ethanol solution (e) ($\approx$ 1540 m/s). He tested the four following combinations: 1. calibration (e) + validation (e); 2. calibration (w) + validation (w); 3. calibration (e) + validation (w); and 4. calibration (w) + validation (e). The expected result was to obtain the best accuracy with the combination using only the ethanol solution. However, the best results were obtained when using the calibration (e) + validation (w) combination. The author explains this result by the fact that the images were of lesser quality when the ethanol solution was used.

Finally, even if all precautions are taken during the calibration, there will always be speed-of-sound distortions when imaging *in vivo*. Indeed, even though the typical assumed speed of sound is 1540 m/s, in human tissues it ranges from approximately 1450 m/s in fat to 1600 m/s in muscles. Furthermore, imaging *in vivo* often means imaging through different layers of tissues with different acoustic properties, compounding the problem. Lindseth et al. (2002) estimated the error resulting from the speed-of-sound uncertainties in their US-based navigation system to be between 0.5 and 3 mm, even after manually controlling the speed of sound of their US machine.

Locating the phantom features in world space

When the position of the phantom in world space is needed for calculations, a few approaches are possible. One is to determine the location of each phantom feature with a tracked pointer. Sato et al. (1998) identified each corner of their 2-D shape alignment phantom with a pointer. The measurements were repeated 6 times and averaged. Welch et al. (2000) went further, averaging over 100 positions. To locate the single wire crossing of their phantom, Hartov et al. (1999) used a similar technique. However, instead of pointing at the cross-wire directly, they pointed at the four end-points of the wires and then computed the location of the crossing. Instead of locating the features directly, some have used drilled holes (divots) on the tank containing the phantom or on the phantom itself. The feature's location relative to these divots is known by construction. Gobbi et al. (1999) used four divots, Amin et al. (2001) used six and Pagoulatos et al. (2001) used 18.

To use the tip of the pointer to localize specific points, the pointer needs to be calibrated as well. Pointer calibration is a problem similar to probe calibration. It involves finding the transformation between the sensor attached on the pointer and the tip of the pointer. This transformation can be found by rotating the pointer while its tip is fixed in a drilled hole. Refer to Leotta et al.

(1997); Hartov et al. (1999) and Zhang (2003) for more details on pointer calibration.

Other solutions not involving pointers also exist. Some groups using the optical tracking technology have put markers directly on their phantom, using it as the fixed reference device. Bouchet et al. (2001) used five active markers and Lindseth et al. (2003c) used four passive markers. The feature's location relative to the markers is then known by construction, as for the method using divots. A numerically controlled milling machine can be used to position features of the phantom to make contact with a needle pointer with very high precision (Lindseth et al. 2003a). Because the reference device needs to be fixed relative to what is being imaged, having the reference incorporated in the phantom makes it possible to move the phantom along with its reference during calibration.

Feature extraction

This section does not cover US segmentation in general, but focuses on phantom feature segmentation as needed for calibration, such as identifying wire crossings, points and lines.

The wall phantom family has the advantage of producing a line in the US image that is easier to segment than other features. Prager et al. (1998b) published a technique to automatically segment their single-wall and Cambridge phantoms. To speed up the process, they sampled only vertical lines at specified regular intervals. An edge-detection operator was then applied after a 1-D Gaussian smoothing and only the pixels over a preselected threshold were considered. In each column, only the point nearest the top of the image was kept. For more robustness to outliers, they chose the RANSAC algorithm (random sample consensus) (Fischler and Bolles 1981) instead of least squares to determine the optimal line. Langø (2000) also implemented an automatic algorithm to detect a line in the US images of his membrane phantom. Edge detection was done with a new and robust wavelet-based detection method (Kaspersen et al. 2001). Both Hook (2003) and Rousseau et al. (2003b) used the Hough transform to detect the line generated by their wall phantom. The first used a threshold on pixel intensity to identify candidate points. The second approach was more sophisticated because it used both gradient and pixel intensity information. To reject outliers, Rousseau et al. (2003b) also added a temporal constraint based on the principle that the line in two consecutive images should have similar parameters. Many years before, Ionescu (1998) used the Hough transform as well, but to detect three intersecting wires forming a triangle (complex cross-wire phantom). To select candidate points, he first applied a pixel intensity threshold and then applied

a Deriche filter (Deriche 1987, 1990), a recursive version of Canny's operator (Canny 1986).

Other phantom types that produce a single dot, such as the point target and the three-wire, or multiple dots, such as the z-fiducials or the other wire phantoms, are usually segmented manually. An improvement proposed by some authors has been to manually define a region-of-interest (ROI) around each dot and to automatically compute the pixel intensity centroid (Carr et al. 2000; Gobbi et al. 1999). Lindseth et al. (2003c) described the most elaborated single-point and multiple-point extraction methods. For their bead phantom (point target category), intensity information was used first to find candidate intensity peaks in the image and then spatial constraints were added to find the peak most likely produced by the bead. The last step was to compute the intensity centroid of the chosen intensity peak. For their Diagonal phantom (other wire phantom category), the preliminary intensity-based candidate feature selection was also used. The intermediate processing was based on trying to match an ideal geometry on a maximum number of candidate points. The ideal geometry was appropriately scaled and then incrementally translated and rotated to find the best match. The final step was again to compute the intensity centroids. A similar procedure was used to detect the extremities (parallel wires) of the z-fiducials of their pyramid phantom (pyramidal shape z-fiducials phantom). The middle point of each z-fiducial was found by intensity-based search with spatial constraints (middle point should be colinear with its two extremities).

Least-squares minimization

Because there is no exact solution to the calibration problem, the minimum residual error (in a least-squares sense) is used to solve the overdetermined system of equations. There are two approaches, the iterative and the closed-form (noniterative) methods. In both approaches, $T_{w \leftarrow s}$ is given by the tracking device and $T_{w \leftarrow i}$ is the unknown calibration matrix, as mentioned previously; refer to eqn (1). The idea behind the closed-form approach is to map the set of points in the images to those on the phantom; hence, $T_{p \leftarrow w}$ must be known. When using the iterative method, the position of the phantom's features in world space is generally unknown, but some authors do take the position information into account (Bouchet et al. 2001; Sato et al. 1998).

Recall that the calibration transformation has a translation, a rotation and a scaling component. Before getting into the details of the different approaches, it is important to mention that the two scaling factors are not treated in the same way by all authors. First, some force the two scaling factors to be equal (Rousseau et al. 2003b; Sato et al. 1998) and others compute them independently (Boctor et al. 2003; Hook 2003). Second, the scaling factors are often computed separately from the translation and rotation. This can be done by computing the ratio of the distances between points on an image and their corresponding distance in world space (Meairs et al. 2000; Nelson and Pretorius 1997). In contrast, Prager et al. (1998b) included the scaling in the minimization process. Finally, some digital setups provide the scaling directly (Lindseth et al. 2003c), simplifying the problem.

A closed-form solution to calibration involves mapping two corresponding point sets, $\{a_j\}$ of the image and $\{b_j\}$ of the phantom with $j = 1..N$. The goal is to find an optimal translation **T** and rotation **R** that map one set onto the other. The point sets are related by the following equation (Eggert et al. 1997):

$$b_j = Ra_j + T + \text{noise}. \tag{8}$$

This equation is often solved by minimizing a least squares error criterion given by Eggert et al. (1997):

$$\sum{}^2 = \sum_{j=1}^{N} \|b_j - Ra_j - T\|^2 \tag{9}$$

or, if the scaling s is included (Horn 1987):

$$\sum{}^2 = \sum_{j=1}^{N} \|b_j - sRa_j - T\|^2. \tag{10}$$

The difficult part of the problem is to find the rotation. A brief description of a few closed-form approaches follows. Arun et al. (1987) proposed a method where the rotations are represented by a 3×3 orthonormal matrix and are found using singular value decomposition (SVD). The translations are in the form of a 3-D vector and are found by aligning the centroids of one set of points with the rotated centroids of the other set. Horn et al. (1988) represented the translation and rotation in the same way, but used the eigenvector associated with the most positive eigenvalues of a symmetrical 4×4 matrix to find the rotations. There is a possible problem with these minimizations over orthogonal matrices: the algorithm is not constrained to return only a rotation. It could actually return a reflection (when the determinant $= -1$). Umeyama (1991) proposed a modified version of the Arun et al. technique that resolves this problem. Kanatani (1994) proposed a simplified version of Umeyama's solution. Other solutions using quaternions have also been published: an older paper by Horn (1987) describes a similar method where rotations are represented by unit quaternions instead and Walker et al. (1991) used dual quaternions to represent translations and rotations together.

Eggert et al. (1997) compared four closed-form algorithms, those of Arun et al. (1987); Horn (1987), Horn et al. (1988) and Walker et al. (1991). They found that no method was superior in all cases. In fact, they believe that no difference should be observed in a real-world application with a low level of noise.

Many iterative algorithms have been published in the calibration literature: the iterative closest point algorithm of Besl and McKay (1992) was used by Welch et al. (2002). The method of Hooke and Jeeves (1961) was used by Bouchet et al. (2001), but the most frequently mentioned technique for iterative optimization in the literature was the Levenberg–Marquardt algorithm (More 1977) and modified versions of it. The iterative methods need initial values to begin the optimization and the method is applied repeatedly until the remaining error becomes smaller than a predefined threshold. The Levenberg–Marquardt algorithm is a combination of the Gauss–Newton and the steepest descent algorithms.

Referring to eqn (4), a set of equations is obtained by scanning the membrane or wall from different positions and orientations. The z component on the left side is always zero. These equations may be written as:

$$F(\vartheta, \varphi) = 0 \qquad (11)$$

where ϑ represents the measured parameters and φ are the unknowns. Hence, we are looking for the parameter vector, φ, that minimizes F. The system of equations, which is overdetermined by making sure the number of equations is greater than the number of unknowns, may be solved with the Levenberg–Marquardt algorithm. Briefly explained, at iteration k, an update $\Delta\varphi$ to the current estimate of φ_k is obtained from the first order Taylor expansion of eqn (11):

$$F(\vartheta, \varphi) \approx F(\vartheta, \varphi_k) + \frac{\partial F(\vartheta, \varphi_k)}{\partial \varphi}(\varphi - \varphi_k) = 0. \quad (12)$$

By rearranging this equation, we get:

$$-F(\vartheta, \varphi_k) = \frac{\partial F(\vartheta, \varphi_k)}{\partial \varphi}(\varphi - \varphi_k) = J(\varphi - \varphi_k), \quad (13)$$

where J represents the Jacobean (*i.e.*, the gradient matrix of F). Each step of the Levenberg–Marquardt algorithm yields the updated parameters (More 1977):

$$\varphi_{k+1} = \varphi_k + (J^T J + \varepsilon I)^{-1} J^T (-F(\vartheta, \varphi_k)). \qquad (14)$$

I is the identity matrix, upper index T represents the transpose of a matrix and ε is a damping factor chosen at each step to stabilize convergence. $-F(\vartheta, \phi_k)$ and J are evaluated at each step of the algorithm for the current estimate φ_k and the process continues until the corrections are sufficiently small.

To check that the equations are significantly independent and, hence, that all unknowns are identifiable, a rank determination may be carried out. This is accomplished by evaluating the Jacobian matrix (*e.g.*, at the initial estimate of the calibration matrix). To determine that this matrix is full rank, a singular value decomposition (SVD) is performed.

As pointed out by Prager et al. (1998b), there are several different values for the angles that can achieve the same global minimum to the iteration process. These solutions are referred to as mirror solutions. This is not the same as a singularity, where an infinite number of solutions exist and the problem is termed ill-conditioned.

Although most calibration problems can be resolved by either technique, some are less flexible. The wall phantoms, for example, require the iterative approach because it is not possible to find the exact equation of the line defined by the intersection of the image plane with the plane of the wall phantom. The same applies for the three-wire phantom, for which it is impossible to know exactly where the US plane is intersecting on each wire. The principle underlying the z-fiducial phantom is to map points known exactly between two coordinate spaces. The closed-form solution is, thus, a more natural approach. Bouchet et al. (2001) worked with such a phantom and compared two closed-form and one iterative approaches using synthetic data with Gaussian noise. They concluded that both closed-form solutions were superior to the iterative optimization, particularly when retrieving the translation parameters and using a small number of points. They also found that their iterative algorithm would not always converge to the optimal transformation when noisy data sets were used. Muratore and Galloway (2001) also compared the iterative and closed-form techniques for their point target phantom (tracked pointer). They found that both techniques gave very similar results. For the point target, complex cross-wire, 2-D shape alignment and other wire phantoms, both solutions have been used. Finally, for the single cross-wire phantom, it was always solved iteratively, except by Hartov et al. (1999). Although they do not present any numerical results, they assert that they compared an iterative and closed-form solution and obtained very similar results for both. They opted for the closed-form solution, because they found it simpler to implement.

When both approaches are possible, it is not clear which one should be used or if it even makes a significant difference to the accuracy. Most authors have not justified why they chose a particular technique over another. It appears mostly to be related to personal preferences, although the closed-form algorithms appear to be simpler to use. According to Eggert et al. (1997), "closed-form solutions are generally superior to iterative methods, in terms of efficiency and robustness, because the latter
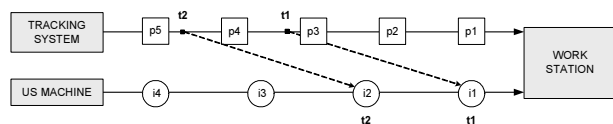
Fig. 9. Illustration of the data flow in a tracked US image setup.

suffer from the problems of not guaranteeing convergence, becoming trapped in local minima of the error function and requiring a good starting estimate."

Temporal calibration

In addition to spatial calibration, some research groups also perform what is known as a temporal calibration, which is the process of synchronizing each US image with its appropriate pose information. Indeed, finding the exact position and orientation of an image is not trivial:

1. The US machine and the tracking system do not generate continuous data. In addition, the rates at which they produce data are generally unequal and may not be constant if, for example, some settings on the US machine, such as the number of foci, are modified.
2. Furthermore, there is no synchronization between the US machine and the tracking system, except for systems where the tracking is integrated into the US scanner (Gronningsaeter et al. 2000). Hence, software on a workstation or external hardware must be used for time-stamping the B-scans and the pose data. When the US images and pose data are time-stamped, they are both already a few milliseconds old, due to internal processing and data transfer. This lateness is called latency and is not the same for both systems. Carr (1996) completed interesting experiments with an electromagnetic tracking system called 3Space (by Polhemus). He found that it took 21.6 ms (using serial communication on a RS-232 port) before receiving positional data after he requested them. This value was broken down into 0.5 ms to transmit the request, 3.5 ms to sample the magnetic field, 2 ms computation time and 15.6 ms to transmit the data. Newer systems using PCI or SCSI interfaces might yield lower temporal latency.

Figure 9 illustrates a typical data flow example. Pose data (p1 to p5) and US images (i1 to i4) are sent to a workstation. If no temporal calibration is applied, image $i1$ taken at time $t1$ will be associated with the pose $p1$, because they arrived at the same time to the workstation (in fact, the correct pose would be between p3 and p4). Furthermore, the positional error would grow proportionally with the probe's speed of movement. Thus,

acquiring images slowly would diminish the pose error associated with an image and, even though it is not a practical setup for imaging other than during calibration, taking images with the probe clamped at each position would bypass the problem.

The general idea of a temporal calibration procedure is first to time-stamp both the pose and image data streams. The latency difference is found experimentally and, afterward, the pose data are interpolated to estimate a pose between two measured pose values. Detailed references on temporal calibration in the tracked US context are given in Prager et al. (1998a); Treece et al. (2003) and the thesis of Gobbi (2003). A brief description of various solutions follows.

Barry et al. (1997) were the first to publish the details of their efforts to synchronize B-scan and pose streams. They used a hardware module that sent a signal to the tracking device every time that a frame arrived and then associated the incoming positional information with the current frame. Although they did not specify how it was done, they measured an offset of 18 ms between the images and the positional data. Barratt et al. (2001a, 2001b) also used a setup where pose data were generated on demand for each B-scan. They wrote custom software that triggered the image capture and tracker reading at a chosen instant in the ECG R-wave of the cardiac cycle. The offset between the two data streams was not mentioned.

Prager et al. (1998a) and Meairs et al. (2000) used similar techniques; they simultaneously created a step input in the image and position data streams by holding the probe still on the skin for a few seconds and then rapidly moving it away. The offset between the two data streams could be estimated by identifying the rapid change in the B-scan and the pose sequences. However, as Prager et al. (1998a) noted, " if images are acquired every t seconds and positions every T seconds, then the offset can be estimated only to an accuracy of $\pm (t + T)/2$ s".

Treece et al. (2003) proposed a new temporal calibration technique based on the same idea, but with an improved accuracy. In their technique, a single-wall phantom was imaged instead of skin. This enabled automatic segmentation of the straight line generated by the bottom of a water tank in each image. In this technique, the probe was immersed in water, held immobile and then moved up and down for a few seconds. A normalization process then gave a stream of distance measurements from the line for both the tracking system data and the US images. For the pose data, only the motion along the direction of maximal movement over the whole sequence was considered. Image distances were linearly interpolated to match in time each pose distance. Finally, the temporal offset was found by com-

Table 2. Single cross-wire phantom category

References	Tracking	Acq.	Coupling	Sol. type
Detmer et al. (1994)	DC mag	A	Water	Iterative
Barry et al. (1997)	AC mag	A (S-video)	Galactose solution	Iterative
Prager et al. (1998b)	AC mag	A	Water at room temp.	Iterative
Hartov et al. (1999)	DC mag	A	Water	Closed-form
Blackall et al. (2000)	Optical	A	Water	–
Lloret et al. (2002)	DC mag	A	Saline at room temp.	Iterative
Boctor et al. (2003)	DC mag	A	Water at 50°C	Iterative

puting the minimum root-mean-square error between the two distance streams.

The approach of Gobbi (2003) had some similarities to that of Treece et al. (2003); the cross-section of a string was imaged by moving a probe from side to side for a few seconds. A principal component analysis was applied to both the pose measurements and B-scans to determine the principal axis of motion. The motion parallel to this axis was then normalized by setting the mean position to zero. The time offset was found by minimizing the least-squares difference between the two normalized motion waveforms. The position components of the pose measurements were then linearly interpolated and the orientation measurement was interpolated by using spherical linear interpolation (SLERP) (Shoemake 1985).

The same year, Nakamoto et al. (2003) also published a temporal calibration method that involved imaging a point target, first holding the probe still and then from different orientations. Contrarily to the method described by Treece et al. (2003) and Gobbi (2003), a spatial calibration must first be performed to be able to localize the imaged point in world space. In their method, the poses were linearly interpolated and the latency was estimated using least squares.

Classification

In this section, all the spatial calibration techniques that the authors could find are classified (see Tables 1 to 11) according to the categories of phantoms that were introduced earlier. In the first column are the references. The inclusion criteria were:

● If different papers by the same first author described very similar spatial calibrations, then they are grouped together on the same line.
● If a scientist described a calibration technique in his/her thesis or a report but never published, then this is included in the list. Otherwise, published articles were generally chosen as the reference. Prager and colleagues, for example, published an internal report in 1997 and an article with similar content in 1998. Only the 1998 article is listed below.
● Only the phantoms used for calibration are listed here. Some authors used different phantoms for testing, but these are not included here.
● If an article refers to another for the calibration technique and does not modify it in any way or does not compare it with other techniques, it was not included.
● Only one paper describing a wire phantom was not detailed enough to be classified (Welch et al. 2002).

The column "description" is sometimes added to further describe the phantom. The column "tracking" identifies the tracking technology that was used. They were classified as "DC mag" (DC electromagnetic), "AC mag" (AC electromagnetic), optical or mechanical arm systems. The column "acq" identifies the acquisition setup, "A" for analog output grabbed by a capture card

Table 3. Single-point target phantom category

References	Description	Tracking	Acq.	Coupling	Sol. type
State et al. (1994)	4-mm bead	Optical	A	Water	Closed-form
Leotta et al. (1995)	1.5-mm brass sphere	DC mag	A	Water	Iterative
Leotta et al. (1997)	1.5-mm brass sphere	DC mag	A	Water	Iterative
Nelson and Pretorius (1997)	1-cm spherical latex balloon	AC mag	A	Water at room temp.	Iterative
Legget et al. (1998)	1.5-mm brass sphere	DC mag	A	Water	Closed-form
Pagoulatos et al. (1998)	1.5-mm stainless steel sphere	DC mag	A	Tissue-mimicking	Iterative
Amin et al. (2001)	1 mm steel ball bearing	Optical	A	Water	Closed-form
Barratt et al. (2001a)	Pin-head	DC mag	D	Water	Iterative
Muratore and Galloway (2001)	Tip of tracked pointer	Optical	A	Water	Both
Lindseth et al. (2003c)	2-mm needle pin head	Optical	D	Water	Closed-form

Table 4. Multiple cross-wires phantom category

References	Description	Tracking	Acq.	Coupling	Sol. type
Trobaugh et al. (1994a)	3 colinear cross-wires	Optical	A	Distilled water at room temp.	Closed-form
Péria et al. (1995)	Triangle	Optical	A	Water	Closed-form
Henry (1997)	Triangle	Optical	A	Water	Closed-form
Ionescu (1998)*	Triangle	Optical	A	Water	Iterative
Meairs et al. (2000)	Triangle + cross-wire	DC mag	A	Water at 37°C	Iterative

* Hook (2003) used a phantom and calibration protocol very similar to that of Ionescu (1998) to compare with his membrane phantom.

and "D" for a direct digital link. In the column "coupling" are listed the coupling media that were used for calibration itself. If a different medium was used for testing, it is not mentioned. Finally, the column "sol. type" classifies the minimization as closed-form or iterative. In each table, references are listed in chronological order.

Unfortunately, this classification does not mention whether or not the speed-of-sound problem was taken into account because many articles do not give enough details (refer to the Speed-of-sound issues section for more detail). The segmentation and temporal calibration were not included in the tables for many reasons. First, categorizing the segmentations as being simply manual, semiautomatic and automatic did not appear to be fair for all the semiautomatic methods that range from almost manual to almost automatic. Furthermore, many papers do not mention how the segmentation was done. As for the temporal calibration, it was only detailed in a few papers and, again, a one-word classification was not trivial. In addition, some temporal calibrations were not described in the same papers as the spatial calibration. For more details on these two topics, refer to the Feature extraction and Temporal calibration sections.

When is a recalibration necessary?

Recalibration is required when the sensor on the probe moves with respect to the probe. This situation can occur if the sensor is temporarily attached on the probe. A permanent sensor fixation is left on the probe between each use, for example, Comeau et al. (2000). Thus, the relative position of the sensor on the probe does not theoretically change. A temporary setup allows the sensor to be removed and reattached after each examination

(Amin et al. 2001; Lindseth et al. 2003c; Sato et al. 1998). A typical utilization of a temporary setup occurs during surgery, where the probe is often covered by a sterile drape because the probe cannot be sterilized. However, passive markers (optical tracking) used in some applications should not be covered. In these cases, the rigid body with the markers is clipped on the probe through the drape or the markers only are clipped through the drape on the rigid body. Ideally, the calibration should be redone every time the sensor is reattached on the probe, but this depends on how precisely it can be reattached. Lindseth et al. (2003c) and Gronningsaeter et al. (2000) eliminated the need for repeated probe calibration by using an adapter that ensured precise and repetitive attachment between the tracking device and the probe, even through the draping. Hence, they could perform the probe calibration once in the laboratory. Amin et al. (2001) evaluated in a quantitative manner the consistency of their calibration procedure when reattaching the sensor on the probe. Their results indicated that their fixation system was precise enough to avoid the necessity to recalibrate each time the sensor was reattached. Treece et al. (2003) found that the accuracy of their system was slightly lowered when remounting the sensor on the probe, going from a 3-D confidence limit of < 0.5 mm to < 0.69 mm.

When using an analog output, changing some settings on the US machine might also affect the calibration matrix. The most obvious is the depth setting that affects the scaling parameter of the calibration. Thus, when using multifrequency probes, either a calibration is executed for every depth setting or one calibration is done at a particular depth and the image is scaled for the other ones (Carr et al. 2000). The pan and zoom operations

Table 5. Multiple-point targets phantom category

References	Description	Tracking	Acq.	Coupling	Sol. type
Kowal et al. (2003)	4 1-mm needle pins	Optical	A	Water at room temp.	Iterative
Leotta (2004)	Planar array of strings and beads	DC mag	A	Water at room temp.	Closed-form

Table 6. 2-D shape alignment phantom category

References	Description	Tracking	Acq.	Coupling	Sol. type
Sato et al. (1998)	3 features	Optical	A	Water at 25°C	Iterative
Berg et al. (1999)	5 features	DC mag	D	Water at 33°C	Closed-form
Langø (2000)	5 features	Optical	D	Water at 22°C	Iterative
Welch et al. (2000, 2001a, 2001b)	3 features	Optical	A	Water	Closed-form
Kowal et al. (2003)	4 features	Optical	A	Water at room temp.	Iterative

also invalidate the original calibration. They affect the translational and scaling parameters (Treece et al. 2003) and, therefore, necessitate a recalibration.

CALIBRATION EVALUATION

This section summarizes the main techniques to evaluate the *in vitro* precision and accuracy of a calibration transformation. These tests are sometimes performed on the same phantoms as the one used for calibration (Detmer et al. 1994), but preferably on a different phantom (Lindseth et al. 2003c; Treece et al. 2003) to minimize bias. Precision is estimated when multiple measurements of the same phenomenon are compared with themselves. Accuracy is estimated when the measurements are compared with a "gold standard" or the best independent measure available. Stability and reproducibility are synonyms of precision. For a more complete discussion, refer to Lindseth et al. (2002; 2003c) and Treece et al. (2003).

Precision

As a simple measure of precision, some authors have done many calibrations with the same technique and provided descriptive statistics for the components of the calibration matrix, such as the standard deviations, root-mean-squares errors, 95% confidence intervals and ranges (Bouchet et al. 2001; Leotta 2004; Pagoulatos et al. 2001; Rousseau et al. 2003b; Treece et al. 2003). It is important to note that all these metrics based on a residual only address questions of precision and not accuracy. Another simple measure, sometimes referred to as calibration reproducibility (Blackall et al. 2000; Lindseth et al. 2003c), was proposed by Prager et al. (1998b) and consisted of comparing the position in world space of the bottom right corner of the image plane when using all the

Table 7. Three-wire phantom category

References	Tracking	Acq.	Coupling	Sol. type
Carr (1996)	AC mag	A	Water	Iterative
Prager et al. (1998b)	AC mag	A	Water at room temp.	Iterative
Carr et al. (2000)	AC mag	A	Water	Iterative
Kowal et al. (2003)	Optical	A	Water at room temp.	Iterative

different combinations of calibration transformations. The method was later used by Blackall et al. (2000); Rousseau et al. (2003b) and Lindseth et al. (2003c). Some authors took more than one point into account: Kowal et al. (2003) considered the two bottom corners, Treece et al. (2003) considered the four corners plus the center of the image, Pagoulatos et al. (2001) considered nine points distributed on the center line of the B-scans and Leotta (2004) considered points specified every 2 cm in depth along the center and the edges of the image up to 16 cm. These two precision measurements are interesting because they can be applied to all techniques because they do not require any test phantom.

A method proposed by Detmer et al. (1994) involves imaging a cross-wire or point target from multiple viewing angles. The generated "points" are then extracted from each image and mapped in world space, ultimately forming a cloud of points. The tightness (or spread) of this cloud is often used to estimate what is called the reconstruction precision. The method has been used frequently (Blackall et al. 2000; Boctor et al. 2003; Hartov et al. 1999; Kowal et al. 2003; Leotta 2004; Leotta et al. 1997; Pagoulatos et al. 1998; Prager et al. 1998b).

Accuracy

Similarly to the experiment to evaluate the reconstruction precision with a cross-wire or point target, it is also possible to estimate the point reconstruction accuracy if the position of the point object is known in world space. Each point in the cloud of points is then compared to the "gold standard" value and residual error vectors can be computed (Blackall et al. 2000; Lindseth et al. 2003c; Muratore and Galloway 2001; Pagoulatos et al. 2001; Trobaugh et al. 1994a). Again, the same principle can be applied to more complex phantoms. Pagoulatos et al. (2001) used their calibration z-fiducials phantom for testing accuracy as well. A phantom with 27 wire-crossings organized in the shape of a cube has also been used to evaluate accuracy (Langø 2000; Lindseth et al. 2003a, 2002). Because the wire-crossings of this phantom were extracted from a reconstructed volume instead of on 2-D images, this technique evaluated the 3-D point reconstruction accuracy. Lindseth et al. (2003a) automated the

Table 8. Wall phantom category

References	Description	Tracking	Acq.	Coupling	Sol. type
Prager et al. (1998b)	Single-wall	AC mag	A	Water at room temp.	Iterative
Prager et al. (1998b)	Cambridge	AC mag	A	Water at room temp.	Iterative
Langø (2000)	Membrane	Optical	D	Water	Iterative
Rousseau et al. (2002, 2003a, 2003b)	Single-wall	DC mag	A	Water*	Iterative
Hook (2003)	Membrane	Optical	A	Water at room temp.	Iterative
Kowal et al. (2003)	Cambridge	Optical	A	Water at room temp.	Iterative
Treece et al. (2003)	Cambridge	Optical	D	Water at 50°C	Iterative

* For these publications, water at room temperature was used without any speed-of-sound correction. Later, for his thesis, Rousseau (2003) did some experiments for the speed-of-sound problem that were mentioned in the Speed-of-sound issues section.

precision and accuracy tests done with this phantom and these tests could thereby be performed quickly for an extensive data set based on a vast number of experimental conditions to thoroughly investigate and validate a 3-D US system. Finally, with a single- or multiple-point objects phantom, it is also possible to evaluate the system bias by computing the mean residual vector (Langø 2000; Lindseth et al. 2003a).

King et al. (1991) evaluated the accuracy of their 3-D US system by measuring the volume and interfeature distances of a testing phantom consisting of pins mounted in the shape of a cylinder sector. Distance and volume measurements have frequently been used since then and are often considered to estimate the reconstruction accuracy. A popular phantom for measuring distances has been a 4 × 2 matrix of small spheres (Blackall et al. 2000; Kowal et al. 2003; Legget et al. 1998; Leotta et al. 1997; Welch et al. 2002) or pins (Barratt et al. 2001a; Prager et al. 1998b). Pagoulatos et al. (2001) and Leotta (2004) used a similar phantom, but with a grid of six beads. The test phantom used by Treece et al. (2003) was made of a tissue-mimicking substance that contained a planar array of 110 2-mm spheres. Boctor et al. (2003) used their Hopkins wire phantom that was also used for calibration. Lindseth et al. (2003c) used their 27 wire-crossings phantom to measure all distances between wire crossings but, again, extracted the features from the 3-D reconstructed volume instead of on the images. They

called that measure the 3-D distance reconstruction accuracy. For volume measurements, water-filled balloons immersed in water or a tissue-mimicking solution were a simple solution adopted by many (Barry et al. 1997; Berg et al. 1999; Pagoulatos et al. 1998). Rousseau et al. (2003b) preferred a manufactured 3-D US phantom (CIRS, Inc., Norfolk, VA) made of a tissue-mimicking material (Zerdine) and containing two ellipsoids.

Other aspects of calibration evaluation

For phantoms where the optimization process estimates both phantom location and spatial calibration parameters simultaneously, the precise movement of the probe while scanning is crucial to the resulting accuracy of the calibration parameters. If this type of phantom is scanned from only one direction, the residual error in the least-squares minimization process may be extremely small, yet the calibration parameters will be highly unconstrained by the minimization process and likely to be inaccurate. Equally, if a point target is always imaged at the midpoint of the B-scans, then, even if the probe orientation is changed, the calibration parameters or the precision measurements will only be valid for the B-scan midpoint and errors at the corners of the B-scans are likely to be much larger.

Hence, a good calibration with these phantoms requires both the exercising of all the degrees of freedom

Table 9. Z-fiducials phantom category

References	*n* of Zs	Tracking	Acq.	Coupling	Sol. type
Comeau et al. (1998)	3	Mechanical arm	A	Water + glycerin	Closed-form
Gobbi et al. (1999)	4	Optical	A	Water at room temp.*	Closed-form
Pagoulatos et al. (1999, 2001)	30	DC mag	A	Distilled water	Closed-form
Comeau et al. (2000)	4	Optical	A	Water	Closed-form
Bouchet et al. (2001)	13	Optical	A	Tissue-mimicking	Both
Zhang et al. (2002)	15	Optical	D	Water at room temp.	Closed-form
Lindseth et al. (2003c)	12	Optical	D	Water	Closed-form

* For this publication, water at room temperature was used without any speed-of-sound correction. Later, for his thesis, Gobbi (2003) used a solution of 10% glycerol in water by volume.

Table 10. Other wire phantom category

References	Tracking	Acq.	Coupling	Sol. type
Liu et al. (1998)	DC mag	A	Water	–
Beasley et al. (1999)	Optical	A	Water at room temp.	Closed-form
Boctor et al. (2003)	DC mag	A	Water at 50°C	Iterative
Lindseth et al. (2003c)	Optical	D	Water	Closed-form

of the probe orientation and also ensuring that the target is seen in all areas of the US image.

RESULTS

It would be very practical if all authors would agree on which calibration technique is the best. Yet, the criteria for defining what is best differs among groups, depending mostly on the type of application of the tracked US images. It would probably be safe to say that most authors agree on the fact that accuracy and precision are the most important evaluation criteria. However, as was seen in the previous section, these criteria can be evaluated in many different ways. Other factors, such as the time and the space required to calibrate, can be important if, for example, the calibration must be done repetitively in an operating room (Kowal et al. 2003). In contrast, the time required for calibration can also be unimportant if calibration is done only once in the laboratory (Lindseth et al. 2003c).

Thus, it is difficult to compare results on the performance of calibration methods from different publications, because they depend on many factors, such as the type of accuracy estimated, the method for calculating the accuracy, the type of tracking device, the depth setting, the range of scanning angles and positions, the quality of the US equipment, etc. Treece et al. (2003) have made an attempt to compare the results obtained by different groups by graphically plotting the 3-D confidence limits for some precision and accuracy measurements. Here, we summarize results only of comparative studies of phantom designs in the text below and in Table 12. We felt that it was too difficult to compare individual published calibration results, due to the reasons listed above. Although there may be certain biases in these comparative studies, they were completed under controlled experimental conditions so as to enable direct comparison of techniques analyzed. However, interpretation of Table 12 should be made with caution, because

of the different experimental methods used by the authors.

Prager et al. (1998b) were the first to compare many types of calibration phantoms. They compared their new single-wall and Cambridge phantoms with the three-wire and single cross-wire phantoms. They concluded that the single cross-wire and Cambridge phantoms produced better results. More specifically, the Cambridge phantom obtained the best r.m.s. error and calibration reproducibility and the cross-wire phantom, the best reconstruction precision and reconstruction accuracy by distance measurements. In terms of time required for the calibration, the Cambridge phantom allowed for a much quicker calibration than the cross-wire phantom.

Blackall et al. (2000) compared their new multimodal registration phantom with the single cross-wire phantom in terms of reproducibility, reconstruction precision, point reconstruction accuracy and reconstruction accuracy by distance measurements. Referring to these four criteria, the single cross-wire phantom appears to be generally better, mostly in terms of reproducibility and reconstruction precision. However, their new method had the advantage of being fully automatic and quicker.

Langø (2000) compared a 2-D shape alignment phantom with a new membrane phantom (one of the wall techniques), evaluating the reconstruction precision (spread) and accuracy (bias). He found that the new membrane phantom had an overall better performance compared with the 2-D shape-alignment technique. Both techniques had a similar spread, but the membrane method resulted in a lower bias of the reconstructed points in 3-D.

Four comparison papers were published in 2003. Boctor et al. (2003) compared their new Hopkins phantom (other wire phantom category) with a single cross-wire phantom. The Hopkins phantom performed better in terms of reconstruction accuracy by distance measurements and the cross-wire phantom was slightly better in

Table 11. Multimodal registration phantom category

References	Tracking	Acq.	Coupling	Sol. type
Blackall et al. (2000)	Optical	A	Gelatin phantom + thin layer of water	Iterative

Ultrasound in Medicine and Biology Volume 31, Number 2, 2005

Table 12. Summary of phantom comparisons

		Precision (mm)			Accuracy (mm)	
		r.m.s. error	Mean calibration reproducibility	Mean point reconstruction precision	Mean point reconstruction accuracy	Mean reconstruction accuracy (distances)
Prager et al. (1998b)	Cross-wire	0.56	1.47	0.04	–	0.04 ± 1.12
	Three-wire	1.04	5.37	−0.15	–	−0.15 ± 2.18
	Single-wall	0.48	3.27	0.14	–	0.14 ± 1.63
	Cambridge	0.34	0.92	0.23	–	0.23 ± 1.33
Blackall et al. (2000)	Cross-wire	–	1.05 ± 0.43	0.80 ± 0.46	1.15 ± 0.40	−0.00019 ± 0.60
	Registration phantom	–	1.84 ± 1.26	1.15 ± 0.62	1.16 ± 0.45	−0.025 ± 0.69
Boctor et al. (2003)	Cross-wire	–	–	0.62 ± 0.29*	–	0.25 ± 1.78*
	Hopkins	–	–	0.72 ± 0.343*	–	0.15 ± 1.63*
Kowal et al. (2003)	Three-wire	0.221	3.2 ± 1.94	2.3 ± 1.23	–	0.3 ± 0.49
	Cambridge	0.160	2.2 ± 2.74	2.4 ± 1.38	–	0.3 ± 0.53
	Pin cage	0.135	2.7 ± 1.59	2.5 ± 1.36	–	0.3 ± 0.58
	Wedge cage	0.151	1.9 ± 1.23	2.2 ± 1.28	–	0.3 ± 0.51
Lindseth et al. (2003c)[†]	Single-point target	–	0.63 ± 0.39 (P) 0.62 ± 0.38 (L)	–	2-D: 0.79 ± 0.39 (P) 2-D: 0.73 ± 0.41 (L) 3-D: 1.00 ± 0.39 (P) 3-D: 1.48 ± 0.35 (L)	3-D: 0.15 ± 0.30 (P) 3-D: 0.23 ± 0.51 (L)
	Diagonal phantom	–	0.38 ± 0.17 (P) 0.44 ± 0.25 (L)	–	2-D: 0.86 ± 0.46 (P) 2-D: 0.77 ± 0.43 (L) 3-D: 0.84 ± 0.36 (P) 3-D: 1.24 ± 0.71 (L)	3-D: 0.10 ± 0.30 (P) 3-D: 0.26 ± 0.46 (L)
	Z-fiducials	–	0.55 ± 0.29 (P) 0.63 ± 0.36 (L)	–	2-D: 1.52 ± 1.35 (P) 2-D: 1.03 ± 0.84 (L) 3-D: 0.81 ± 0.43 (P) 3-D: 1.15 ± 0.43 (L)	3-D: 0.16 ± 0.33 (P) 3-D: 0.25 ± 0.45 (L)
Leotta (2004)	Single-point target	–	–	0.94	–	−0.10 ± 0.70
	Multiple point target	–	–	0.96*	–	−0.10 ± 0.68*

* These are the average values for the all trials.

[†] Lindseth et al. (2003c) performed their tests on three types of probes, but only partially with the intraoperative probe. Consequently, only the results obtained with the flat phased-array (P) and the flat linear-array (L) probes are reported here. Second, some measures were done on images (2-D) and others on reconstructed volumes (3-D). Both types of measures are reported here.

terms of reconstruction precision. Using three different probes, Hook (2003) wrote a report comparing the performances of a membrane phantom and a multiple cross-wire phantom described by Ionescu (1998) from the same laboratory. He concluded that the membrane phantom produced smaller r.m.s. errors, smaller maximum errors and a more centralized error distribution. Lindseth et al. (2003c) also used three types of probes to compare three phantoms; they were a single point target, a pyramid-shaped z-fiducials and a phantom in the other wire phantom category, which they called the Diagonal phantom. Five quality measures were obtained, but the 3-D point reconstruction accuracy, referred to as the 3-D navigation accuracy, was considered to be the most important. In terms of this criterion, the z-fiducials phantom performed best, followed by the Diagonal phantom. However, the z-fiducial phantom was not superior according to the four other criteria. The Diagonal phantom had the best calibration reproducibility and the single point target, the best point reconstruction accuracy. Kowal et al. (2003) proposed two new phantoms; they were the Four Edge with pin cage (multiple point target category) and the Four Edge with wedge cage (2-D alignment category). They were compared to a three-wire and a Cambridge phantom. The three-wire phantom had a weaker performance in terms of mean r.m.s. error and variance. The two Four Edge methods were considerably quicker and required less space. The Four Edge with wedge cage phantom also had a slightly better calibration reproducibility. Nevertheless, the difference between the phantoms was not significant when looking at the reconstruction precision and the reconstruction accuracy by distance measurements. Finally, Leotta (2004) compared a single-point target phantom with their new multiple-point target phantom. Both phantoms were similar in terms of point reconstruction precision and reconstruction accuracy by distance measurements.

CONCLUSION

US probe calibration is an obligatory step to build 3-D volumes from 2-D images acquired in a freehand US system. Calibration finds the transformation that relates the image plane to the sensor attached on the probe.

Many authors have included a brief literature review in their articles, but this paper is the first comprehensive overview of what has been done in the field. Calibration is a process with multiple components that were covered in the various sections of this review. A very brief summary of each of them follows.

The most popular tracking systems used for 3-D US are the electromagnetic and optical technologies. The optical tracking systems allow more accurate calibrations, but are also more expensive.

The traditional method for acquiring images is by digitizing them with a frame-grabber. Although digital images appear to have more potential, no studies have been published to confirm whether or not using digital images really improves the performance of calibration.

A number of different spatial calibration phantoms were described and classified. Each design has advantages and disadvantages over the others in terms of ease of use, accuracy, precision, number of images required, etc. There is no agreement as to which phantom design is the best.

Speed-of-sound-related errors in calibration can be significant, particularly when imaging at large depth settings where the distortion is greater. Simple solutions were presented to calibrate in a medium with a speed of sound similar to the one assumed by most US machines (1540 m/s).

Early on, feature extraction in the US images was done manually, but now this is increasingly being automated. Lines in US images appear to be easiest to segment but, if the quality of the images is sufficient, points can also be automatically extracted.

There exist two approaches to resolve the overdetermined system of equations of the calibration problem, the iterative or the noniterative (closed-form) solutions. Some types of phantoms can be resolved by both, but others are more naturally resolved by one or the other. The minimization technique does not seem to have a major impact on precision or accuracy; however, iterative techniques may suffer from local minima.

The tracking system and the US machine are not synchronized in time. Moving the probe very slowly minimizes the positional error for every B-scan. However, faster movement is required during real acquisitions such as, for example, during surgery. Temporal calibration allows a better coherence of the positional and image streams, potentially improving the quality of the calibration and also enabling more rapid movement during acquisition.

Finally, many criteria have been proposed to evaluate the quality of a calibration technique. Precision and accuracy are probably the most important criteria and multiple approaches to test these criteria were presented in the review. Some precision measurements do not require any phantom equipment and, hence, could be used to compare calibration techniques between different groups.

3-D US is more powerful than 2-D US, mainly because it facilitates understanding of the anatomy. Compared with other imaging modalities, it has the important advantage of being real-time, safe and inexpensive. Calibration is an essential step to acquire tracked 2-D images and has a major impact on the quality of the reconstructed volume to enable precise 3-D visualization, planning and accurate image-guided interventions.

REFERENCES

Amin DV, Kanade T, Jaramaz B, et al. Calibration method for determining the physical location of the ultrasound image plane. In: Lecture notes in computer science, MICCAI. Vol. 2208. Utrecht, The Netherlands: Springer, 2001:940.

Anderson ME, McKeag MS, Trahey GE. The impact of sound speed errors on medical ultrasound imaging. J Acoust Soc Am 2000;107:3540–3548.

Arbel T, Morandi X, Comeau RM, et al. Automatic non-linear MRI-ultrasound registration for the correction of intra-operative brain deformations. In: Lecture notes in computer science, MICCAI. Vol. 2208. Utrecht, The Netherlands: Springer, 2001:913–922.

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

Barratt DC, Davies AH, Hughes AD, et al. Accuracy of an electromagnetic three-dimensional ultrasound system for carotid artery imaging. Ultrasound Med Biol 2001a;27:1421–1425.

Barratt DC, Davies AH, Hughes AD, et al. Optimisation and evaluation of an electromagnetic tracking device for high-accuracy three-dimensional ultrasound imaging of the carotid arteries. Ultrasound Med Biol 2001b;27:957–968.

Barry CD, Allott CP, John NW, et al. Three-dimensional freehand ultrasound: Image reconstruction and volume analysis. Ultrasound Med Biol 1997;23:1209–1224.

Beasley RA, Stefansic JD, Herline AJ, et al. Registration of ultrasound images. SPIE Proc 1999;3658:125–132.

Berg S, Torp H, Martens D, et al. Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data. Ultrasound Med Biol 1999;25:745–753.

Besl P, McKay N. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Machine Intell 1992;14:239–256.

Bilaniuk N, Wong G. Speed of sound in pure water as a function of temperature. J Acoust Soc Am 1993;93:1609–1612.

Birkfellner W, Watzinger F, Wanschitz F, et al. Systematic distortions in magnetic position digitizers. Med Phys 1998;25:2242–2248.

Blackall JM, Rueckert D, Maurer CR, et al. An image registration approach to automated calibration for freehand 3D ultrasound. In: Lecture notes in computer science. Pittsburgh, PA: Springer, 2000;462–471.

Boctor EM, Jain A, Choti M, et al. A rapid calibration method for registration and 3D tracking of ultrasound images using spatial localizer. SPIE Proc 2003;5035:521–532.

Bouchet LG, Meeks SL, Goodchild G, et al. Calibration of three-dimensional ultrasound images for image-guided radiation therapy. Phys Med Biol 2001;46:559–577.

Brendel B, Winter S, Rick A, et al. Registration of 3D CT and ultrasound datasets of the spine using bone structures. Comput Aided Surg 2002;7:146–155.

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

Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Machine Intell 1986;8:679–698.

Carr JC. Surface reconstruction in 3D medical imaging. Christchurch: University of Canterbury, 1996.

Carr JC, Stallkamp JL, Fynes MM, et al. Design of a clinical free-hand 3D ultrasound system. SPIE Proc 2000;3982:14–25.

Chassat F, Lavallée S. An experimental protocol for accuracy evaluation of 6D localizers for computer-assisted surgery: Application to four optical localizers. In: Lecture notes in computer science, MICCAI. Vol. 1496. Cambridge, MA: Springer, 1998:277–284.

Cinquin P, Bainville E, Barbe C, et al. Computer assisted medical interventions. IEEE Eng Med Biol Magazine 1995;14:254–263.

Comeau RM, Fenster A, Peters TM. Integrated MR and ultrasound imaging for improved image guidance in neurosurgery. SPIE Proc 1998;3338:747–754.

Comeau RM, Sadikot AF, Fenster A, et al. Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery. Med Phys 2000;27:787–800.

Crivianu-Gaita D, Miclea F, Gaspar A, et al. 3D reconstruction of prostate from ultrasound images. Int J Med Inf 1997;45:43–51.

Deriche R. Optimal edge detection using recursive filtering. First International Conference on Computer Vision, London. 1987:501–505.

Deriche R. Fast algorithms for low-level vision. IEEE Trans Pattern Anal Machine Intell 1990;12:78–87.

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

Dey D, Gobbi DG, Slomka PJ, et al. Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: Creating stereoscopic panoramas. IEEE Trans Med Imaging 2002;21:23–30.

Dudley NJ, Gibson NM, Fleckney MJ, et al. The effect of speed of sound in ultrasound test objects on lateral resolution. Ultrasound Med Biol 2002;28:1561–1564.

Eggert DW, Lorusso A, Fisher RB. Estimating 3-D rigid body transformations: A comparison of four major algorithms. Machine Vision Applic 1997;9:272–290.

Fenster A, Downey DB. Three-dimensional ultrasound imaging. Annu Rev Biomed Eng 2000;2:457–475.

Fischler MA, Bolles RC. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Comm ACM 1981;24:381–395.

Gobbi DG, Comeau RM, Peters TM. Ultrasound probe tracking for real-time ultrasound/MRI overlay and visualization of brain shift. In: Lecture notes in computer science, MICCAI. Vol. 1679. Cambridge, UK: Springer, 1999:920–927.

Gobbi DG. Brain deformation correction using interactive 3D ultrasound imaging. Ph.D. thesis. University of Western Ontario, London, ONT, Canada, 2003.

Goldstein A. The effect of acoustic velocity on phantom measurements. Ultrasound Med Biol 2000;26:1133–1143.

Goss SA, Johnston RL, Dunn F. Comprehensive compilation of empirical ultrasonic properties of mammalian tissues. J Acoust Soc Am 1978;64:423–457.

Gronningsaeter A, Kleven A, Ommedal S, et al. SonoWand, an ultrasound-based neuronavigation system. Neurosurgery 2000;47:1373–1380.

Hartov A, Eisner SD, Roberts DW, et al. Error analysis for a free-hand three-dimensional ultrasound system for neuronavigation. Neurosurg Focus 1999;6:.

Henry D. Outils pour la modélisation de structures et la simulation d'examens échographiques. Ph.D. thesis, Université Joseph Fourier, Grenoble, 1997.

Hoffmann AL, Laguna MP, de la Rosette JJ, et al. Quantification of prostate shrinkage after microwave thermotherapy: A comparison of calculated cell-kill versus 3D transrectal ultrasound planimetry. Eur Urol 2003;43:181–187.

Hook I. Probe calibration for 3D ultrasound image localization. Internship report. Grenoble, France, 2003.

Hooke R, Jeeves TA. "Direct search" solution of numerical and statistical problems. J Assoc Comput Mach 1961;8:212–229.

Horn BKP. Closed-form solution of absolute orientation using unit quaternions. J Opt Soc Am A 1987;4:629–642.

Horn BKP, Hilden HM, Negahdaripour S. Closed-form solution of absolute orientation using orthonormal matrices. J Opt Soc Am A 1988;5:1127–1135.

Hughes SW, D'Arcy TJ, Maxwell DJ, et al. Volume estimation from multiplanar 2D ultrasound images using a remote electromagnetic position and orientation sensor. Ultrasound Med Biol 1996;22:561–572.

Ionescu G. Segmentation et recalage d'images échographiques par utilisation de connaissances physiologiques et morphologiques. Ph.D. thesis, Université Joseph Fourier, Grenoble, 1998.

Kanatani K. Analysis of 3-D rotation fitting. IEEE Trans Pattern Anal Machine Intell 1994;16:543–549.

Kelly IM, Gardener JE, Brett AD, et al. Three-dimensional US of the fetus. Work in progress. Radiology 1994;192:253–259.

Khadem R, Yeh CC, Sadeghi-Tehrani M, et al. Comparative tracking error analysis of five different optical tracking systems. Comput Aided Surg 2000;5:98–107.

Kindratenko VV. A survey of electromagnetic position tracker calibration techniques. Virtual Reality Res Dev Applic 2000;5:169–182.

King DL, King DL Jr, Shao MY. Evaluation of in vitro measurement accuracy of a three-dimensional ultrasound scanner. J Ultrasound Med 1991;10:77–82.

Kowal J, Amstutz CA, Caversaccio M, et al. On the development and comparative evaluation of an ultrasound B-mode probe calibration method. Comput Aided Surg 2003;8:107–119.

Langø T. Ultrasound guided surgery: Image processing and navigation. Ph.D. thesis, Norwegian University of Science and Technology, Trondheim, 2000.

Legget ME, Leotta DF, Bolson EL, et al. System for quantitative three-dimensional echocardiography of the left ventricle based on a magnetic-field position and orientation sensing system. IEEE Trans Biomed Eng 1998;45:494–504.

Leotta DF. An efficient calibration method for freehand 3-D ultrasound imaging systems. Ultrasound Med Biol 2004;30:999–1008.

Leotta DF, Detmer PR, Gilja OH, et al. Three-dimensional ultrasound imaging using multiple magnetic tracking systems and miniature magnetic sensors. IEEE Ultrason Sympos 1995;:1415–1418.

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.

Li PC, Li CY, Yeh WC. Tissue motion and elevational speckle decorrelation in freehand 3D ultrasound. Ultrason Imaging 2002;24:1–12.

Light ED, Davidsen RE, Fiering JO, et al. Progress in two-dimensional arrays for real-time volumetric imaging. Ultrason Imaging 1998;20:1–15.

Lindseth F, Bang J, Langø T. A robust and automatic method for evaluating accuracy in 3-D ultrasound-based navigation. Ultrasound Med Biol 2003a;29:1439–1452.

Lindseth F, Kaspersen JH, Ommedal S, et al. Multimodal image fusion in ultrasound-based neuronavigation: Improving overview and interpretation by integrating preoperative MRI with intraoperative 3D ultrasound. Comput Aided Surg 2003b;8:49–69.

Lindseth F, Langø T, Bang J, et al. Accuracy evaluation of a 3D ultrasound-based neuronavigation system. Comput Aided Surg 2002;7:197–222.

Lindseth F, Tangen GA, Langø T, et al. Probe calibration for freehand 3-D ultrasound. Ultrasound Med Biol 2003c;29:1607–1623.

Liu J, Gao X, Zhang Z, et al. A new calibration method in 3D ultrasonic imaging system. 20th Annual Conference. IEEE Eng Med Biol 1998;20:839–841.

Lloret D, Serrat J, Lopez AM, et al. Motion-induced error correction in ultrasound imaging. In: 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT 2002), Padova, Italy, 2002.

Martin K, Spinks D. Measurement of the speed of sound in ethanol/water mixtures. Ultrasound Med Biol 2001;27:289–291.

Meairs S, Beyer J, Hennerici M. Reconstruction and visualization of irregularly sampled three- and four-dimensional ultrasound data for

cerebrovascular applications. Ultrasound Med Biol 2000;26:263–272.

Meyer K, Biocca FA. A survey of position trackers. Presence: Teleoperators and virtual environments. 1992;1:173–200.

More JJ. The Levenberg–Marquardt algorithm: Implementation and theory. In: Lecture notes in mathematics. Vol. 630. 1977:105–116.

Muratore DM, Galloway RL. Beam calibration without a phantom for creating a 3-D freehand ultrasound system. Ultrasound Med Biol 2001;27:1557–1566.

Nakamoto M, Sato Y, Nakada K, et al. A temporal calibration method for freehand 3D ultrasound system: A preliminary result. London, UK: CARS, 2003.

Nelson TR, Pretorius DH. Interactive acquisition, analysis and visualization of sonographic volume data. Int J Imag Syst Tech 1997;8:26–37.

Nelson TR, Pretorius DH. Three-dimensional ultrasound imaging. Ultrasound Med Biol 1998;24:1243–1270.

Pagoulatos N, Edwards WS, Haynor DR, et al. Calibration and validation of free-hand 3D ultrasound systems based on DC magnetic tracking. SPIE Proc 1998;3335:59–71.

Pagoulatos N, Haynor DR, Kim Y. Fast calibration for 3D ultrasound imaging and multimodality image. First Joint BMES/EMBS Conference (Serving Humanity, Advancing Technology), Atlanta, GA, 1999. Proc IEEE 1999;:1065.

Pagoulatos N, Haynor DR, Kim Y. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. Ultrasound Med Biol 2001;27:1219–1229.

Pennec X, Cachier P, Ayache N. Tracking brain deformation in time sequences of 3D US images. Pattern Recog Lett 2003;24:801–813.

Péria O, Chevalier L, François-Joubert A, et al. Using a 3D position sensor for registration of SPECT and US images of the kidney. Lecture Notes Comput Sci 1995;905:23–29.

Prager RW, Gee A, Berman L. Stradx: Real-time acquisition and visualization of freehand three-dimensional ultrasound. Med Image Anal 1998a;3:129–140.

Prager RW, Gee AH, Treece GM, et al. Sensorless freehand 3-D ultrasound using regression of the echo intensity. Ultrasound Med Biol 2003;29:437–446.

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

Rohling R, Fung W, Lajevardi P. PUPIL. Programmable ultrasound platform and interface library, MICCAI 2003. In: Lecture Notes in Computer Science. Vol. 2879. Montreal, QUE, Canada: Springer, 2003:424–431.

Rohling R, Munger P, Hollerbach JM, et al. Comparison of relative accuracy between a mechanical and an optical position tracker for image-guided neurosurgery. J Image Guided Surg 1995;1:30–34.

Rohling RN, Gee AH, Berman L. Three-dimensional spatial compounding of ultrasound images. Med Image Anal 1997;1:177–193.

Rousseau F. Méthodes d'analyse d'images et de calibration pour l'échographie 3D en mode main-libre. Ph.D. thesis, Université de Renne I, Renne, 2003.

Rousseau F, Barillot C. Quality assessment of electromagnetic localizers in the context of 3D ultrasound. INRIA 2002.

Rousseau F, Hellier P, Barillot C. A fully automatic calibration procedure for freehand 3D ultrasound. IEEE International Symposium on Biomedical Imaging, Washington DC, 2002.

Rousseau F, Hellier P, Barillot C. Méthode de calibration pour système échographique 3D main-libre. Congrès Francophone de Vision par Ordinateur. Gerardmer, France: ORASIS'03, 2003a:7–14.

Rousseau F, Hellier P, Barillot C. Robust and automatic calibration method for 3D freehand ultrasound, MICCAI 2003. In: Lecture notes in computer science. Vol. 2879. Montreal, QUE, Canada: Springer, 2003b:440–448.

Sato Y, Nakamoto M, Tamaki Y, et al. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. IEEE Trans Med Imaging 1998;17:681–693.

Schmerber S, Chassat F. Accuracy evaluation of a CAS system: Laboratory protocol and results with 6D localizers, and clinical experiences in otorhinolaryngology. Comput Aided Surg 2001;6:1–13.

Shoemake K. Animating rotations with quaternion curves. ACM SIGGRAPH Comput Graph 1985;19:245–254.

State A, Chen DT, Brandt A, et al. Case study: Observing a volume rendered fetus within a pregnant patient. In: Proceedings of IEEE Visualization. Los Alamitos, CA: IEEE Computer Society Press, 1994;364–368.

Treece GM, Gee AH, Prager RW, et al. High-definition freehand 3-D ultrasound. Ultrasound Med Biol 2003;29:529–546.

Trobaugh JW, Richard WD, Smith KR, et al. Frameless stereotactic ultrasonography: Method and applications. Comput Med Imaging Graph 1994a;18:235–246.

Trobaugh JW, Trobaugh DJ, Richard WD. Three-dimensional imaging with stereotactic ultrasonography. Comput Med Imaging Graph 1994b;18:315–323.

Tuthill TA, Krucker JF, Fowlkes JB, et al. Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. Radiology 1998;209:575–582.

Umeyama S. Least-squares estimation of transformation parameters between two points patterns. IEEE Trans Pattern Anal Machine Intell 1991;13:376–380.

Unsgaard G, Ommedal S, Muller T, et al. Neuronavigation by intraoperative three-dimensional ultrasound: Initial experience during brain tumor resection. Neurosurgery 2002;50:804–812.

Walker MW, Shao L, Volz RA. Estimating 3-D location parameters using dual number quaternions. CVGIP: Image Understanding 1991;54:358–367.

Welch JN, Bax MR, Mori K, et al. A fast and accurate method of ultrasound probe calibration for image-guided surgery. 16th International Congress of Computer Aided Radiology and Surgery (CARS), Paris, France, 2002:1078.

Welch JN, Johnson JA, Bax MR, et al. A real-time freehand 3D ultrasound system for image-guided surgery. Delft, Holland: Ultrasonics International, 2001a;1601–1604.

Welch JN, Johnson JA, Bax MR, et al. Real-time freehand 3D ultrasound system for clinical applications. Visualiz Display Image-Guided Procedures 2001b;4319:724–730.

Welch N, Johnson JA, Bax MR, et al. A real-time freehand 3D ultrasound system for image-guided surgery. Ultrasonic Symposium 2000;2:1601–1604.

West JB, Maurer CR Jr. Designing optically tracked instruments for image-guided surgery. IEEE Trans Med Imaging 2004;23:533–545.

Wiles AD, Thompson DG, Frantz DD. Accuracy assessment and interpretation for optical tracking systems. SPIE Proc 2004;5367.

Zhang Y, Rohling R, Pai DK. Direct surface extraction from 3D freehand ultrasound images. IEEE Visualization 2002;:45–52.

Zhang Y. Direct surface extraction from 3D freehand ultrasound images. Ph.D. thesis, University of British Columbia, Vancouver, 2003.