

---

**Freehand 3D Ultrasound  
Calibration: A Review**

P-W. Hsu, R. W. Prager  
A. H. Gee and G. M. Treece

**CUED/F-INFENG/TR 584**

December 2007

University of Cambridge  
Department of Engineering  
Trumpington Street  
Cambridge CB2 1PZ  
United Kingdom

Email: [pwh24@cam.ac.uk](mailto:pwh24@cam.ac.uk), [rwp/ahg/gmt11@eng.cam.ac.uk](mailto:rwp/ahg/gmt11@eng.cam.ac.uk)

---

# Freehand 3D Ultrasound Calibration: A Review

Po-Wei Hsu, Richard W. Prager, Andrew H. Gee and Graham M. Treece

University of Cambridge  
Department of Engineering  
Trumpington Street  
Cambridge CB2 1PZ

## Abstract

Freehand three-dimensional ultrasound is a technique for acquiring ultrasonic data of a 3D volume by recording the trajectory of the ultrasound probe using a position sensor. In planning and registration, a freehand ultrasound systems is used to track a two-dimensional probe. Probe calibration is necessary to find the rigid body transformation from the coordinate system of the B-scan to that of the mobile part of the position sensor. Numerous techniques for this have been developed over the past decade. In this review, we give a comprehensive description of existing calibration techniques and classify them according to the mathematical principles on which they are based. We give a thorough analysis of these approaches based on their accuracy, ease of use, reliability, and speed of calibration. To ensure consistency, these comparisons are done by the authors based on experimental results and not on figures quoted in previous papers.

## 1 Introduction

Three dimensional (3D) ultrasound imaging is a medical imaging modality that allows the clinician to obtain a 3D model of the anatomy, possibly in real-time (Nelson and Pretorius, 1998; Fenster et al., 2001). A 3D ultrasound imaging system has many clinical applications, some of these are: obstetrics (Gonçalves et al., 2005), gynecology (Alcazar, 2005), breast biopsy (Fenster et al., 2004b), cardiology (Fenster et al., 2004a), fetal cardiology (Yagel et al., 2007), neurosurgery (Unsgaard et al., 2006), radiology (Meeks et al., 2003) and surgery (Rygh et al., 2006).

One of the techniques to build a 3D ultrasonic system is to track a position sensor attached to a conventional ultrasound probe—a freehand 3D ultrasound system (Gee et al., 2003). This technique allows a large volume to be recorded and visualized in a fixed global coordinate system. As the probe is swept over the anatomy, the trajectory of the probe is recorded by the attached position sensor. The volume of the anatomy can be constructed by matching the ultrasonic data with its corresponding position in space. In addition to acquiring a 3D volume, freehand ultrasound systems also have applications in planning (Eulenstein et al., 2004) and registration (Penney et al., 2004).

However, the mobile part of the position sensor records the 3D location of the sensor  $S$ , rather than the scan plane  $I$ , relative to its stationary counterpart  $W$  as shown in Figure 1. It is therefore necessary to find the position and orientation of the scan plane with respect to the electrical centre of the position sensor. This rigid-body transformation  $T_{S \leftarrow I}$  comprises six parameters—three translations in the direction of the  $x$ ,  $y$  and  $z$ -axes and the three rotations, azimuth, elevation and roll, about these axes. This transformation is determined through a process called probe calibration (Mercier et al., 2005). The stationary part of the position sensor is often called the world coordinate system, and the term position sensor is used to mean its mobile counterpart. We will also follow these conventions in this paper. In general, a transformation involves both a rotation and a translation in 3D space. For brevity, we will use the notation  $T_{B \leftarrow A}$  to mean a rotational transformation followed by a translation from the coordinate system  $A$  to coordinate system  $B$ .

Another issue before we can construct a volume in space is to determine the scales of the B-scans. A point  $p^I = (u, v, 0)^t$  in a B-scan image, where  $u$  and  $v$  are the column and row indices, typically has units in pixels rather than in millimetres. A scaling factor  $T_s = \begin{pmatrix} s_u & 0 & 0 \\ 0 & s_v & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ,

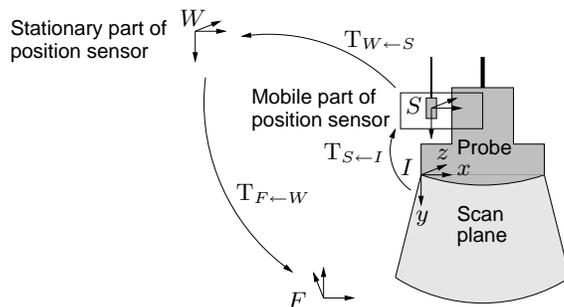


Figure 1: The coordinates associated with a freehand 3D ultrasound system.

where  $s_u$  and  $s_v$  are the scales in millimetres per pixel, is necessary to change the unit of the point to metric units by  $p^I = T_s p^{I'}$ . In this paper, we will use the notation  $p^A$  to denote the coordinates of a point  $p$  in the coordinate system  $A$ .

If both the calibration and the image scales are known, each point can be mapped to 3D space by:

$$p^F = T_{F \leftarrow W} T_{W \leftarrow S} T_{S \leftarrow I} T_s p^{I'}. \quad (1)$$

In the above equation,  $T_{W \leftarrow S}$  can be read from the position sensor. The transformation from the world coordinate system to a phantom coordinate system  $T_{F \leftarrow W}$  is in fact not necessary in 3D image analysis. Most of the time, it is nevertheless included for convenience. Should it be removed, all analysis on the resulting 3D image will remain correct. However, the anatomy may appear at an absurd orientation. We will see later in this paper, how the choice of  $T_{F \leftarrow W}$  will help us to find the calibration parameters.

In this paper, we start by classifying each calibration technique accordingly to its principles. This is followed by a discussion of the metrics used to assess calibration quality. Finally, we compare the calibration techniques focusing on ease of use, speed of calibration and reliability. The comparison was performed based on our own experimental results, rather than figures quoted from previous papers to eliminate factors caused by differences in 3D ultrasound systems and user expertise.

## 2 Probe Calibration

Before we start introducing the different methods to calibrate a probe, we briefly outline a device that is often used in modern probe calibration (Anagnostoudis and Jan, 2005). This is a 3D localizer, often called a pointer or a stylus. Figure 2(a) shows one such stylus, consisting of a round shaft. On one end, it has position sensing devices that can be tracked by the position tracking system, at the other end, it is sharpened to a point. The localizer can report the location of its tip in 3D space, hence we can get the location of any point in space by pointing the stylus at the target.

Figure 2(b) shows the coordinate system involved when using a stylus. If the position of its tip  $r^L$  is known in the stylus's coordinate system  $L$ , then the position of the tip in 3D space is given by:

$$r^W = T_{W \leftarrow L} r^L, \quad (2)$$

where  $T_{W \leftarrow L}$  is provided by the position sensor. The position of the tip  $r^L$  may be supplied by the manufacturer (Muratore and Galloway Jr., 2001). When this position is not known, it can be determined by a pointer calibration (Leotta et al., 1997).

During a pointer calibration, the stylus is rotated about its tip while the position sensor's readings are recorded. Since the tip of the stylus remains stationary throughout the pointer calibration process, its location  $r^W$  in 3D space is therefore fixed. We can then use Equation 2 to

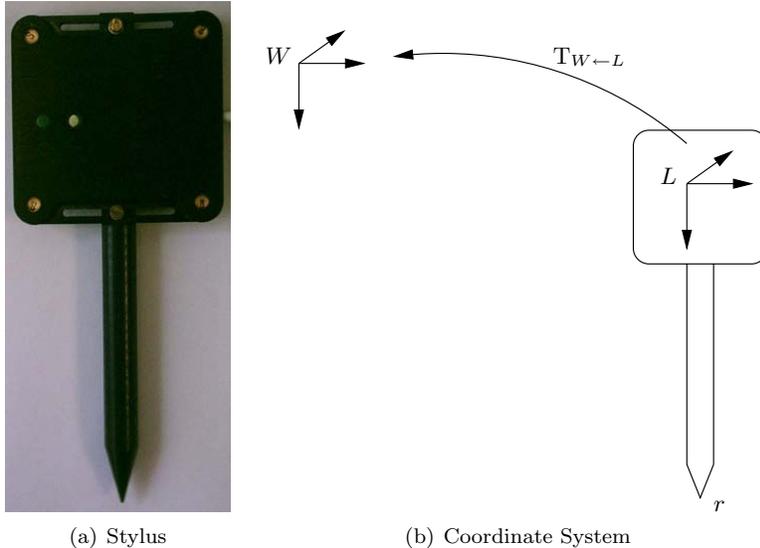


Figure 2: A 3D localizer.

solve for the position of the stylus tip, by minimizing

$$\sum_i \left| \overline{r_i^W} - T_{W \leftarrow L_i} r^L \right|, \quad (3)$$

where  $|\cdot|$  denotes the usual Euclidean norm on  $\mathbb{R}^3$  and  $\overline{a_i}$  the mean of  $(a_i)$ . We also used the notation  $r^L$  instead of  $r^{L_i}$  since  $r$  is invariant in every  $L_i$ .

The stylus is nevertheless prone to errors. These include errors from inaccurate tracking and pointer calibration. The accuracy of the pointer calibration is dependent on the size of stylus. Pointer calibrations typically have RMS errors between 0.6 and 0.9mm, but errors up to 1.5mm have been quoted (Hsu et al., 2007b). The tracking error is dependent on the tracking system. A typical optical tracking system, such as the Polaris, has a tracking error of 0.35mm. In general, a stylus has a positioning uncertainty of approximately 1mm.

The stylus has become popular in probe calibration because of its ability to locate points in space. Such a stylus is often part of the package when purchasing the position sensor for a freehand ultrasound system, so it is available for probe calibration.

## 2.1 Point Phantom

A common approach to perform probe calibration is to scan an object with known dimensions (a phantom). This phantom can be as simple as a point target. Indeed, this was one of the first phantoms (Detmer et al., 1994; State et al., 1994; Trobaugh et al., 1994) used for this purpose and continues to be used to this day (Barratt et al., 2006; Krupa, 2006). Calibrating with a point phantom can be divided into two classes, calibration with the aid of a stylus or without a stylus.

### 2.1.1 Point Phantom Without a Stylus

The point phantom can be formed by a pair of cross-wires (Detmer et al., 1994; Barry et al., 1997; Huang et al., 2005; Krupa, 2006) or a spherical bead-like object (State et al., 1994; Leotta et al., 1997; Legget et al., 1998; Pagoulatos et al., 1999; Barratt et al., 2006). Trobaugh et al. (1994) and Meairs et al. (2000) imaged a phantom with multiple point targets one at a time, but their theory is no different to the case when only a single target is used. The point phantom  $p$  is scanned, and its location  $p^{I'} = (u, v, 0)^t$  segmented in the B-scan. Now, if we position the phantom coordinate

system so that its origin coincides with the point phantom as shown in Figure 3, then Equation 1 becomes

$$\mathbb{T}_{F \leftarrow W} \mathbb{T}_{W \leftarrow S} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (4)$$

This is an equation with 11 unknowns—two scale factors, six calibration parameters and three parameters from  $\mathbb{T}_{F \leftarrow W}$ . Only the three translations in  $\mathbb{T}_{F \leftarrow W}$  need to be determined, since we do not care about the orientation of  $F$ , hence we can set the three rotations in  $\mathbb{T}_{F \leftarrow W}$  to arbitrary values, such as zeroes. If we capture  $N$  images of the point phantom from many directions and orientations, we can find these 11 unknowns by minimizing

$$f_{\text{point1}} = \sum_{i=1}^N \left| \mathbb{T}_{F \leftarrow W} \mathbb{T}_{W \leftarrow S_i} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s p^{I_i} \right|, \quad (5)$$

with the three rotations in  $\mathbb{T}_{F \leftarrow W}$  set to zero. This function can be minimized using iterative optimisation algorithms, such as the Levenberg-Marquardt algorithm (More, 1977). After the calibration parameters and the scales are found, the transformation  $\mathbb{T}_{F \leftarrow W}$  may be discarded and replaced with an alternative  $\mathbb{T}_{F \leftarrow W}$  that is convenient for visualization.

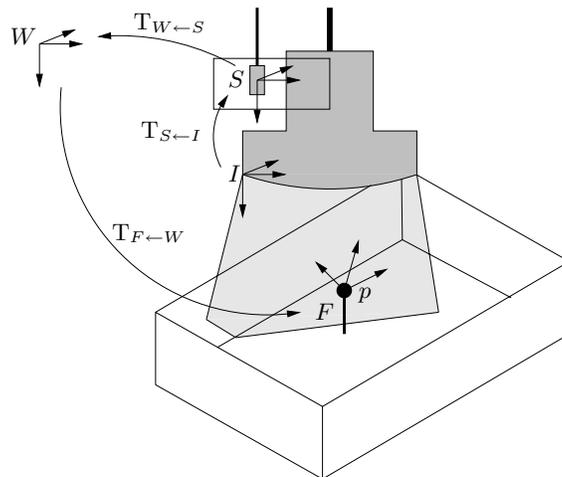


Figure 3: The geometry of a point phantom.

### 2.1.2 Point Phantom With a Stylus

When a stylus is available, we can find the position of the point phantom  $p^W$  in world space by pointing the stylus at the phantom. This approach was followed by P eria et al. (1995), Hartov et al. (1999), Amin et al. (2001) and Viswanathan et al. (2004). If the scales are unknown, the calibration can be solved by minimizing

$$f_{\text{point2}} = \sum_{i=1}^N \left| p^W - \mathbb{T}_{W \leftarrow S_i} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s p^{I_i} \right|. \quad (6)$$

There is little to be gained over Equation 5, since the minimum of this function needs to be found by an iterative minimization algorithm. Viswanathan et al. (2004) implemented an alternative solution form used in robotics (Andreff et al., 2001) involving Kronecker products (Brewer, 1978) to solve the calibration parameters and the image scales, but an iterative minimization algorithm is still required.

In some rare cases, the scales may be supplied by the manufacturer (Boctor et al., 2003) or by accessing the raw ultrasound signals (Hsu et al., 2006), but this requires special arrangements with the supplier. Otherwise, the scales can be obtained explicitly by using the distance measurement tool provided by the ultrasound machines (Hsu et al., 2006).

If the scales can be found (Péria et al., 1995), then the segmented image of the point phantom is known in millimetres:  $p^I = T_s p^{I'}$ . After the point has been located in world space by the stylus, it can be mapped to the sensor’s coordinate system by the inverse of the position sensor readings, i.e.  $p^S = T_{W \leftarrow S}^{-1} p^W$ . This means that the point phantom is known in the two coordinate system  $I$  and  $S$ , and we want to find a transformation  $T_{S \leftarrow I}$  that best transforms  $\{p^{I_i}\}$  to  $\{p^{S_i}\}$ . This can be found by minimizing

$$f_{\text{point3}} = \sum_{i=1}^N \left| T_{W \leftarrow S_i}^{-1} p^W - T_{S \leftarrow I} T_s p^{I_i} \right|. \quad (7)$$

Unlike the case with the previous function, the minimum of  $f_{\text{point3}}$  can be found in a closed-form, provided that the point has been scanned at three non-collinear locations in the B-scans. Péria et al. (1995) scanned three distinct points, but this is not necessary. The most popular solution to find the minimum of  $f_{\text{point3}}$  is the singular value decomposition technique devised by Arun et al. (1987), and modified by Umeyama (1991). Eggert et al. (1997) detailed the strengths and weaknesses of the different solution forms.

### 2.1.3 Point Phantom Variants

There are three major difficulties when using the point phantom described above. Most importantly, the images of the phantom need to be segmented manually. Although some automatic algorithms may exist (Hsu et al., 2007c), segmentation of isolated points in ultrasonic images are seldom reliable. This is evident from the fact that all of the above mentioned research groups who use a point target segmented their phantom manually. This makes the calibration process long and tiresome; it can take up to two hours depending on the number of points to be segmented. Secondly, it is very difficult to align the point phantom precisely with the scan plane. The finite thickness of the ultrasound beam makes the target visible in the B-scans even if the target is not precisely at the elevational centre of the scan plane. This error can be up to several millimetres depending on the beam thickness and the ability of the user to align the scan plane with the phantom. Finally, the phantom also needs to be scanned from a sufficiently diverse range of positions, and its location spread throughout the B-scan images. This is to ensure the resulting system of constraints is not under-determined with multiple solutions (Prager et al., 1998).

There are several phantoms that are designed to overcome the segmentation and alignment problems of the point phantom, while still based on the same mathematical principles. Liu et al. (1998) scanned a pyramid transversely. The pyramid appears as a triangle of varying sizes in the B-scans, depending on where the pyramid is scanned. The side lengths of the triangle are used to find the precise intersection of the scan plane with the pyramid. The three points of intersection act as three distinct point targets.

Brendel et al. (2004) scanned a sphere with a known diameter. The centre of the sphere acts as the virtual point phantom. The sphere appears as a circle in the B-scans and can be segmented automatically by using a Hough transform (Hough, 1959). Alignment is ensured providing the circle has the correct diameter. However, the lack of good visual feedback in the B-scans means that alignment is difficult. Sauer et al. (2001) scanned five spheres and manually fitted their image to a circle with the corresponding diameter. Gooding et al. (2005) placed a cross-wire through the centre of the sphere to ensure good alignment, while maintaining automatic segmentation. Hsu et al. (2007c) scanned a phantom consisting of two cones joining at a circle. The centre of this circle serves as the virtual point phantom as shown in Figure 4. Alignment of the scan plane with the circle is aided by the cones, so that even a slight misalignment can be detected. The fact that circles can be segmented automatically makes calibration simpler and quicker to perform.

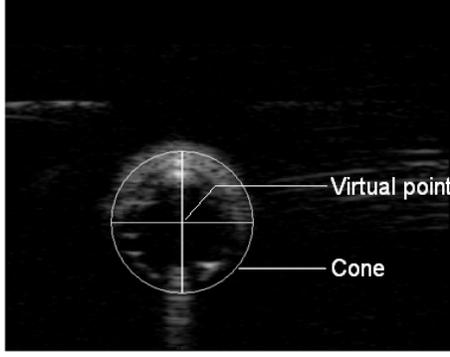


Figure 4: An image of the cone phantom.

## 2.2 Stylus

When a stylus is available, it is possible to perform calibration by just using the stylus. Instead of scanning a point phantom and finding its location with a stylus, the tip of the stylus itself can be scanned. Muratore and Galloway Jr. (2001) were the first to perform probe calibration with a stylus, and Zhang et al. (2006) also followed their approach. The calibration process is almost identical to the one where a point target is used. The tip of the stylus is scanned from many positions and orientations. This places constraints on the calibration parameters. If the image scales are unknown, a function similar to  $f_{\text{point2}}$  is minimized, the only difference being that the point target is now free to move around in space. The function to be minimized is

$$f_{\text{stylus}} = \sum_{i=1}^N \left| \mathbb{T}_{W \leftarrow L_i} r^L - \mathbb{T}_{W \leftarrow S_i} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s p_i^{I'} \right|, \quad (8)$$

where  $\mathbb{T}_{W \leftarrow L_i}$  is the stylus's location in space.

This technique is equivalent to a point phantom and is subject to most of its disadvantages. Hence alignment is a major source of error. Hsu et al. (2007c) designed a Cambridge stylus with a thick shaft. This shaft is thinned at a point precisely 20mm above the tip of the stylus. Instead of locating the stylus's tip in the scan plane, this thinned point is located. Any misalignment is detected visually.

Khamene and Sauer (2005) solve the alignment problem by attaching a rod to a position sensor, as shown in Figure 5. Both ends of the rod are pointer calibrated, and their locations in space are therefore  $\mathbb{T}_{W \leftarrow L} r_1^L$  and  $\mathbb{T}_{W \leftarrow L} r_2^L$ . The rod is scanned at an arbitrary location, and the point of intersection segmented in the B-scan. This point's location in world space is governed by Equation 1, and lies on the line joining the two ends of the rod. The distance from the point  $p^W$  to the line segment  $r_1^W r_2^W$  is

$$\frac{|(r_2^W - r_1^W) \times (r_1^W - p^W)|}{|r_2^W - r_1^W|}.$$

This distance must be zero, hence

$$\left| (\mathbb{T}_{W \leftarrow L} r_2^L - \mathbb{T}_{W \leftarrow L} r_1^L) \times (\mathbb{T}_{W \leftarrow L} r_1^L - \mathbb{T}_{W \leftarrow S} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s p^{I'}) \right| = 0.$$

The  $\times$  in the above equations denotes the cross product of two vectors in  $\mathbb{R}^3$ . Calibration can be found by minimizing

$$f_{\text{rod}} = \sum_{i=1}^N \left| (\mathbb{T}_{W \leftarrow L_i} r_2^L - \mathbb{T}_{W \leftarrow L_i} r_1^L) \times (\mathbb{T}_{W \leftarrow L_i} r_1^L - \mathbb{T}_{W \leftarrow S_i} \mathbb{T}_{S \leftarrow I} \mathbb{T}_s p_i^{I'}) \right|. \quad (9)$$

This is an equation with six unknowns—the six calibration parameters. Hsu et al. (2007c) pointed out that for a reliable optimisation, the scales needs to be found explicitly and fixed before optimisation.

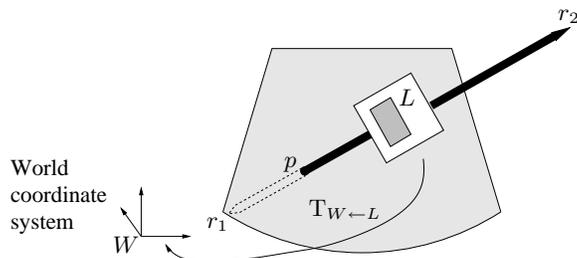


Figure 5: The geometry of a rod stylus.

### 2.3 Three-wire Phantom

The three-wire phantom is solely used by Carr (1996). Instead of mounting a pair of cross-wires in the fluid, three mutually orthogonal wires are used. These three wires form the three principal axes of the phantom coordinate system as shown in Figure 6. Each wire is scanned along its length individually. Suppose that the wire defining the  $x$ -axis is being scanned, then the point on the wire that is being scanned must satisfy

$$\mathbf{T}_{F \leftarrow W} \mathbf{T}_{W \leftarrow S} \mathbf{T}_{S \leftarrow I} \mathbf{T}_s \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} p_x^F \\ p_y^F = 0 \\ p_z^F = 0 \end{pmatrix}. \quad (10)$$

The  $y$  and  $z$  coordinates of  $p^F$  give rise to two equality constraints. If  $N_x$ ,  $N_y$  and  $N_z$  points were recorded along the  $x$ ,  $y$  and  $z$ -axes of the phantom coordinate system in that order, calibration can be solved by minimizing

$$f_{3\text{-wire}} = \sum_{i=1}^{N_x} \left( (p_{i_y}^F)^2 + (p_{i_z}^F)^2 \right) + \sum_{i=N_x+1}^{N_x+N_y} \left( (p_{i_x}^F)^2 + (p_{i_z}^F)^2 \right) + \sum_{i=N_x+N_y+1}^{N_x+N_y+N_z} \left( (p_{i_x}^F)^2 + (p_{i_y}^F)^2 \right). \quad (11)$$

This approach involves solving for 14 variables. These are the two scales, six calibration parameters and the six parameters that define the phantom coordinate system.

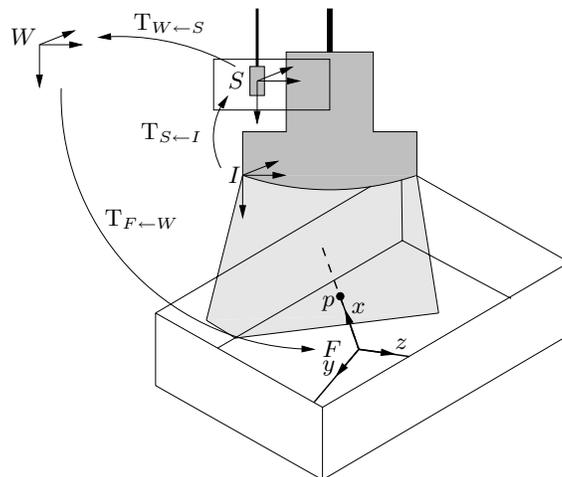


Figure 6: The geometry of a three-wire phantom.

This technique does not require the user to align the scan plane with the phantom and potentially speeds up the calibration process. Segmentation remains slow since manual intervention is

required. The user also needs to keep track of which wire is being scanned. The phantom may need to be precision manufactured to ensure that the wires are straight and orthogonal to each other.

## 2.4 Plane Phantom

Instead of scanning a point, it is possible to scan a plane. The design complexity of the plane varies from the floor of a container (Prager et al., 1998), a plexiglass plate (Rousseau et al., 2005), a nylon membrane (Langø, 2000) to a precision-made Cambridge phantom (Prager et al., 1998) and its variants (Varandas et al., 2004; Ali and Logeswaran, 2007).

The plane appears as a straight line in the B-scans. If we align the phantom coordinate system so that its  $xy$ -plane coincides with the plane phantom as shown in Figure 7, then every point on the line in the image must satisfy, by Equation 1:

$$\mathbf{T}_{F \leftarrow W} \mathbf{T}_{W \leftarrow S} \mathbf{T}_{S \leftarrow I} \mathbf{T}_s \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} p_x^F \\ p_y^F \\ p_z^F = 0 \end{pmatrix}. \quad (12)$$

The equation for the  $z$  coordinate of the phantom coordinate system is the required constraint on the calibration parameters. For each segmented line, we get two independent constraints by choosing any two points on the line. Choosing more points does not add any further information. The calibration parameters are solved by minimizing

$$f_{\text{plane}} = \sum_{i=1}^N \left( (p_{1i_z}^F)^2 + (p_{2i_z}^F)^2 \right), \quad (13)$$

where  $N$  is the number of images of the plane. The above equation is a function of 11 variables—two scales, six calibration parameters and three parameters from  $\mathbf{T}_{F \leftarrow W}$ . These three parameters consists of two rotations and one translation. Since we only require the  $xy$ -plane to coincide with the plane phantom, the two translations in the plane and the rotation about a normal of the plane will be absent in the equation.

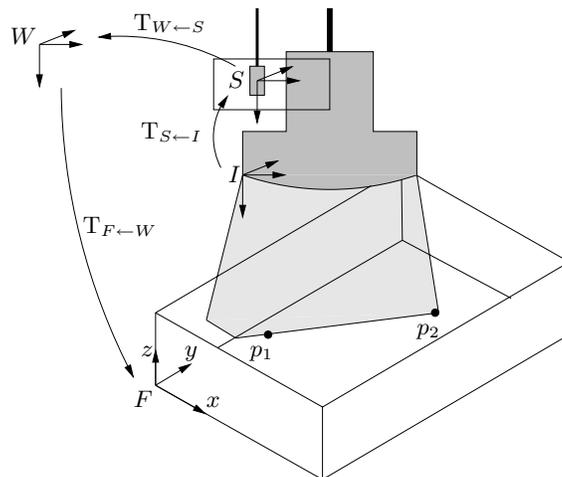


Figure 7: The geometry of a plane phantom.

The plane technique is attractive because it enables an automatic segmentation algorithm to be used, making the calibration process rapid to perform. The plane appears as a straight line in the B-scans. There are several automatic algorithms for segmenting a line, such as the Hough transform (Hough, 1959) and wavelet-based techniques (Kaspersen et al., 2001). Prager et al.

(1998) implemented a simplified version of the line detection algorithm by Clarke et al. (1996) and used the RANSAC algorithm to reject outliers (Fischler and Bolles, 1981).

A major drawback of using this approach, similar to the case of a point phantom, is that the phantom needs to be scanned from a wide range of angles and positions (Prager et al., 1998). In particular, the user is required to scan the phantom obliquely, as shown in Figure 8(a). Due to the thick ultrasound beam, point  $B$  is encountered by the ultrasound pulse before point  $A$ . The echo from point  $B$  makes the plane appear at an incorrect position. The user is subsequently required to scan the plane at the same angle on both sides of the normal to limit this error (Prager et al., 1998). Furthermore, much of the ultrasound energy is reflected away from the plane. The echo received by the probe is therefore weak, making segmentation at these positions difficult.

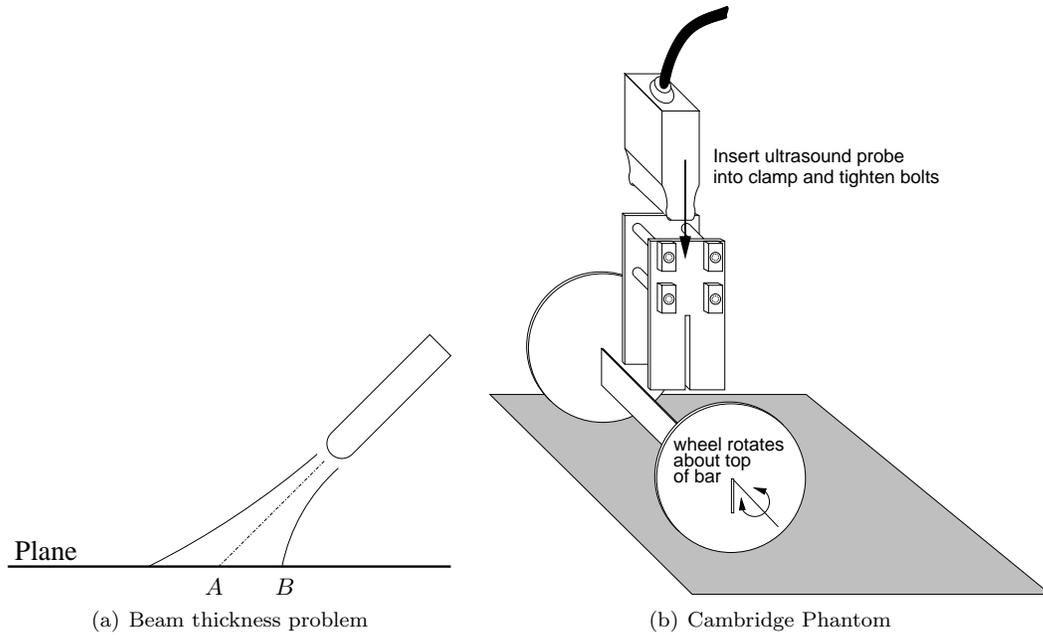


Figure 8: The beam thickness problem and solution.

It is possible to use a Cambridge phantom shown in Figure 8(b). The user is required to mount the probe onto the clamp, so that the scan plane is aligned with the slit of the clamp and hence with the brass bar. In order to assist the user to align the scan plane, a set of wedges (Figure 12(c)) can be placed on the brass bar. The user then aligns the scan plane with the wedges. In either case, aligning the scan plane with the brass bar may be difficult. The phantom is moved around in space by translating the phantom or rotating the wheels so that the phantom remains in contact with the floor of the container. Since the top of the brass bar joins the centre of the wheels, it always remains at a fixed height above the floor. The top of the brass bar serves as a virtual plane that is scanned. The advantage of using the Cambridge phantom is that a strong and clear reflection is received from the brass bar, irrespective of the probe position. The user can scan the plane from different angles and still get a clear image. However, the user is still required to scan the phantom from a wide range of positions and angles. Calibrating with a plane phantom is therefore a skilled task and requires the user to be experienced. From our experience of supporting medical physicists and clinicians, an incorrect calibration is often obtained because the phantom has not been scanned from a sufficiently diverse set of positions and orientations. Although an eigenvalue metric exists to detect whether the system of equations is under-constrained (Hsu et al., 2006), the user is still required to be sufficiently trained.

Dandekar et al. (2005) used two parallel wires to mimic a plane phantom. The virtual plane is formed by the unique plane that passes through the two parallel wires. The idea is to scan the set of two wires; the points of intersection of the wires with the scan plane are chosen as the points

$p_1$  and  $p_2$  in Figure 7. The phantom can be moved freely in the container so that both wires always intersect the scan plane. This phantom has the advantage that the beam thickness effect is minimized. When the plane is being scanned at an oblique angle, the plane no longer appears at an incorrect depth. The user therefore does not need to ensure that scans from the same angle to both sides of the normal were captured. However, the phantom needs to be precision manufactured to ensure that the wires are parallel. Most importantly, the wires need to be manually segmented. This sacrifices the rapid segmentation advantage of the plane phantom, making calibration, once again, a time consuming process. The user is still required to follow the same complex protocol and scan the phantom from a wide variety of positions and angles.

## 2.5 Two-plane Phantom

Boctor et al. (2003) designed a phantom with a set of parallel wires forming two orthogonal planes. When the set of wires is being scanned, it appears as distinct dots in the shape of a cross. If we align the phantom coordinate system with the orthogonal planes as shown in Figure 9, then a point  $p_h$  lying on the horizontal axis of the cross lies on the  $xz$ -plane of the phantom coordinate system and must satisfy:

$$\begin{pmatrix} p_x^F \\ p_y^F = 0 \\ p_z^F \end{pmatrix} = T_{F \leftarrow W} T_{W \leftarrow S} T_{S \leftarrow I} T_s p_h^{I'}. \quad (14)$$

The  $y$  coordinate in the above equation is a constraint on the calibration parameter. A similar constraint can be obtained for each point  $p_v$  on the vertical axis of the cross. Suppose that  $N$  images of the phantom are captured, each consisting of  $M_h$  points on the horizontal axis and  $M_v$  points on the vertical axis of the cross, then the calibration parameter and the scales can be found by minimizing

$$f_{2\text{-plane}} = \sum_{i=1}^N \left( \sum_{j=1}^{M_h} (p_{i,hj_y}^F)^2 + \sum_{j=1}^{M_v} (p_{i,vj_x}^F)^2 \right), \quad (15)$$

where  $p_{i,hj}$  and  $p_{i,vj}$  denote the  $j^{\text{th}}$  point on the horizontal and vertical axis of the cross in the  $i^{\text{th}}$  image. This equation consists of 13 variables, only the translation in the  $z$ -axis of the phantom coordinate system can be arbitrary.

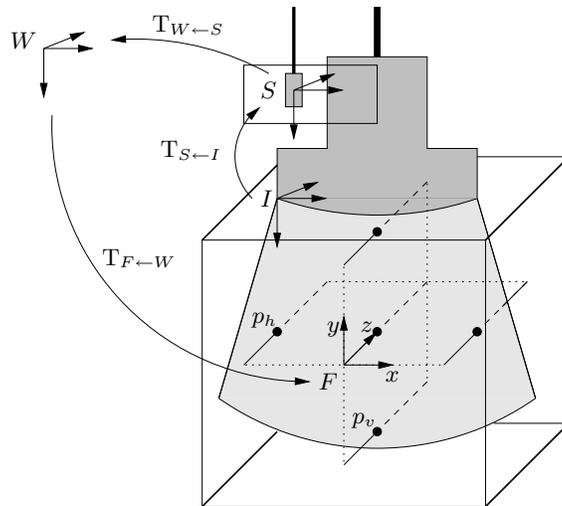


Figure 9: The geometry of a two-plane phantom.

An advantage of the two-plane phantom is that the set of wires appear as a cross in the ultrasound image. This information can be used to automatically segment the wires. Just as in the case with a point and a plane phantom, the phantom needs to be scanned from a wide variety of positions to constrain the calibration parameters.

It may be possible to generalize this idea and scan the faces of a cube with the phantom coordinate system suitably defined. Points on each face of the cube need to satisfy the equation for that plane and this places a constraint on the calibration parameters. Calibration can be solved by minimizing a similar equation to  $f_{\text{plane}}$  and  $f_{2\text{-plane}}$ . However, nobody has yet applied this calibration technique in a freehand 3D ultrasound system.

## 2.6 Two-dimensional Alignment Phantom

When calibration is performed using a point phantom with the aid of a stylus, with known scales, calibration only needs three non-collinear points to be positioned in the scan plane. If it is possible to align the scan plane with three such points at the same time, then even one frame is sufficient for calibration. Sato et al. (1998) was the first to use such a phantom. They scanned a thin board with three vertices as shown in Figure 10. The location of these vertices is determined by using a stylus. The scan plane is then aligned with these vertices, and each vertex is segmented in the B-scan. Since the distance between each pair of vertices is known, and we can find their distance in pixels from the ultrasound images, the scale factors can be easily computed. The calibration parameters can be solved in a closed-form by minimizing a function similar to  $f_{\text{point3}}$ . If we have captured  $N$  images of a two-dimensional alignment phantom with  $M$  fiducial points, calibration is found by minimizing

$$f_{2D} = \sum_{i=1}^N \sum_{j=1}^M \left| T_{W \leftarrow S}^{-1} p_j^W - T_{S \leftarrow I} T_s p_j^{I'} \right|. \quad (16)$$

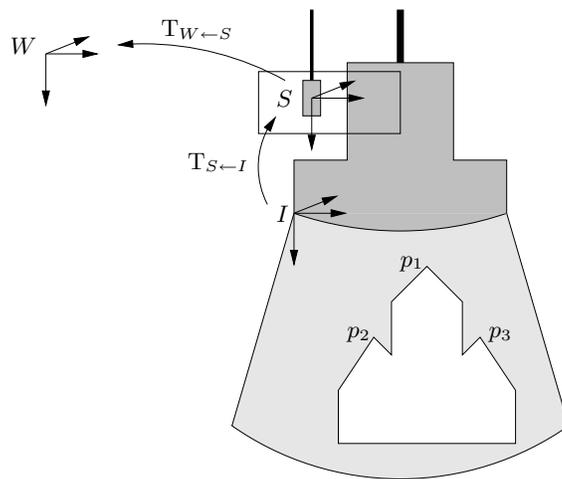


Figure 10: The coordinates of a 2D alignment phantom.

Several other groups use similar two-dimensional alignment phantoms with a variety of shapes and different numbers of fiducial points. Berg et al. (1999) aligned a jagged membrane with five corners and Welch et al. (2000) used an acrylic board with seven vertices. Beasley et al. (1999) scanned a ladder of wires with three weights fitted on the strings. Lindseth et al. (2003b) scanned a diagonal phantom, with the 9 fiducial points formed by cross-wires. Leotta (2004) fitted 21 spherical beads on parallel wires at different axial depths. The main disadvantage of this phantom is the requirement to align the phantom with the point fiducials, which can be very difficult. An advantage is that only one frame ( $N = 1$  in Equation 16) is theoretically needed for probe calibration, and a large number of fiducial points can be captured with just a few frames.

### 2.6.1 Z-phantom

The Z-fiducial phantom was designed so that the user is not required to align the scan plane with the 2D phantom (Comeau et al., 1998, 2000). The phantom consists of a set of wires in a ‘Z’ shape, as shown in Figure 11. The end points of the ‘Z’ wire configuration  $w_1, w_2, w_3$  and  $w_4$  can be found in space using a stylus. A typical Z-phantom may have up to 30 such ‘Z’ configurations. Instead of pointing the stylus at each end of the wire, there are usually a number of fixed locations (divots) on the phantom. The ‘Z’ shaped wire configurations are precision manufactured relative to these divots and the positions of the divots are located in space by using a stylus (Gobbi et al., 1999; Pagoulatos et al., 2001; Boctor et al., 2004; Zhang et al., 2004; Hsu et al., 2007a). It may be possible to attach a position sensor to the Z-phantom (Lindseth et al., 2003b; Chen et al., 2006), and precision manufacture the wires relative to the position sensor (Lindseth et al., 2003a). This requires the coordinates of the position sensor to be known. Calibration then requires two objects to be tracked simultaneously, but otherwise there is little difference between the two approaches. In each case, the end points of the wires can be found in space.

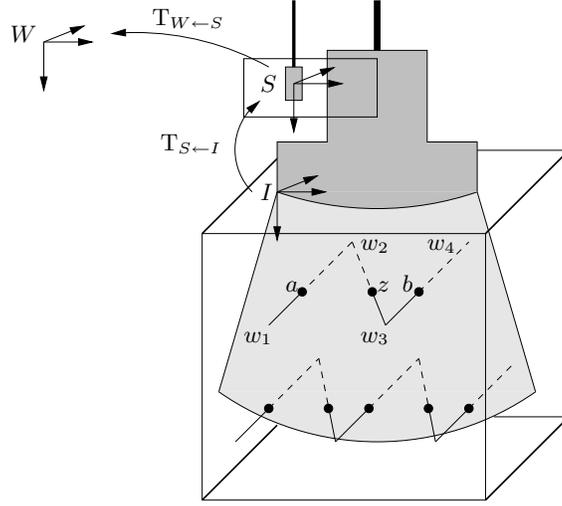


Figure 11: The geometry of a Z-phantom.

When the Z-phantom is scanned, the scan plane intersects the wire  $w_1w_2w_3w_4$  at  $a, z$  and  $b$ . These points are segmented in the ultrasound images. Assuming the image scales are known, the distances  $|z - b|$  and  $|a - b|$  can be measured off the B-scan images. The location of  $z$  is given by:

$$\begin{aligned} z^W &= w_3^W + \frac{|z - w_3|}{|w_2 - w_3|} (w_2^W - w_3^W) \\ &= w_3^W + \frac{|z - b|}{|a - b|} (w_2^W - w_3^W), \end{aligned} \quad (17)$$

since  $\triangle aw_2z$  and  $\triangle bw_3z$  are similar. If  $N$  images of the phantom are captured, each consisting of  $M$  Z-fiducials, then the calibration parameters can be found by minimizing

$$f_{\text{Z-phantom}} = \sum_{i=1}^N \sum_{j=1}^M \left| \mathbf{T}_{W \leftarrow S_i}^{-1} z_{ij}^W - \mathbf{T}_{S \leftarrow I} \mathbf{T}_s z_{ij}^{I_i} \right|, \quad (18)$$

where  $z_{ij}$  is the  $j^{\text{th}}$  Z-fiducial in the  $i^{\text{th}}$  frame. This function differs slightly from  $f_{2D}$  since the Z-fiducials are at different positions, depending on the scan plane, while 2D alignment phantoms are fixed in space.

The Z-phantom has the advantage that it does not require alignment of the scan plane with the phantom. It also maintains other advantages of a 2D alignment phantom, e.g. only one frame

is needed for calibration. However, the scale factors can no longer be measured off the B-scan images, and need to be found using other approaches as described in section 2.1.2.

It may be possible to segment the wires automatically. Chen et al. (2006) simplified their phantom to just two ‘Z’ wire configurations. Their segmentation algorithm involves finding two sets of parallel wires. Hsu et al. (2007a) mounted a membrane on top of their phantom, which is treated as a plane and can be segmented automatically. The wires are at known locations below the membrane, and this information is used to find the wires. This allows calibration to be completed in just a few seconds.

## 2.6.2 Mechanical Instrument

Gee et al. (2005) built a mechanical instrument that performs probe calibration by calibrating the position sensor and the scan plane separately to the gantry on the instrument. Since the two calibrations are independent, once the position sensor is calibrated, the depth and zoom settings can be changed and only the scan plane needs to be re-calibrated each time. This is achieved by using a specialized probe holder. Both the position sensor and the ultrasound probe are attached to the probe holder. The probe holder is positioned onto the phantom during calibration, and removed when the calibration is complete.

The phantom’s coordinate system is defined by its gantry  $G$ , where the probe holder  $H$  is mounted at a fixed location, as shown in Figure 12(a). The transformation  $T_{H \leftarrow G}$  is therefore constant and determined by the geometry. The position sensor is mounted onto the probe holder at a fixed location as well. The transformation  $T_{S \leftarrow H}$  is therefore also fixed. In order to find this transformation, the probe holder is placed into the sensor’s volume while the position sensor is attached. A stylus is then used to locate fixed landmarks on the probe holder and record the corresponding locations in the sensor’s coordinate system. Since this part of the calibration process is independent of the probe, replacing the probe or changing any of the ultrasound settings will not affect the relative position of the sensor to the gantry.

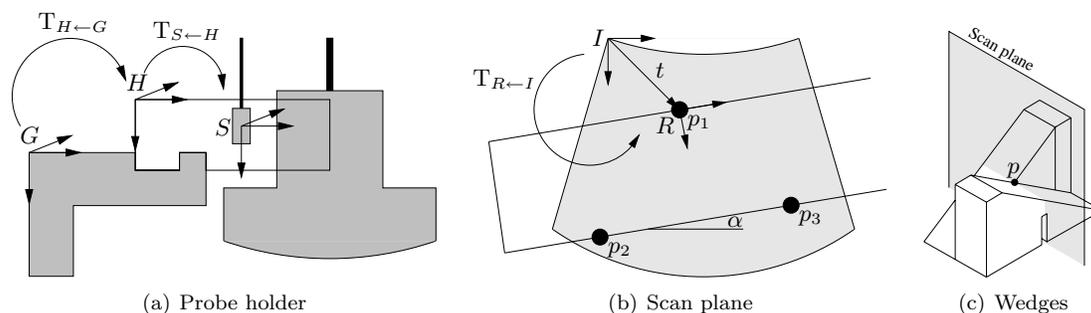


Figure 12: Geometry of the mechanical device for calibration.

In order to calibrate the scan plane, the user is required to align the scan plane with a 2D phantom by adjusting a set of micrometers. The 2D phantom consists of two parallel wires, with three sets of wedges  $p_1, p_2$  and  $p_3$  mounted on these wires at known locations. The coordinate system of the wires  $R$  is defined so that its origin coincides with  $p_1$ , as shown in Figure 12(b). Once the wires and these wedges are aligned with the scan plane, the image scales are found from the known distance between these wedges, as in the case with other 2D phantoms. If we rely on the user to ensure that  $p_2$  is to the left of  $p_3$ , the transformation  $T_{R \leftarrow I}$  only has three degrees of freedom—two translations  $t$  and a rotation  $\alpha$ . The translation is found from the location of  $p_1$  in the B-scan, and the rotation is found from the orientation of  $p_2$  and  $p_3$ .

The three sets of wedges also help in aligning the scan plane, rather than merely placing landmarks on the two wires. A set of wedges is shown in Figure 12(c). It consists of two triangular blocks. When the scan plane is aligned perfectly with the two wedges, a symmetrical reflection will be obtained in the ultrasound image. The surface of the wedges are roughened to ensure a

strong reflection. This visual feedback allows the 2D plane to be aligned with the scan plane to a high degree of accuracy.

Now, once the two calibrations are complete, the transformation that relates the wires to the gantry  $T_{G \leftarrow R}$  is simply read off the micrometers. Calibration is found as a series of transformations mapping from the B-scan to the wires, then to the gantry, the probe holder and finally to the position sensor, as shown in Figure 13. Calibration is therefore given by

$$T_{S \leftarrow I} = T_{S \leftarrow H} T_{H \leftarrow G} T_{G \leftarrow R} T_{R \leftarrow I}. \quad (19)$$

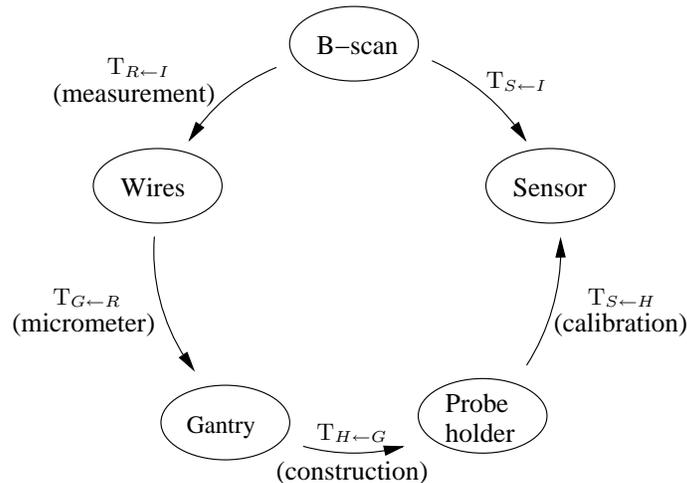


Figure 13: Principle of the mechanical device for calibration.

## 2.7 Image Registration

Another technique to calibrate a probe is image registration. When a point phantom is used for probe calibration, the point is scanned from different positions and orientations. The 3D image of the point can be constructed by using an assumed calibration and image scales. An iterative optimisation algorithm is implemented to find the calibration and scales so that the constructed image *best fits* the model. Here, *best fit* is measured by the amount of variation of the reconstructed point. Once the *best fit* has been found, the required calibration is the corresponding values that result in the least variation of the reconstructed point. This idea is used in other phantoms as well, each one using a different measure to define what is the *best fit*. For example, the plane phantom requires the reconstructed points to lie on a plane. Thus *best fit* is measured by the deviation of the reconstructed points from a particular plane. What these techniques have in common is that particular points of the phantom are selected, and *best fit* is measured as a function of the deviation of these points from their ideal location.

Blackall et al. (2000) built a gelatin phantom with tissue mimicking properties. The geometric model of the phantom is acquired by a magnetic resonance (MR) scan. The phantom is scanned with a freehand 3D ultrasound system. A 3D image of the phantom can be reconstructed by using an assumed calibration and Equation 1. An iterative optimisation algorithm is implemented to find the calibration and the image scales where the reconstructed image *best fits* the MR model. The similarity measure between two 3D volumes  $A$  and  $B$  is given by their mutual information (Studholme et al., 1999):

$$I(A, B) = \frac{H(A) + H(B)}{H(A, B)}, \quad (20)$$

where  $H(A)$  and  $H(B)$  denote the marginal entropies of the images and  $H(A, B)$  represents their joint entropy.

This technique is dependent on the image quality of the phantom and the similarity measure used. The impact of choosing another similarity measure (Pluim et al., 2003; Zitořa and Flusser, 2003) is unknown.

## 2.8 3D Probe Calibration

Although it is not the main focus of this paper to investigate calibrations for a 3D probe (a mechanically swept or a 2D array probe), we mention in passing that all the techniques that are used to calibrate a 2D probe are equally valid for the calibration of 3D probes. In fact, the exact same phantoms have been used, such as the point phantom (Sawada et al., 2004; Poon and Rohling, 2007) and the Z-phantom (Bouchet et al., 2001). The mathematical principles remain the same. However, since a 3D probe is used, a 3D image of the phantom is obtained. This is useful for segmenting the phantom. Lange and Eulenstein (2002) and Hastenteufel et al. (2003) used an image registration technique. Poon and Rohling (2005) provided a detailed discussion comparing calibrations using the various phantoms, including a three-plane phantom that has not been used to calibrate conventional 2D probes.

## 3 Calibration Quality Assessment

Probe calibration is a critical component of every freehand 3D ultrasound system, and its quality has a direct impact on the performance of the imaging system. It is therefore crucial to quantify the accuracy achievable with each calibration technique. Unfortunately, there has not been an agreed standard for assessing calibration quality. As a result, every research group may assess calibration quality differently, depending on what is available and convenient. Comparing calibration qualities between different research groups is therefore not straightforward. The quoted figures need to be interpreted on an individual basis, e.g. some may quote standard deviation and others may quote the 95% confidence interval. Nevertheless, we may classify all quality measures broadly into two classes, namely precision and accuracy.

### 3.1 Precision

One of the first measures used was formulated by Detmer et al. (1994) and used by various other research groups (Leotta et al., 1997; Prager et al., 1998; Blackall et al., 2000; Meairs et al., 2000; Muratore and Galloway Jr., 2001; Brendel et al., 2004; Dandekar et al., 2005). Now commonly named the reconstruction precision (RP), this measure is calculated by scanning a point phantom  $p$  from different positions and orientations. The point phantom is segmented in the B-scans and reconstructed in 3D space by using Equation 1. If  $N$  images of the point are captured, we get a cloud of  $N$  points spread in world space. Reconstruction precision is measured by the spread of this cloud of points, i.e.

$$\mu_{RP1} = \frac{1}{N} \sum_{i=1}^N \left| \mathbf{T}_{W \leftarrow S_i} \mathbf{T}_{S \leftarrow I} \mathbf{T}_s p_i^{I'} - \overline{p_i^W} \right|. \quad (21)$$

This equation can be generalized to include multiple calibrations:

$$\mu_{RP2} = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M \left| \mathbf{T}_{W \leftarrow S_i} \mathbf{T}_{S \leftarrow I_j} \mathbf{T}_s p_i^{I'} - \overline{p_{ij}^W} \right|, \quad (22)$$

where  $N$  is the number of images of the point phantom and  $M$  the number of calibrations.

Reconstruction precision measures the point reconstruction precision of the entire system, rather than calibration itself. This is dependent on a lot of factors, such as position sensor error, alignment error and segmentation error. Nevertheless, it is not unrelated to calibration. If calibration is far from correct, and the point phantom has been scanned from a sufficiently diverse

set of positions, then every image of the point will be mapped to an incorrect location, resulting in a huge spread and subsequently a large reconstruction error. A good reconstruction precision is therefore necessary, but unfortunately not sufficient, for a good calibration.

An alternative measure, based on the same idea as reconstruction precision, is called calibration reproducibility (CR) (Prager et al., 1998). Calibration reproducibility measures the variability in the reconstructed position of points in the B-scan. Suppose that only a single frame of the phantom is captured, and that  $N$  calibrations were performed, we can map the single point in space by using the different calibrations. Now, it is not necessary to reconstruct the point in world space, since the transformation  $T_{W \leftarrow S}$  is independent of calibration. Equivalently, the reconstruction can be done in the sensor’s coordinate system. This removes position sensing variations. Furthermore, the point phantom image itself is unnecessary. Imaging a point introduces alignment and segmentation errors. Instead, we may conveniently assume that a point has been imaged without scanning such a point physically, and that we have perfectly aligned and segmented its location  $p^{I'}$  on the B-scan. Calibration reproducibility is computed as follows:

$$\mu_{\text{CR}} = \frac{1}{N} \sum_{i=1}^N \left| T_{S \leftarrow I_i} T_s p^{I'} - \overline{p^{S_i}} \right|. \quad (23)$$

Clearly, the measure for calibration reproducibility is not just dependent on the calibrations ( $T_{S \leftarrow I_i}$ ), but also on the point  $p^{I'}$ . When Prager et al. (1998) first used this measure, they chose  $p^{I'}$  to be the centre of the image. Many research groups also gave the variation at the centre of the image (Meairs et al., 2000; Lindseth et al., 2003b). Pagoulatos et al. (2001) quoted variations for multiple points down the middle of the image. When there is an error in the calibration, say one of the rotational parameters, often the scan plane is incorrect by a rotation about some axis near the centre of the image. This means that points near the centre of the image are still roughly correct, but errors measured by points towards the edges are more visible. Therefore, many papers in recent years quote calibration reproducibility for a point at a corner of the image (Blackall et al., 2000; Rousseau et al., 2005), points along the left and right edges of the image (Leotta, 2004) and the four corners of the image (Treece et al., 2003; Gee et al., 2005; Hsu et al., 2006). Leotta (2004) and the Cambridge group (Treece et al., 2003; Gee et al., 2005; Hsu et al., 2006) also quoted the spread in the centre of the image. Brendel et al. (2004) gave the maximum variation of every point in the B-scan. Calibration reproducibility is a measure solely based on calibration, and does not incorporate errors like the position sensor or human skills such as alignment and segmentation. For this reason, calibration reproducibility has started to become the norm when precision is measured. In some papers, precision is simply defined as calibration reproducibility and referred to as “precision”.

Some research groups give the variation of the six calibration parameters (Amin et al., 2001; Boctor et al., 2003, 2004; Viswanathan et al., 2004). Other research groups give the variation of the three calibration translations and each entry in the rotational transformation matrix (Pagoulatos et al., 2001; Leotta, 2004), this is not appropriate since these values are not independent. In any case, interpreting these results is difficult since it is the variation due to the combination of these six parameters that is useful.

## 3.2 Accuracy

Precision measures the spread of a point in some coordinate system. This does not measure calibration accuracy as there may be a systematic error. In fact, it is almost impossible to measure calibration accuracy since the true calibration is unknown. If there was a technique that was able to give us the exact error, then this technique could be used to find the calibration parameters in the first place. Gee et al. (2005) measured their accuracy by considering the error in each component of their instrument. However, they have to assume that the scan plane can be aligned with their phantom without a systematic bias.

Many research groups quote accuracy for the entire freehand 3D ultrasound system. The calibration accuracy can then be deduced or inferred from the system accuracy, with a careful

quantization of every error source in the system evaluation (Lindseth et al., 2002). In fact, this is the ultimate accuracy that is important to a clinician, who is interested in the performance of the system, rather than some individual component. However, in such an environment, the accuracy of interest would be the *in vivo* accuracy. This is again difficult to assess. The reason is not only because it is difficult and inconvenient to scan a live patient in the laboratory, but the shape of the real anatomical structure is unknown. This is why the ultrasound system was built in the first place. Some research groups produce *in vivo* images in their papers (Meairs et al., 2000; Ali and Logeswaran, 2007), but merely as examples of images constructed by their system. As a result, accuracy experiments are often performed on artificial phantoms in a well controlled environment. Note that there are many papers on freehand 3D ultrasound systems as a whole. Although these papers may include probe calibration, their goal is to evaluate the accuracy of their system, rather than the calibration. We have thus excluded these accuracy assessments in this section. Their methods will favour clinical quantities, such as volume and *in vivo* images.

*In vitro* accuracy is nevertheless very different to *in vivo* accuracy. First, the image of the phantom usually has a better quality than in *in vivo* images. Unlike the speckle in *in vivo* images, phantom images have a clear border and segmentation is usually more accurate. For this reason, Treece et al. (2003) scanned a tissue mimicking phantom when assessing the accuracy of their system. Scanning *in vivo* is also subject to tissue deformation due to probe pressure (Treece et al., 2002). Furthermore, sound travels at different speeds as it passes through the various tissue layers, which does not occur in *in vitro* experiments. For a given system, the *in vitro* accuracy is generally better than the *in vivo* accuracy. Nevertheless, *in vitro* accuracy defines what can be achieved with such a system in an ideal environment.

### 3.2.1 Point Reconstruction Accuracy

Point reconstruction accuracy (PRA) is probably the most objective measure for accuracy. However, it is only recently, with the increased use of the stylus, that this technique has become widely used. A point  $p$  is scanned and its location reconstructed in 3D space. The 3D location of the point phantom is usually verified by the stylus (Blackall et al., 2000; Muratore and Galloway Jr., 2001; Pagoulatos et al., 2001). The only exception being Lindseth et al. (2003b), who precision manufactured their point phantom relative to the position sensor. Point reconstruction accuracy is given by the discrepancy between the reconstructed image and the stylus reading, i.e.

$$\mu_{\text{PRA}} = p^W - T_{W \leftarrow S} T_{S \leftarrow I} T_s p^I. \quad (24)$$

This is a measurement of system accuracy, and includes errors from every component of the system. The main error is due to manual misalignment of the scan plane with the point phantom used for accuracy assessment. As described before, when calibrating with a point phantom, manual alignment is difficult due to the thick beam width. There are other sources of error such as segmentation error and position sensor error, these should not be neglected. Of course, for better measurement, the point should be scanned from different positions and at different locations in the B-scan. A large number of images should be captured and the results averaged.

It is important that the image of the point phantom is scanned at different locations in the B-scans. This is because if the calibration was performed by capturing a series of images incorrectly in one region of the B-scan, then calibration would be most accurate for points near the same region of the B-scan. If the image of the point phantom used for accuracy assessment is again captured at the same region, the measured accuracy will appear to be higher than the true accuracy. In order to find the true calibration accuracy, the point phantom needs to be imaged at different locations throughout the B-scan.

Note that it is bad practice to use the same phantom that was used for calibration to assess its accuracy, especially when the location of the point fiducial is dependent on phantom construction (Liu et al., 1998; Chen et al., 2006). This means that point reconstruction accuracy is not very appropriate to assess a calibration performed using a point target, if the same phantom and the same algorithm are used. This is because if there is a flaw in the construction of the phantom,

such errors will cause an offset in the calibration. The same error will occur during accuracy assessment, and will remain unnoticed.

### 3.2.2 Distance Accuracy

Before the stylus was developed enabling the evaluation of point reconstruction accuracy, many groups assessed accuracy by measuring the distances between objects (Leotta et al., 1997; Prager et al., 1998; Blackall et al., 2000; Boctor et al., 2003; Lindseth et al., 2003b; Leotta, 2004; Dandekar et al., 2005; Hsu et al., 2006; Krupa, 2006). This technique is popular because the experiment is easy to set up. A phantom is manufactured with distinct landmarks. Even though the exact location of these landmarks are unknown in 3D space, the distances between the landmarks are known. This means that when the phantom is scanned and its image reconstructed in 3D space, we can compute the distances between the landmarks and see whether the computed distance is correct. The error measure is

$$\mu_{\text{Distance}} = |p_1^W - p_2^W| - \left| T_{W \leftarrow S_1} T_{S \leftarrow I} T_s p_1^{I'} - T_{W \leftarrow S_2} T_{S \leftarrow I} T_s p_2^{I'} \right|. \quad (25)$$

The idea behind this measure is that should a line be scanned, with an incorrect calibration, the image of the line in 3D space should be distorted. However, this depends on the way in which the phantom is scanned. Very often, when assessing the accuracy by distance measurement, a single sweep of the phantom is performed in one direction, as shown in Figure 14(a). Accuracy assessment performed in this way is incorrect. If the calibration is wrong, then the whole line will be incorrect in the same way. Each point will be offset by the same value and the reconstructed image will appear to be correct. What the user ends up assessing is the resolution of the ultrasound system. It is therefore not surprising that many research groups quote a high distance measurement accuracy.

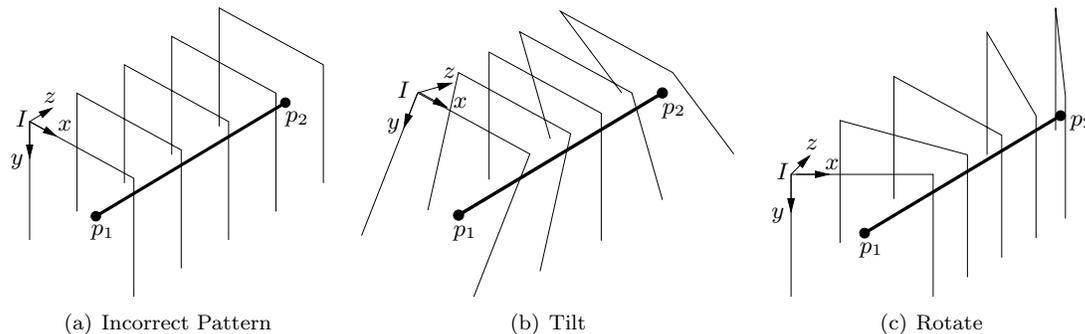


Figure 14: The different types of scanning pattern during accuracy assessment.

In order to successfully detect an incorrect calibration, the line should be scanned by tilting or rotating the probe in different directions, as shown in Figure 14(b) and (c). This ensures that calibration errors will map different points on the line in different directions. The line image will be a distorted curve for incorrect calibrations, and the distance between the two end points will be incorrect.

### 3.2.3 Volume Measurement

Some researchers produced a phantom with a known volume (Rousseau et al., 2005; Dandekar et al., 2005). The phantom is scanned and reconstructed in world space. The volume of the phantom can be calculated from their 3D imaging system. The computed volume is then compared with the known volume, and the difference quoted.

The advantage of this measure is that it gives the user an expected error for volume measurements. Often in diagnosis, the clinician is concerned about the volume of some anatomy, rather

than its absolute location. Hence errors in volume measurements are important. As in the case for distance accuracy, the position of the phantom may be incorrect. Also, the volume of such an object may be correct even if the calibration is incorrect, unless the phantom has been scanned with the probe rotated in some direction.

### 3.3 Comparison

It is very difficult to compare results quoted from different research groups, because of the differences in each measure. Treece et al. (2003) analyzed these differences and made an attempt to compare the results from different research group. However, even for calibration reproducibility, which does not contain user induced errors other than errors from calibration itself, it is difficult to compare results across different groups. Different calibrations are probably performed at different depth settings. Furthermore, the size of the cropped B-scan is probably different since a different ultrasound machine is used. Point reconstruction accuracy is highly dependent on the point phantom that is imaged. This has a direct impact on the ability to align the scan plane with the phantom accurately. Distance and volume measurements are highly dependent on the scan motion, which in turn is solely dependent on the user. Even so, many papers fail to describe the probe motion when performing such an assessment. This means that distance or volume measurement results are unlikely to be meaningful. Due to these differences, it has become common practice to give multiple accuracy measures (Lindseth et al., 2003b).

## 4 Choosing a Phantom

In this section, we will try to answer the question “what is the best way to calibrate a freehand 3D ultrasound system for a particular application?” At first glance, the question may seem trivial to answer: one should simply choose the most accurate technique. However, this accuracy is dependent on many factors, such as the probe type, user skill and calibration time.

### 4.1 Accuracy

Table 1 shows the precision (CR) and accuracy (PRA) achievable with the different phantoms. These results are from our research group, using the same precision and accuracy measures, on the same accuracy assessment phantom and with similar ultrasound settings when performing the calibrations. The figures in this table are therefore directly comparable. Where a citation is missing in the first column, we have performed calibrations with the same ultrasound machine and settings as the ones used by Hsu et al. (2007c), so that the results are comparable.

For precision, we have given the variation at the centre of the B-scan as well as the mean of the variations at the four corners and the centre of the B-scan. The PRA is computed by scanning the tip of a 1.5mm thick wire (Hsu et al., 2007c). We scanned the wire tip at five different regions in the B-scans—near the four corners and the centre of the B-scan. The probe is rotated through a full revolution about the lateral axis at six different positions. Five images of the wire are taken at each probe position and in each region of the B-scan. The PRA for images captured near the centre of the B-scan as well as the mean of every point captured are given in the table.

From the table it can be seen that the calibration performed using the Cambridge phantom is the most accurate, closely followed by the cone phantom. The Cambridge stylus and the plane phantom produce suboptimal accuracies, and the spherical stylus produces the worst accuracy. The best precision is obtained by calibrating with the mechanical instrument designed by Gee et al. (2005). The Cambridge phantom is least precise when calibrating at 3cm.

Table 2 shows the quoted precision reported by the various research groups. We have not included accuracy measurements for reasons outlined in the previous section. Each group uses different phantoms to assess their point reconstruction accuracy, and such accuracies are therefore dependent on the phantom used. From the table it can be seen that all the values for precision

Table 1: Precision and accuracy of calibrations performed by our research group. The precision and accuracy measurements are in millimetres.

Phantom	Probe	Depth	Precision (CR)		Accuracy (PRA)	
			Centre	Mean	Centre	Mean
Point (Cone) (Hsu et al., 2007c)	Linear	3cm	0.27	0.59	1.86	1.77
Stylus (Spherical) (Hsu et al., 2007c)	Linear	3cm	0.31	0.44	3.07	3.63
Stylus (Cambridge) (Hsu et al., 2007c)	Linear	3cm	0.45	0.61	1.52	2.18
Plane	Linear	3cm	0.39	0.57	2.46	2.28
Cambridge Phantom	Linear	3cm	0.83	0.88	1.56	1.67
Mechanical Instrument (Gee et al., 2005)	Linear	6cm	0.15	0.19	—	—
Mechanical Instrument (Gee et al., 2005)	Curvilinear	12cm	0.24	0.44	—	—
Z-phantom	Curvilinear	8cm	0.47	0.78	—	—
(Hsu et al., 2007a)	Curvilinear	15cm	1.07	1.54	—	—

are of the same order. Precision increases as the depth setting becomes shallower. This is exactly what we would expect.

Table 2: Precision of calibrations performed by the various research groups. The precision measurements are in millimetres.

Phantom	Probe	Depth	Precision (CR)		
			Centre	Corner	Mean
Point (Mears et al., 2000)	Linear	3.5cm	1.81	—	—
Point (Lindseth et al., 2003b)	Linear	8cm	0.62	—	—
Point (Sphere) (Hsu et al., 2007b)	Linear	3cm	0.31	0.47	0.44
Point (Cone) (Hsu et al., 2007b)	Linear	3cm	0.27	0.67	0.59
Plane (Rousseau et al., 2003)	Linear	—	—	0.89	—
Plane (Rousseau et al., 2005)	Sector	—	—	2.75	—
Plane	Linear	3cm	0.39	0.61	0.57
Cambridge phantom	Linear	3cm	0.83	0.89	0.88
2D alignment (Lindseth et al., 2003b)	Linear	8cm	0.44	—	—
2D alignment (Leotta, 2004)	Sector	10cm	0.67	1.32	1.19
Z-phantom (Lindseth et al., 2003b)	Linear	8cm	0.63	—	—
Z-phantom (Hsu et al., 2007a)	Curvilinear	8cm	0.47	0.86	0.78
Z-phantom (Hsu et al., 2007a)	Curvilinear	15cm	1.07	1.66	1.54
Image registration (Blackall et al., 2000)	Linear	4cm	—	1.05	—

## 4.2 Calibration Factors

There are several factors that should be taken into account when choosing a particular phantom. The most important factors, other than the precision and accuracy requirements, are the type of probe, difficulty of the calibration procedure and the calibration time.

### 4.2.1 Probe Type

There is a large difference between calibrating a linear and a curvilinear probe. Curvilinear probes usually have a lower frequency and are used for imaging at a higher depth setting. It is generally less accurate to calibrate a probe at a higher depth setting, since the image degrades away from the focus and constraining a larger image is more difficult. Despite all these effects, it is still very different to calibrate a linear and a curvilinear probe, even at the same depth. Some phantoms, such as the point phantom, may be equally suitable to calibrate both a linear and a curvilinear probe. On the other hand, 2D alignment phantoms are more suitable for a curvilinear probe, and the Cambridge phantom is more suitable to calibrate a linear probe. A 2D alignment phantom, particularly the Z-phantom, requires getting as many fiducials as possible into the same B-scan frame. Although it is theoretically possible to scan just a part of the phantom repeatedly with a linear probe, this defeats the purpose of using such a phantom. On the other hand, using a plane phantom to calibrate a curvilinear probe may be difficult. If calibration is not performed in a solution where sound travel at a speed similar to soft tissue, the distortions will cause the plane to appear as a curve, and not a line. The image needs to be rectified for accurate segmentation. A simple solution is to calibrate in hot water to match the sound speed in water and in soft tissue. If the Cambridge phantom is used to calibrate a probe at a high depth setting, the reverberation due to the clamp may degrade the image so badly that the brass bar is undetectable.

### 4.2.2 Ease of Use

The point phantom is difficult to use in the sense that the user needs to align the scan plane with the phantom. This requires a certain amount of skill and experience. If a stylus is available, the phantom only needs to be scanned at three non-collinear locations in the B-scans. So, a novice user will perform a slightly worse calibration, but probably not much worse than an expert. The 2D alignment phantoms require precise alignment of the whole phantom with the scan plane. This may be difficult to achieve for a novice user. The user will probably take a very long time to complete the task, but as in the case of the point phantom with a stylus, the calibration should be fairly reliable. In contrast, the Z-phantom does not need any alignment. Not much skill or experience is required to calibrate a probe. The accuracy achieved by an expert and a beginner should be similar. On the other hand, a point and a plane phantom are difficult to use. It is crucial to scan the phantom from a wide variety of positions. An inexperienced user usually neglects one or more of the required probe motions leading to an incorrect calibration. The Cambridge phantom requires the user to mount the probe accurately, which is also a skilled task.

### 4.2.3 Calibration Time

The time needed for calibration is dependent on the image quality and segmentation. Images of a point phantom often need to be segmented manually. Nevertheless, automatic segmentation algorithms have been implemented. The automatic segmentation of the plane phantom makes it attractive to use, as the time is shortened considerably. The Z-phantom can calibrate a probe in seconds, outperforming every other phantom in this sense.

## 4.3 Phantom Comparison

We want to answer the question “which phantom is most suitable for probe calibration?” We have listed and discussed the factors that should be taken into account when choosing such a phantom. Based on these factors, we will now compare the strengths and weaknesses of the four major phantoms, namely: point phantom, stylus, plane phantom, Z-phantom and their variants.

The two variants of the point phantom are the cone phantom and the spherical stylus used by Hsu et al. (2007c). Although the cone phantom is physically not a point phantom, it is based on the mathematical principle of a point phantom and has been classified as such in this review. From our perspective, this phantom shows what can be achieved when the point can be aligned and segmented accurately. We will use the results from the spherical stylus to represent a typical

point phantom. This phantom is subject to typical alignment and segmentation problems with point phantoms. The only advantage being that the point phantom can be moved around, which is unlikely to be a huge advantage. The stylus to be compared will be the Cambridge stylus. This shows what a stylus can achieve with a good alignment. The two variants of the plane phantom are the Cambridge phantom and a plexiglass plate.

In this paper, we will disregard some phantoms used by individual groups in the comparison. These phantoms include the three-wire phantom, two-plane phantom, ordinary 2D alignment phantom and the mechanical instrument. The main problem with the three-wire phantom is that a large number of frames is necessary for an accurate calibration. Manual segmentation is also required. Due to these drawbacks, this phantom has not been used in the last decade. The two-plane phantom works on the same principle as a plane phantom, and can be classified and compared as such. For a 2D alignment phantom, it is difficult to align the scan plane with the whole 2D phantom. It is probably easier to scan individual points one-by-one on such a phantom. For this reason, the 2D alignment phantom is inferior to the point phantom. The mechanical instrument is expensive to manufacture, making it uneconomical to be purchased for a freehand 3D ultrasound system. Also, the position sensor needs to be mounted at a fixed position relative to the phantom. This means that either a specific probe holder needs to be used, or the probe holder needs to be calibrated as well. Neither of these approaches offers a straightforward solution.

Table 4.3 ranks the six phantoms according to the different factors that are deemed important for calibration. For each factor, the phantoms are ranked from 1 to 6, where 6 is given to the least suitable phantom. The table is drawn up based on our experience with the phantoms.

Table 3: Probe calibration factors for the different phantoms.

Factor	Point		Stylus	Plane		2D alignment
	Sphere	Cone	Cambridge	Plate	Cambridge	Z-phantom
Precision	1	2	2	2	6 <sup>†</sup>	5
Accuracy	5	1	3	3	1	5
Easy to use (Novice)	4	2	2	4	4	1
Easy to use (Expert)	5	2	2	5	2	1
Segmentation	6	5	2	2	1	2
Speed	6	3	2	4	4	1
Reliability	4	2	2	4	4	1
Phantom simplicity	1	4	3	1	4	4
Linear probe	✓	✓	✓	✓	✓	✗ <sup>‡</sup>
Curvilinear probe	✓	✓	✓	✓	✗ <sup>‡</sup>	✓

<sup>†</sup> This is based on the precision at 3cm. The precision is better when calibrating at a higher depth.

<sup>‡</sup> It is possible, but difficult, to use these phantoms to calibrate the corresponding probe.

From the table we see that the Z-phantom is the easiest to use. Calibration can be completed within seconds. The calibration performed by a novice should be reliable and have a similar accuracy to that obtained by an expert. However, the precision and accuracy achievable by the phantom is among the worst of all the available phantoms. In contrast, the plane phantoms are difficult to use. The user needs to be sufficiently trained in order to use a plane phantom. The accuracy of a Cambridge phantom is nevertheless the best among the available phantoms. The Cambridge phantom also becomes very easy to use if the user is sufficiently skilled. The plexiglass plate also achieves moderate accuracy and is simple to make. The cone phantom is also very accurate. Not much training is required to use this phantom. However, the phantom needs to be aligned manually by the user, and the segmentation requires human intervention to mark the search region. The Cambridge stylus and the point target lie in the middle. They are not particularly simple nor very difficult to use, and can produce calibrations in a reasonable time. Automatic segmentation is also possible with good image quality. The Cambridge stylus produces

better accuracy with a slightly more complicated design. Most phantoms are suitable to calibrate both a linear and a curvilinear probe. The Z-phantom may be more suitable for a curvilinear probe, so that a large number of fiducials can be captured in the same frame, enabling very rapid calibration. The Cambridge phantom is not suitable for a curvilinear probe at high depth since the reverberation effect from the clamp corrupts the image so badly that the plane cannot be detected.

## 5 Conclusion

In this paper, we have classified all the phantoms used for freehand 3D ultrasound calibration by their mathematical principles. The strengths and weaknesses of each phantom are discussed. The different measures used to assess the calibration quality are analyzed and the accuracy of each phantom quantified. In the end, we have pointed out the situations where a particular phantom may be more suitable than others. Unfortunately, there is no single phantom that outperforms the rest. The Cambridge phantom and the cone phantom are the most accurate. The Cambridge phantom is most difficult to use for a novice user, but easy to use for an expert. The Z-phantom is easiest to use and produces a calibration within seconds, but its accuracy remains poor. The other phantoms lie between these extremes, offering moderate accuracy, ease of use and phantom complexity.

## References

- J. L. Alcazar. Three-dimensional ultrasound in gynecology: current status and future perspectives. *Current Women's Health Reviews*, 1(1), 2005.
- A. Ali and R. Logeswaran. A visual probe localization and calibration system for cost-effective computer-aided 3D ultrasound. *Computers in Biology and Medicine*, 37:1141–1147, 2007.
- D. V. Amin, T. Kanade, B. Jaramaz, A. M. DiGioia III, C. Nikou, R. S. LaBarca, and J. E. Moody Jr. Calibration method for determining the physical location of the ultrasound image plane. In *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 2208, pages 940–947. Springer, 2001.
- A. Anagnostoudis and J. Jan. Use of an electromagnetic calibrated pointer in 3D freehand ultrasound calibration. In *Radioelektronika*, 2005.
- N. Andreff, R. Horaud, and B. Espiau. Robot hand-eye calibration using structure from motion. *International Journal of Robotics Research*, 20(3):228–248, 2001.
- K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.
- D. C. Barratt, G. P. Penney, C. S. K. Chan, C. M. Slomczykowski, T. J. Carter, P. J. Edwards, and D. J. Hawkes. Self-calibrating 3D-ultrasound-based bone registration for minimally invasive orthopedic surgery. *IEEE Transactions on Medical Imaging*, 25(3):312–323, 2006.
- C. D. Barry, C. P. Allott, N. W. John, P. M. Mellor, P. A. Arundel, D. S. Thomson, and J. C. Waterton. Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in Medicine & Biology*, 23(8):1209–1224, 1997.
- R. A. Beasley, J. D. Stefansic, A. J. Herline, L. Guttierrez, and R. L. Galloway Jr. Proceedings of spie. volume 3658, pages 125–132, 1999.
- S. Berg, H. Torp, D. Martens, E. Steen, S. Samstad, I. Høivik, and B. Olstad. Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data. *Ultrasound in Medicine & Biology*, 25(5):745–753, 1999.
- J. M. Blackall, D. Rueckert, C. R. Maurer Jr., Hill D. L. G. Penney, G. P., and D. J. Hawkes. An image registration approach to automated calibration for freehand 3D ultrasound. In *Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 1935, pages 462–471. Springer, 2000.
- E. M. Boctor, A. Jain, M. A. Choti, R. H. Taylor, and G. Fichtinger. A rapid calibration method for registration and 3D tracking of ultrasound images using spatial localizer. In *Proceedings of SPIE*, volume 5035, pages 521–532, 2003.
- E. M. Boctor, A. Viswanathan, M. Choti, R. H. Taylor, G. Fichtinger, and G. Hager. A novel closed form solution for ultrasound calibration. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, volume 1, pages 527–530, 2004.
- L. G. Bouchet, S. L. Meeks, G. Goodchild, F. J. Bova, J. M. Buatti, and W. A. Friedman. Calibration of three-dimensional ultrasound images for image-guided radiation therapy. *Physics in Medicine and Biology*, 46:559–577, 2001.
- B. Brendel, S. Winter, and H. Ermert. A simple and accurate calibration method for 3D freehand ultrasound. In *Biomedizinische Technik BMT 2004*, 2004.
- J. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, 25(9):772–781, 1978.

- J. C. Carr. *Surface reconstruction in 3D medical imaging*. PhD thesis, University of Canterbury, Christchurch, New Zealand, 1996.
- T. K. Chen, P. Abolmaesumi, A. D. Thurston, and R. E. Ellis. Automated 3D freehand ultrasound calibration with real-time accuracy control. In *Proceedings of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 4190, pages 899–906. Springer, 2006.
- J. C. Clarke, S. Carlsson, and A. Zisserman. Detecting and tracking linear features efficiently. In *Proceedings of the British Machine Vision Conference 1996*, pages 415–424, Edinburgh, 1996. British Machine Vision Association.
- R. M. Comeau, A. Fenster, and T. M. Peters. Integrated MR and ultrasound imaging for improved image guidance in neurosurgery. In *Proceedings of SPIE*, volume 3338, pages 747–754, 1998.
- R. M. Comeau, A. F. Sadikot, A. Fenster, and T. M. Peters. Intraoperative ultrasound for guidance and tissue correction in image-guided neurosurgery. *Medical Physics*, 27(4):787–800, 2000.
- S. Dandekar, Y. Li, J. Molloy, and J. Hossack. A phantom with reduced complexity for spatial 3-D ultrasound calibration. *Ultrasound in Medicine & Biology*, 31(8):1083–1093, 2005.
- P. R. Detmer, G. Bashein, T. Hodges, K. W. Beach, E. P. Filer, D. H. Burns, and D. E. Stradness Jr. 3D ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation. *Ultrasound in Medicine & Biology*, 20(9):923–936, 1994.
- D. W. Eggert, A. Lorusso, and R. B. Fisher. Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, 9:272–290, 1997.
- S. Eulenstein, T. Lange, M. Hünerbein, P. M. Schlag, and H. Lamecker. Ultrasound-based navigation system incorporating preoperative planning for liver surgery. *International Congress Series*, 1268:758–763, 2004.
- A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46:R67–R99, 2001.
- A. Fenster, A. Landry, D. B. Downey, R. A. Hegele, and J. D. Spence. 3D ultrasound imaging of the carotid arteries. *Current Drug Targets—Cardiovascular & Hematological Disorders*, 4(2):161–175, 2004a.
- A. Fenster, K. J. M. Surry, G. R. Mills, and D. B. Downey. 3D ultrasound guided breast biopsy system. *Ultrasonics*, 42:769–774, 2004b.
- M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- A. H. Gee, R. W. Prager, G. H. Treece, and L. H. Berman. Engineering a freehand 3D ultrasound system. *Pattern Recognition Letters*, 24:757–777, 2003.
- A. H. Gee, N. E. Houghton, G. M. Treece, and R. W. Prager. A mechanical instrument for 3D ultrasound probe calibration. *Ultrasound in Medicine & Biology*, 31(4):505–518, 2005.
- D. G. Gobbi, R. M. Comeau, and T. M. Peters. Ultrasound probe tracking for real-time ultrasound/MRI overlay and visualization of brain shift. In *Proceedings of the 2nd International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 1679, pages 920–927. Springer, 1999.
- L. F. Gonçalves, W. Lee, J. Espinoza, and R. Romero. Three- and 4-dimensional ultrasound in obstetric practice. Does it help? *Journal of Ultrasound in Medicine*, 24:1599–1624, 2005.

- M. J. Gooding, S. H. Kennedy, and J. A. Noble. Temporal calibration of freehand three-dimensional ultrasound using image alignment. *Ultrasound in Medicine & Biology*, 31(7):919–927, 2005.
- A. Hartov, S. D. Eisner, D. W. Roberts, K. D. Paulsen, L. A. Platenik, and M. I. Miga. Error analysis for a free-hand three-dimensional ultrasound system for neuronavigation. *Neurosurgical Focus*, 6(3), 1999.
- M. Hastenteufel, S. Mottl-Link, I. Wolf, R. de Simone, and H-P. Meinzer. A method for the calibration of 3D ultrasound transducers. In *Proceedings of SPIE*, volume 5029, pages 231–238, 2003.
- P. V. C. Hough. Machine analysis bubble chamber pictures. In *International Conference on High Energy Accelerators and Instrumentation*, pages 554–556. CERN, 1959.
- P-W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece. Rapid, easy and reliable calibration for freehand 3D ultrasound. *Ultrasound in Medicine & Biology*, 32(6):823–835, 2006.
- P-W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece. Real-time freehand 3D ultrasound. *Ultrasound in Medicine & Biology*, in press, 2007a.
- P-W. Hsu, R. W. Prager, N. E. Houghton, A. H. Gee, and G. M. Treece. Accurate fiducial location for freehand 3D ultrasound calibration. In *Proceedings of SPIE*, volume 6513, 2007b.
- P-W. Hsu, G. M. Treece, R. W. Prager, N. E. Houghton, and A. H. Gee. Comparison of free-hand 3D ultrasound calibration techniques using a stylus. *Ultrasound in Medicine & Biology*, submitted, 2007c.
- Q. H. Huang, Y. P. Zheng, M. H. Lu, and Z. R. Chi. Development of a portable 3D ultrasound imaging system for musculoskeletal tissues. *Ultrasonics*, 43:153–163, 2005.
- J. H. Kaspersen, T. Langø, and F. Lindseth. Wavelet-based edge detection in ultrasound images. *Ultrasound in Medicine & Biology*, 27(1):89–99, 2001.
- A. Khamene and F. Sauer. A novel phantom-less spatial and temporal ultrasound calibration method. In *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 3750, pages 65–72. Springer-Verlag, 2005.
- A. Krupa. Automatic calibration of a robotized 3D ultrasound imaging system by visual servoing. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 4136–4141, 2006.
- T. Lange and S. Eulenstein. Calibration of swept-volume 3-D ultrasound. In *Proceedings of Medical Image Understanding and Analysis*, volume 3, pages 29–32, 2002.
- T. Langø. *Ultrasound guided surgery: image processing and navigation*. PhD thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2000.
- M. E. Legget, D. F. Leotta, E. L. Bolson, J. A. McDonald, R. W. Martin, X-N. Li, C. M. Otto, and F. H. Sheehan. System for quantitative three-dimensional echocardiography of the left ventricle based on a magnetic-field position and orientation sensing system. *IEEE Transactions on Biomedical Engineering*, 45(4):494–504, 1998.
- D. F. Leotta. An efficient calibration method for freehand 3-D ultrasound imaging systems. *Ultrasound in Medicine & Biology*, 30(7):999–1008, 2004.
- D. F. Leotta, P. R. Detmer, and R. W. Martin. Performance of a miniature magnetic position sensor for three-dimensional ultrasound imaging. *Ultrasound in Medicine & Biology*, 23(4): 597–609, 1997.

- F. Lindseth, T. Langø, and J. Bang. Accuracy evaluation of a 3D ultrasound-based neuronavigation system. *Computer Aided Surgery*, 7:197–222, 2002.
- F. Lindseth, J. Bang, and T. Langø. A robust and automatic method for evaluating accuracy in 3-D ultrasound-based navigation. *Ultrasound in Medicine & Biology*, 29(10):1439–1452, 2003a.
- F. Lindseth, G. A. Tangen, T. Langø, and J. Bang. Probe calibration for freehand 3-D ultrasound. *Ultrasound in Medicine & Biology*, 29(11):1607–1623, 2003b.
- J. Liu, X. Gao, Z. Zhang, S. Gao, and J. Zhou. A new calibration method in 3D ultrasonic imaging system. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 20, pages 839–841, 1998.
- S. Meairs, J. Beyer, and M. Hennerici. Reconstruction and visualization of irregularly sampled three- and four-dimensional ultrasound data for cerebrovascular applications. *Ultrasound in Medicine & Biology*, 26(2):263–272, 2000.
- S. L. Meeks, J. M. Buatti, L. G. Bouchet, F. J. Bova, T. C. Ryken, E. C. Pennington, K. M. Anderson, and W. A. Friedman. Ultrasound-guided extracranial radiosurgery: technique and application. *International Journal of Radiation Oncology, Biology, Physics*, 55(4):1092–1101, 2003.
- L. Mercier, T. Langø, F. Lindseth, and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine & Biology*, 31(4):449–471, 2005.
- J. J. More. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical Analysis, Lecture Notes in Mathematics*, volume 630, pages 105–116. Springer-Verlag, 1977.
- D. M. Muratore and R. L. Galloway Jr. Beam calibration without a phantom for creating a 3-D freehand ultrasound system. *Ultrasound in Medicine & Biology*, 27(11):1557–1566, 2001.
- T. R. Nelson and D. H. Pretorius. Three-dimensional ultrasound imaging. *Ultrasound in Medicine & Biology*, 24(9):1243–1270, 1998.
- N. Pagoulatos, W. S. Edwards, D. R. Haynor, and Y. Kim. Interactive 3-D registration of ultrasound and magnetic resonance images based on a magnetic position sensor. *IEEE Transactions on Information Technology in Biomedicine*, 3(4):278–288, 1999.
- N. Pagoulatos, D. R. Haynor, and Y. Kim. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. *Ultrasound in Medicine & Biology*, 27(9):1219–1229, 2001.
- G. P. Penney, J. M. Blackall, M. S. Hamady, T. Tabharwal, A. Adam, and D. J. Hawkes. Registration of freehand 3d ultrasound and magnetic resonance liver images. *Medical Image Analysis*, 8:81–91, 2004.
- O. Péria, L. Chevalier, A. François-Joubert, J-P. Caravel, S. Dalsoglio, S. Lavallée, and P. Cinquin. Using a 3D position sensor for registration of SPECT and US images of the kidney. In *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine, Lecture Notes in Computer Science*, volume 905, pages 23–29, 1995.
- J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.
- T. C. Poon and R. N. Rohling. Comparison of calibration methods for spatial tracking of a 3-D ultrasound probe. *Ultrasound in Medicine & Biology*, 31:1095–1108, 2005.
- T. C. Poon and R. N. Rohling. Tracking a 3-D ultrasound probe with constantly visible fiducials. *Ultrasound in Medicine & Biology*, 33(1):152–157, 2007.

- R. W. Prager, R. N. Rohling, A. H. Gee, and L. Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine & Biology*, 24(6):855–869, 1998.
- F. Rousseau, P. Hellier, and C. Barillot. Robust and automatic calibration method for 3D freehand ultrasound. In *Proceedings of the 6th International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 2879, pages 440–448. Springer, 2003.
- F. Rousseau, P. Hellier, and C. Barillot. Confhuisus: A robust and fully automatic calibration method for 3D freehand ultrasound. *Medical Image Analysis*, 9(1):25–38, 2005.
- O. M. Rygh, T. A. Nagelhus Hernes, F. Lindseth, T. Selbekk, T. Brostrup, and T. B. Müller. Intraoperative navigated 3-dimensional ultrasound angiography in tumor surgery. *Surgical Neurology*, 66(6):581–592, 2006.
- Y. Sato, M. Nakamoto, Y. Tamaki, T. Sasama, I. Sakita, Y. Nakajima, M. Monden, and S. Tamura. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging*, 17(5):681–693, 1998.
- F. Sauer, A. Khamene, B. Bascle, L. Schinunang, F. Wenzel, and S. Vogt. Augmented reality visualization of ultrasound images: system description, calibration, and features. In *Proceedings of the IEEE and ACM International Symposium on Augmented Reality*, pages 30–39, 2001.
- A. Sawada, K. Yoda, M. Kokubo, T. Kunieda, Y. Nagata, and M. Hiraoka. A technique for noninvasive respiratory gated radiation treatment system based on a real time 3D ultrasound image correlation: a phantom study. *Medical Physics*, 31(2), 2004.
- A. State, D. T. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: observing a volume rendered fetus within a pregnant patient. In *Proceedings of the Conference on Visualization '94, IEEE Visualization*, pages 364–368, California, 1994. IEEE Computer Society Press.
- C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.
- G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. *Medical Image Analysis*, 6:199–214, 2002.
- G. M. Treece, A. H. Gee, R. W. Prager, C. J. C. Cash, and L. H. Berman. High-definition freehand 3-D ultrasound. *Ultrasound in Medicine & Biology*, 29(4):529–546, 2003.
- J. W. Trobaugh, W. D. Richard, K. R. Smith, and R. D. Bucholz. Frameless stereotactic ultrasonography: method and applications. *Computerized Medical Imaging and Graphics*, 18(4): 235–246, 1994.
- S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.
- G. Unsgaard, O. M. Rygh, T. Selbekk, T. B. Müller, F. Kolstad, F. Lindseth, and T. A. Nagelhus Hernes. Intra-operative 3D ultrasound in neurosurgery. *Acta Neurochirurgica*, 148(3): 235–253, 2006.
- J. Varandas, P. Baptista, J. Santos, R. Martins, and J. Dias. VOLUS — a visualization system for 3D ultrasound data. *Ultrasonics*, 42:689–694, 2004.
- A. Viswanathan, E. M. Boctor, R. H. Taylor, G. Hager, and G. Fichtinger. Immediate ultrasound calibration with three poses and minimal image processing. In *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 3217, pages 446–454. Springer, 2004.

- J. N. Welch, J. A. Johnson, M. R. Bax, R. Badr, and R. Shahidi. A real-time freehand 3D ultrasound system for image-guided surgery. *IEEE Ultrasonics Symposium*, 2:1601–1604, 2000.
- S. Yagel, S. M. Cohen, I. Shapiro, and D. V. Valsky. 3D and 4D ultrasound in fetal cardiac scanning: a new look at the fetal heart. *Ultrasound in Obstetrics and Gynecology*, 29(1):81–95, 2007.
- H. Zhang, F. Banovac, A. White, and K. Cleary. Freehand 3D ultrasound calibration using an electromagnetically tracked needle. 6141, 2006.
- W. Y. Zhang, R. N. Rohling, and D. K. Pai. Surface extraction with a three-dimensional freehand ultrasound system. *Ultrasound in Medicine & Biology*, 30(11):1461–1473, 2004.
- B. Zitořa and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003.