Freehand Three-Dimensional Ultrasound Calibration

許伯偉
Po-Wei Hsu

Department of Engineering
Girton College
University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

December 2007
Acknowledgements

I would like to thank:

Dr. Richard Prager, my supervisor, for giving me the freedom to investigate the topics that interested me, guiding my research and encouraging me to keep on schedule.

Dr. Andrew Gee and Dr. Graham Treece, for a level of support throughout which was well beyond the call of duty.

James Housden and Joel Lindop, colleagues in the Medical Imaging Group, for valuable scientific discussions.

Mr. Patrick Gosling, for maintaining the computing facility and solving many technical problems of the ultrasound system.

my schoolmates from Girton College, junior scholars from the Engineering Department and schoolfellows from the university for their friendship.

my parents, for everything.
Abstract

Freehand three-dimensional (3D) ultrasound is a technique for acquiring ultrasonic data of a 3D volume with application in radiotherapy and surgery planning. As the clinician sweeps the ultrasound probe over the anatomy, the trajectory of the probe is recorded using an attached position sensor. However, the position sensor tracks the position of the sensor, rather than the ultrasound B-scan. It is therefore necessary to find the rigid body transformation from the coordinate system of the B-scan to that of the position sensor. This transformation is found by probe calibration.

During probe calibration, the user usually scans an object with known dimensions (a phantom). An image of the phantom is reconstructed in 3D space. The correct rigid body calibration transformation is found when the reconstructed image matches its model. There are several designs of the phantom, each with its own strengths and weaknesses. In a clinical environment, some of the factors considered when choosing a phantom are accuracy, reliability, speed and ease of use. If probe calibration is inaccurate, the reconstructed image of the anatomy will be distorted, leading to incorrect distance and volume measurements. A difficult calibration routine may result in an unreliable calibration if the user is not suitably trained. In this thesis, we investigate how to achieve an accurate, reliable and rapid calibration.

Our first step is to improve the reliability of an accurate calibration technique from the literature. The Cambridge phantom, a variant of plane-based calibration, has been shown to be one of the most accurate calibration techniques. Unfortunately, calibrations performed using plane-based techniques are often unreliable. The user is required to scan the phantom in a complex way. Inexperienced users often neglect one of the required scanning motions which results in an unreliable calibration. We show how it is possible to provide feedback
on the reliability of the calibration. This allows the user to rectify an unreliable calibration.

Having achieved an accurate and reliable calibration, we now search for a fast and easy calibration technique. We study a class of two-dimensional alignment phantoms—the Z-fiducial phantom. Probe calibration using such a phantom only requires a single image of the phantom. However, calibration speed using this phantom is impeded by the necessity of segmenting isolated points on the phantom reliably, which requires human intervention. We solve this problem by mounting a thin rubber membrane on top of the phantom. The membrane is segmented automatically and the phantom features can be easily located as they are at known positions relative to the membrane. This enables us to segment isolated points automatically at the full PAL frame rate of 25Hz, enabling calibration to be completed in a few seconds.

In addition to improving existing calibration techniques, we present two novel phantoms—the cone phantom and the Cambridge stylus. They are both simple in design, easy to use and produce accurate calibrations. The cone phantom produces calibrations with accuracies matching the Cambridge phantom. The Cambridge stylus is small in size and can be carried around conveniently. These phantoms offer alternatives to the Cambridge phantom and the Z-phantom, ensuring calibration reliability and simplicity, while producing accurate calibrations.

In this dissertation, we show how to achieve a good reliability with one of the most accurate calibration techniques—the plane-based calibrations. We also show how to achieve rapid calibration by using the Z-phantom, allowing calibration to be performed in a few seconds. We present the cone phantom and the Cambridge stylus. Both of them are easy to use and are capable of achieving a good accuracy.
Declaration

This dissertation is the result of my own original work and does not include anything done in collaboration with others, except where acknowledged. It has not been submitted in whole or in part for a degree at any other university. It contains approximately 50 figures and 40000 words including appendices, footnotes, tables and references. The following publications were derived from this work:

Journal Articles


Book Chapters


Contents

List of Figures x

List of Tables xii

Nomenclature xiii

1 Introduction 1
   1.1 3D Ultrasonic Imaging ................................... 1
   1.2 Freehand 3D Ultrasound System ............................ 3
       1.2.1 Tracking System ................................... 3
           1.2.1.1 Mechanical Arms ............................ 4
           1.2.1.2 Electromagnetic Sensors .................. 4
           1.2.1.3 Optical Sensors ............................ 5
       1.2.2 Image Acquisition System ........................... 5
       1.2.3 Patient Movement ................................ 6
       1.2.4 Probe Calibration ................................. 7

2 Probe Calibration 10
   2.1 Point Phantom ........................................ 12
       2.1.1 Point Phantom Without a Stylus ................ 12
       2.1.2 Point Phantom With a Stylus .................... 13
       2.1.3 Point Phantom Variants ........................ 15
   2.2 Stylus ................................................. 16
   2.3 Three-wire Phantom .................................. 17
   2.4 Plane Phantom ....................................... 18
   2.5 Two-plane Phantom .................................. 23
2.6 Two-dimensional Alignment Phantom .............................................. 25
  2.6.1 Z-phantom ........................................................................... 26
  2.6.2 Mechanical Instrument ............................................................... 28
2.7 Image Registration ........................................................................ 30
2.8 3D Probe Calibration ..................................................................... 31

3 Calibration Phantom Evaluation ....................................................... 33
  3.1 Calibration Quality Assessment ...................................................... 33
    3.1.1 Precision ........................................................................... 33
    3.1.2 Accuracy ........................................................................... 36
      3.1.2.1 Point Reconstruction Accuracy ...................................... 37
      3.1.2.2 Distance Accuracy ....................................................... 38
      3.1.2.3 Volume Measurement .................................................... 40
    3.1.3 Calibration Quality Comparison ............................................. 40
  3.2 Other Phantom Evaluation Factors ................................................. 41
    3.2.1 Probe Type ......................................................................... 41
    3.2.2 Ease of Use ......................................................................... 42
    3.2.3 Calibration Reliability ............................................................. 42
    3.2.4 Calibration Time ................................................................... 43
  3.3 Calibration Technique Comparison ................................................. 43
    3.3.1 Accuracy ............................................................................ 44
    3.3.2 Time .................................................................................. 44
    3.3.3 Ease of Use ......................................................................... 45

4 Reliable Calibration ........................................................................... 46
  4.1 Improvement to Single Wall Calibrations ....................................... 48
    4.1.1 Speed of Sound .................................................................. 48
    4.1.2 Explicit Scale Estimation ...................................................... 50
    4.1.3 Feedback on Calibration Reliability ...................................... 50
  4.2 Calibration Experiments ............................................................... 52
    4.2.1 Calibration Comparison ......................................................... 55
    4.2.2 Curvature of Valid Calibrations ............................................ 57
    4.2.3 Incomplete Scanning Patterns .............................................. 62
  4.3 Calibration Protocol Simulation ..................................................... 64
CONTENTS

A Plane-based Phantom 112
A.1 The Curvilinear Probe ........................................ 112
A.2 Solution of the Criterion Function ....................... 115
A.3 Verification of the sphere distances ................... 116
A.4 Eigenvalues of Valid Calibrations ..................... 117

B Z-fiducial Phantom 119
B.1 Rejection Criterion for the Z-phantom ................. 119
B.2 Correcting the Elevational Offset .................... 119

C Cambridge Stylus and Cone Phantom 121

D Algorithms and Source Code 123
D.1 Mathematical Functions ........................... 123
   D.1.1 Levenberg-Marquardt Algorithm .................. 123
   D.1.2 Open Source Libraries ........................ 123
D.2 Plane-based Calibration Simulation .................. 124
D.3 Segmentation Algorithms .............................. 124
   D.3.1 Plane-based Phantom ............................ 124
   D.3.2 Z-phantom ..................................... 125
   D.3.3 Cone Phantom .................................. 125
   D.3.4 Cambridge Stylus ............................. 125

References 140
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A 3D reconstruction of the brachial plexus.</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>The coordinates associated with a freehand 3D ultrasound system.</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>A 3D localizer.</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>B-scan images of the point phantoms.</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>The geometry of a point phantom.</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>The geometry of a pyramid phantom.</td>
<td>16</td>
</tr>
<tr>
<td>2.5</td>
<td>The geometry of a rod stylus.</td>
<td>17</td>
</tr>
<tr>
<td>2.6</td>
<td>The geometry of a three-wire phantom.</td>
<td>19</td>
</tr>
<tr>
<td>2.7</td>
<td>B-scan image of a three-wire phantom.</td>
<td>19</td>
</tr>
<tr>
<td>2.8</td>
<td>The geometry of a plane phantom.</td>
<td>20</td>
</tr>
<tr>
<td>2.9</td>
<td>The beam thickness problem and solution.</td>
<td>21</td>
</tr>
<tr>
<td>2.10</td>
<td>B-scan images of the Cambridge phantom.</td>
<td>23</td>
</tr>
<tr>
<td>2.11</td>
<td>The geometry of a two-plane phantom.</td>
<td>24</td>
</tr>
<tr>
<td>2.12</td>
<td>The geometry of a 2D alignment phantom.</td>
<td>26</td>
</tr>
<tr>
<td>2.13</td>
<td>The geometry of a Z-phantom.</td>
<td>27</td>
</tr>
<tr>
<td>2.14</td>
<td>Geometry of the mechanical device for calibration.</td>
<td>29</td>
</tr>
<tr>
<td>2.15</td>
<td>B-scans of a set of wedges.</td>
<td>30</td>
</tr>
<tr>
<td>2.16</td>
<td>Principle of the mechanical device for calibration.</td>
<td>30</td>
</tr>
<tr>
<td>3.1</td>
<td>Scanning patterns used during accuracy assessment.</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Distortion due to sound speed errors.</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Probe motions for plane-based calibration.</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Scanning patterns used in accuracy assessment.</td>
<td>56</td>
</tr>
<tr>
<td>4.4</td>
<td>Improved precision of plane-based calibrations.</td>
<td>56</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>Accuracy of plane-based calibrations.</td>
<td>57</td>
</tr>
<tr>
<td>4.6</td>
<td>Precisions of plane-based calibrations.</td>
<td>58</td>
</tr>
<tr>
<td>4.7</td>
<td>Distance accuracy of plane-based calibrations.</td>
<td>59</td>
</tr>
<tr>
<td>4.8</td>
<td>The variation of the system error.</td>
<td>60</td>
</tr>
<tr>
<td>4.9</td>
<td>Second minimum eigenvalue of the Hessian.</td>
<td>61</td>
</tr>
<tr>
<td>4.10</td>
<td>Calibration precision as calibration becomes more constrained.</td>
<td>61</td>
</tr>
<tr>
<td>4.11</td>
<td>The minimum eigenvalues when a motion is missing.</td>
<td>62</td>
</tr>
<tr>
<td>4.12</td>
<td>Comparison of calibrations when some motions are missing.</td>
<td>63</td>
</tr>
<tr>
<td>4.13</td>
<td>Minimum set of scanning patterns required for a plane-based calibration.</td>
<td>66</td>
</tr>
<tr>
<td>5.1</td>
<td>The Z-phantom.</td>
<td>73</td>
</tr>
<tr>
<td>5.2</td>
<td>Ultrasound image of the Z-phantom.</td>
<td>74</td>
</tr>
<tr>
<td>5.3</td>
<td>The top view of a Z-fiducial intersecting the scan plane.</td>
<td>75</td>
</tr>
<tr>
<td>5.4</td>
<td>Precision and accuracy of the Z-phantom.</td>
<td>80</td>
</tr>
<tr>
<td>5.5</td>
<td>Motions used to scan the membrane.</td>
<td>83</td>
</tr>
<tr>
<td>5.6</td>
<td>Improved precision and accuracy.</td>
<td>85</td>
</tr>
<tr>
<td>6.1</td>
<td>The calibration phantoms.</td>
<td>88</td>
</tr>
<tr>
<td>6.2</td>
<td>B-scan images of the calibration phantoms.</td>
<td>91</td>
</tr>
<tr>
<td>6.3</td>
<td>Misaligned B-scan images of the calibration phantoms.</td>
<td>91</td>
</tr>
<tr>
<td>6.4</td>
<td>Calibration styli to be compared.</td>
<td>93</td>
</tr>
<tr>
<td>6.5</td>
<td>B-scan images of the various styli.</td>
<td>94</td>
</tr>
<tr>
<td>6.6</td>
<td>The calibration protocol.</td>
<td>95</td>
</tr>
<tr>
<td>6.7</td>
<td>Probe calibration precisions.</td>
<td>96</td>
</tr>
<tr>
<td>6.8</td>
<td>Precision of calibrations in the three principal axes.</td>
<td>97</td>
</tr>
<tr>
<td>6.9</td>
<td>B-scan of the wire used for point reconstruction accuracy assessment.</td>
<td>97</td>
</tr>
<tr>
<td>6.10</td>
<td>Different estimates of the probe calibration accuracy.</td>
<td>98</td>
</tr>
<tr>
<td>6.11</td>
<td>Point reconstruction accuracy calculation.</td>
<td>99</td>
</tr>
<tr>
<td>A.1</td>
<td>Sound speed correction for a curvilinear probe.</td>
<td>113</td>
</tr>
<tr>
<td>A.2</td>
<td>Estimated error of temperature correction routine.</td>
<td>114</td>
</tr>
<tr>
<td>A.3</td>
<td>Effect of temperature correction routines on precision.</td>
<td>116</td>
</tr>
</tbody>
</table>
## List of Tables

4.1 Probes and other settings used in the calibrations. . . . . . . . . . . 53

5.1 Precision of the Z-phantom. . . . . . . . . . . . . . . . . . . . . . 81
5.2 Comparison of calibration precision and accuracy. . . . . . . . . . . 85

7.1 Our calibration precision and accuracy. . . . . . . . . . . . . . . . 105
7.2 Other research group’s calibration precisions. . . . . . . . . . . . . 106
7.3 Probe calibration factors for the different phantoms. . . . . . . . . . 108

A.1 Setting differences of the ultrasound machine. . . . . . . . . . . . 115
A.2 Estimation of the curvilinear probe centre. . . . . . . . . . . . . . 115
A.3 Distance between the spheres. . . . . . . . . . . . . . . . . . . . . . 117

C.1 Precision of differentiated probe calibrations. . . . . . . . . . . . . 121
C.2 Precision of undifferentiated probe calibration. . . . . . . . . . . . 122
C.3 Point reconstruction accuracies in different regions of the B-scan.  122
C.4 Bias and spread of the point reconstruction accuracies. . . . . . . . 122
Nomenclature

Roman Symbols

\( f \)  A function
\( F \)  Coordinate system of the phantom
\( G \)  Coordinate system of the gantry
\( H \)  Coordinate system of the probe holder
\( I \)  Coordinate system of the B-scan in metric units
\( I' \) Coordinate system of the B-scan in pixels
\( L \)  Coordinate system of the stylus
\( M, N \) Total number
\( p \)  A point
\( r \)  Tip of a stylus
\( R \)  Coordinate system of the wires
\( s \)  Scale factor
\( S \)  Coordinate system of the mobile part of the position sensor
\( t \)  Transpose of a matrix
\( T \)  Transformation
\( u, v \) Column and row indices of the B-scan
\( W \)  Coordinate system of the stationary part of the position sensor
\( x, y, z \) Three principal axes of a coordinate system

Greek Symbols

\( \mu \)  Measure of accuracy and precision
\( \mu_{\text{PRA}} \) Measure of Accuracy
\( \mu_{\text{CR}}, \mu_{\text{RP}} \) Measure of Precision
Superscripts

$p^A$ Coordinates of $p$ in coordinate system $A$

Subscripts

$i, j$ Subscript index

$p_x, p_y, p_z$ The $x, y, z$ coordinate of the point $p$

$T_{B\leftarrow A}$ Transformation from coordinate system $A$ to coordinate system $B$

$T_s$ Scale factor, equivalently $T_{I\leftarrow I'}$

Other Symbols

$|\cdot|$ Euclidean norm on $\mathbb{R}^3$

$\{\cdot\}$ A set

$\vec{a}$ A ray

$\bar{a}_i$ Mean of $(a_i)$

$\times$ Cross product of two vectors in $\mathbb{R}^3$

Acronyms

2D Two-dimensional
3D Three-dimensional
AC Alternating current
CMM Coordinate measuring machine
CR Calibration reproducibility
CT Computed tomography
DC Direct current
DOF Degrees of freedom
ECG Electrocardiogram
MR Magnetic resonance
PRA Point reconstruction accuracy
RF Radio-frequency
RP Reconstruction precision
Chapter 1

Introduction

1.1 3D Ultrasonic Imaging

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 (Fenster et al., 2001; Nelson & Pretorius, 1998). A 3D ultrasound imaging system has many clinical applications, some of these are: obstetrics (Gonçalves et al., 2005), gynaecology (Alcázar, 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).

There are in general four techniques to construct a 3D volume using ultrasound devices. They are:

1. 3D ultrasound probes,
2. mechanically swept probes,
3. sensorless freehand techniques, and
4. freehand techniques with a position sensor.

A 3D ultrasound probe consists of a two-dimensional (2D) array of transducer elements (Light & Smith, 2004; Light et al., 1998; Lu et al., 2006; Smith et al., 1991; Yen & Smith, 2004). These probes have the advantage that real-time
imaging is possible. However, these probes have lower resolution compared to 2D ultrasound. This is because it is difficult to wire a large number of elements in a confined space.

A mechanically swept probe uses a motor to move a one-dimensional array within the probe housing (Brandl et al., 1999). A 3D image is formed by combining the 2D slices.

When using the freehand technique, the clinician guides a conventional 2D ultrasound probe over the anatomy to gather non-uniformly spaced B-scans. If the geometric orientation of each B-scan is known in 3D space, then the volumetric data can be reconstructed. In the sensorless approach, the geometric orientation is approximated by using speckle decorrelation (Chang et al., 2003; Chen et al., 1997; Gee et al., 2006). Since each frame is located with respect to the previous frame, cumulative errors and drift are inevitable.

Another technique 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). As the probe is swept over the anatomy, the trajectory of the probe is recorded by the attached position sensor. The volume of anatomy can be constructed by matching the ultrasonic data with its corresponding position in space. This technique allows a large volume to be recorded and visualized in a fixed global coordinate system.

A freehand 3D ultrasound system is particularly useful when it is required to locate the subject in a fixed external coordinate system. In radiotherapy planning, such as partial breast irradiation (Coles et al., 2007), the tumour can be localized using a freehand 3D ultrasound system and the image co-registered with a planning Computed Tomography (CT). This allows additional information on the size and location of the surgical bed to be integrated into the radiotherapy planning system.

A mechanically swept probe or a 2D array may be useful when scanning a small region of interest. However, they are limited to a small field of view and cannot be used when it is not possible to scan the anatomy in a single sweep of the probe. Freehand 3D ultrasound is thus useful for scanning large objects such as the liver (Treece et al., 2001) or scanning around a curved surface such as the neonatal foot (Cash et al., 2005b) and the brachial plexus (Cash et al.,
Figure 1.1 shows a lateral and an anteroposterior view of the brachial plexus acquired using freehand 3D ultrasound.

Figure 1.1: An example of a surface rendered 3D reconstruction of the brachial plexus in relation to the carotid artery, subclavian artery and the first rib (Cash et al., 2005a).

### 1.2 Freehand 3D Ultrasound System

In a freehand 3D ultrasound system, the probe is tracked by the tracking system. The choice of the tracking system is therefore one of the factors to be considered when constructing a freehand 3D ultrasound system.

#### 1.2.1 Tracking System

There are generally three types of tracking systems, namely:

1. mechanical arms,
2. electromagnetic sensors, and
3. optical sensors.
1.2 Freehand 3D Ultrasound System

1.2.1.1 Mechanical Arms

The ultrasound probe can be attached to a mechanical arm with six degrees of freedom (DOF). The geometric properties of the mechanical arm are precisely defined and there are encoders at each articulation to report the angles in real-time. The exact position of the end of the arm may then be computed from the geometric model of the arm. This technology has the advantage that the position of the arm is always available, with no possibility of obstruction unlike with other non-contact tracking systems. An accuracy of 0.1mm has being achieved by the multi-articular portable coordinate measuring machine (CMM) (Mitutoyo, Japan). The manufacturer has quoted a measuring envelope between 1.8–3.6m for the portable CMM. It is also possible to fix the probe in a certain position in space with this technology. The major draw-back of such systems is that they are expensive and considered cumbersome. Each mechanical arm can only track one object at any one time. The accuracy may also be affected if there is a force being applied to the arm without moving it (Bosch, 1995).

1.2.1.2 Electromagnetic Sensors

The receiver of an electromagnetic system is placed on the probe to measure the induced electrical current when moved within the magnetic field generated by the transmitter. The magnetic field may be generated by alternating currents (AC) or direct currents (DC). The accuracy of an AC electromagnetic device is 0.7mm for a Fastrak (Polhemus, U.S.A.), 0.7–1.9mm for an Aurora (Northern Digital Inc, Canada) and 1.8mm for a DC Flock of Birds (Ascension Technology Corporation, U.S.A). A limitation of the electromagnetic tracking system is that the receiver needs to be placed near the transmitter, within the generated magnetic field. The accuracy of the Aurora rapidly degrades to 1.3–1.6mm when the receiver is placed 400-500mm from the generator. The performance of electromagnetic sensors can be impeded by the presence of metallic (AC sensors) and ferromagnetic material (DC sensors) (Birkfellner et al., 1998). Hummel et al. (2006) showed that an error of 9.4mm can occur when a stainless steel rod is placed 10mm away from the transmitter. It therefore remains a challenging task for such systems to
1.2 Freehand 3D Ultrasound System

be used in medical surgery, where various metal objects are continuously being moved around.

1.2.1.3 Optical Sensors

Optical tracking systems consist of multiple cameras observing one or more targets. The target consists of passive or active markers with a predefined non-symmetrical geometry on a rigid structure. At least three markers are necessary so that the position and orientation of the target can be accurately located in 3D space. In systems where more than three markers are used, there will be redundant data and therefore a more accurate estimation of the target’s position. Accuracies of 0.35mm for the Polaris (Northern Digital Inc, Canada) and 0.15mm for the Certus (Northern Digital Inc, Canada) are quoted by the manufacturer. They also claim that there is no noticeable difference whether active or passive markers are used (Wiles et al., 2004). The standard measurement length covered by the Polaris is 1.5m, but the length covered by the Certus can be as long as 4.5m. A disadvantage of using an optical tracking system is that a direct line of sight needs to be maintained between the cameras and the tracked targets.

1.2.2 Image Acquisition System

In ultrasonic imaging, the beam former transmits ultrasound pulses into the body, which are reflected by the tissue and then received by the transducer elements. This is the raw radio-frequency (RF) signal. For tissue imaging, the amplitude of the backscattered RF signal is a function of its depth. Hence an internal image can be formed after time gain compensation and band-pass filtering. This is passed to a scan converter unit to transform the signals into a viewable image.

In modern freehand 3D ultrasound, the ultrasonic data is transferred to a computer for display and processing. One way to transfer the images from the ultrasound machine to the computer is to digitize the video output port of the ultrasound machine by using a frame grabber card (Huang et al., 2005; Prager et al., 1999; Varandas et al., 2004).

It is also possible to transfer the internal ultrasonic data directly to the computer (Gronningsaeter et al., 2000; Treece et al., 2003). Unlike a video output
port that is readily available on many ultrasound machines, access to the internal ultrasound signals may require a special arrangement with the ultrasound manufacturer. Scan conversion thus becomes the responsibility of the computer software to display the data on the computer monitor.

The computer that is used for image processing and display is also responsible for synchronizing data from the ultrasound machine and the tracking system. The computer sends requests to the associated hardware and time-stamps the ultrasonic and position data when it arrives. However, these data arrive with an unknown delay after having been generated by the hardware. A temporal calibration may need to be performed to find the relative delay between the two data streams so that matching B-scans and positions are available for further analysis (Gooding et al., 2005; Nakamoto et al., 2003; Prager et al., 1999; Rousseau et al., 2006; Treece et al., 2003). The residual misalignment after calibration is typically less than 10ms (Treece et al., 2003). Since the patient is scanned between 10–20mm/s, the error due to temporal misalignment is approximately 0.1–0.2mm.

1.2.3 Patient Movement

When a patient is being scanned using a freehand 3D ultrasound system, it is important that the patient does not move during image acquisition. This is different to a mechanically swept probe or a 2D array, where there is little distortion if the patient moves together with the probe.

During image-guided radiosurgery, where the treatment beam is aligned to the patient, a conformable plastic mask can be used to fix the patient onto the surgery bed. Murphy et al. (2003) showed that the patient remained, on average, within 0.85mm of the intended position over short durations of 1–5 minutes.

If freehand 3D ultrasound is used to scan areas near the thoracic cavity, such as breasts and lungs, the images will be affected by breathing artifacts. It is necessary to complete the scan in a single breath hold of approximately 10–15 seconds to minimize breathing artifacts. Murphy et al. (2003) showed in their study that targets in the lungs remained within 2mm of their imaged location. This margin includes patient movements mentioned in the previous paragraph.
Anatomical movement may also be caused by pulsatile motion due to blood flow. Position fluctuation increases to 2.6mm in the pancreas as a result of aortic pulsation (Murphy et al., 2003). Distortions caused by blood flow pulsation may be reduced by gating the image acquisition system to an electrocardiogram (ECG) (Barry et al., 1997; Meairs et al., 2000), so that each image is captured at the same point during each cardiac cycle.

Another type of patient movement is caused by the image acquisition process itself. The clinician inevitably exerts a pressure on the skin when scanning the patient with the ultrasound probe. This causes the anatomy to deform locally where the pressure is exerted. Nevertheless, probe pressure artifacts may be partially corrected using a combination of image based and position sensing techniques (Treece et al., 2002).

### 1.2.4 Probe Calibration

One complication of any freehand 3D ultrasound system is that the tracking system records the 3D location of the position sensor $S$, rather than the scan plane $I$, relative to its stationary counterpart $W$ as shown in Figure 1.2. It is therefore necessary to find the position and orientation of the scan plane with respect to the coordinate system of the position sensor. This rigid-body transformation $^{1}T_{S \leftarrow I}$ 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.

The stationary part of the tracking system defines the world coordinate system, and the term position sensor is used to mean its mobile counterpart. We will also follow these conventions in this thesis. In general, a transformation involves both a rotation and a translation in 3D space. For brevity, we will use the notation

$$^{1}T_{P} = \begin{pmatrix}
\cos \alpha \cos \beta & \cos \alpha \sin \beta \sin \gamma - \sin \alpha \cos \gamma & \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma \\
\sin \alpha \cos \beta & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma \\
- \sin \beta & \cos \beta \sin \gamma & \cos \beta \cos \gamma
\end{pmatrix} p + 
\begin{pmatrix}
t_{x} \\
t_{y} \\
t_{z}
\end{pmatrix},$$

where $t_{x}, t_{y}$ and $t_{z}$ are the translations and $\gamma, \beta$ and $\alpha$ are the rotations about the $x, y$ and $z$ axes.
T_{B\rightarrow A} to mean a rotational transformation followed by a translation from the coordinate system \(A\) to coordinate system \(B\).

Figure 1.2: The coordinates associated with a freehand 3D ultrasound system.

Another issue before we can construct a volume in space is to determine the scales of the B-scans. A point \(p'' = (u, v, 0)^t\) 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 \(T_s = \begin{pmatrix} s_u & 0 & 0 \\ 0 & s_v & 0 \\ 0 & 0 & 0 \end{pmatrix}\), 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 \(p' = T_s p''\). In this thesis, 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:

\[
p^F = T_{F\rightarrow W} T_{W\rightarrow S} T_{S\rightarrow I} T_s p''.
\]

In the above equation, \(T_{W\rightarrow S}\) can be read from the position sensor. The transformation from the world coordinate system to a phantom coordinate system \(T_{F\rightarrow W}\) 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 thesis, how the choice of \(T_{F\rightarrow W}\) will help us to find the calibration parameters.

Accurate calibration is vital for accurate freehand 3D ultrasound, since it is one of the dominant errors in the system (Treece et al., 2003). If the probe is
incorrectly calibrated, the system can produce dangerously plausible, but inaccurate images and measurements.

The aims of this thesis are to provide reliable, rapid, easy and accurate calibration techniques. The thesis is organized in the following manner to address these aims. We will start by reviewing different calibration techniques in Chapter 2. Chapter 3 discusses how calibration accuracies are assessed. Chapters 4–6 of this thesis focus on improving current calibration techniques in terms of calibration reliability, speed, ease of use and accuracy. Finally, we will give a comprehensive comparison of current state-of-the-art calibration techniques in Chapter 6 and conclude in Chapter 7.
Chapter 2

Probe Calibration

Probe calibration has been an active research topic for many years (Mercier et al., 2005). In this chapter, we will review existing probe calibration techniques and classify each technique accordingly to its principles.

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 & Jan, 2005). This is a 3D localizer, often called a pointer or a stylus. Figure 2.1(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 2.1(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:

\[
\begin{align*}
  r^W &= T_{W\leftarrow L} r^L, \\
     &\quad (2.1)
\end{align*}
\]

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 & 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. A pointer calibration is fairly simple to
perform compared to a probe calibration. A valid pointer calibration can be obtained provided that the stylus is rotated at as many orientations as possible within the tracking system’s volume. 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 2.1 to solve for the position of the stylus tip, by minimizing

$$\sum_i |r_i^W - T_{W\rightarrow L} r_i^L|,$$

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 errors include tracking the stylus and pointer calibration errors. 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 (Khamene & Sauer, 2005; Leotta, 2004; Leotta et al., 1997; Pagoulatos et al., 1998). The tracking error is dependent on the tracking system. A typical optical tracking system, such as the Polaris, has a RMS tracking error of 0.35mm. In general, a stylus has a positioning uncertainty of approximately 1mm. This does not include hand-eye uncertainties that depend on the point to be localized and user expertise.
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 (Barry et al., 1997; Detmer et al., 1994; Huang et al., 2005; Krupa, 2006) or a spherical bead-like object (Barratt et al., 2006; Legget et al., 1998; Leotta et al., 1997; Pagoulatos et al., 1999; State et al., 1994). An example of B-scan images of the cross-wire and spherical point phantoms is shown in Figure 2.2. 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 \( p'' = (u, v, 0)^t \) 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 2.3, then Equation 1.1 becomes

\[
T_{F\leftarrow W}T_{W\leftarrow S}T_{S\leftarrow I}T_s \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.3}
\]

This is an equation with 11 unknowns—two scale factors, six calibration parameters and three parameters from \( T_{F\leftarrow W} \). Only the three translations in \( T_{F\leftarrow W} \) need to be determined, since we do not not care about the orientation of \( F \), hence we can set the three rotations in \( T_{F\leftarrow W} \) to arbitrary values, such as zeroes. If we
2.1 Point Phantom

capture $N$ images of the point phantom from many directions and orientations, we can find these 11 unknowns by minimizing

$$f_{\text{point1}} = \sum_{i=1}^{N} \left| T_{F \leftarrow W} T_{W \leftarrow S_i} T_{S \leftarrow I} T_{S} p_{i}' \right|,$$  \hspace{1cm} (2.4)

with the three rotations in $T_{F \leftarrow W}$ 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 $T_{F \leftarrow W}$ may be discarded and replaced with an alternative $T_{F \leftarrow W}$ that is convenient for visualization.

![Cross-wire images](image1.png)

(a) Cross-wire (aligned)  \hspace{1cm} (b) Cross-wire (misaligned)  \hspace{1cm} (c) Spherical phantom

Figure 2.2: B-scan images of the point phantoms.

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 Péria 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

$$f_{\text{point2}} = \sum_{i=1}^{N} \left| p^W - T_{W \leftarrow S_i} T_{S \leftarrow I} T_{S} p_{i}' \right|.$$ \hspace{1cm} (2.5)

There is little to be gained over Equation 2.4, 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.,
2.1 Point Phantom

2001) involving Kronecker products (Brewer, 1978) to solve the calibration parameters and the image scales, but an iterative minimization algorithm is still required.

Sometimes the scales may be supplied by the manufacturer (Boctor et al., 2003). If the scales can be found (Péria et al., 1995), then the segmented image of the point phantom is known in millimetres: \( p^l = T_s p^l' \). 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. \( p^S = T_{W→S}^{-1} p^W \). This means that the point phantom is known in the two coordinate system \( I \) and \( S \), and we want to find a transformation \( T_{S→I} \) that best transforms \( \{p^l_i\} \) to \( \{p^S_i\} \). This transformation is also found by minimizing Equation 2.5. However, since the image scales are known, a closed form solution is available provided that the point phantom 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 \( f_{point2} \) with known scales 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.

![Figure 2.3: The geometry of a point phantom.](image-url)
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, segmentation of isolated points in ultrasonic images is seldom reliable. This is evident from the fact that most 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\(^1\) centre of the scan plane. This misalignment 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 as shown in Figure 2.4. 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, shown as circles in Figure 2.4, act as three distinct point targets that have been scanned.

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

\(^1\)In this thesis, we follow the convention where the lateral and axial axis correspond to the principal axes of the scan plane, where the axial axes is parallel to the direction of propagation of the ultrasound wave, and the elevational axis is perpendicular to the scan plane.
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 a good alignment, while maintaining automatic segmentation. The fact that circles can be segmented automatically makes calibration simpler and quicker to perform.

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 & 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

$$f_{\text{stylus}} = \sum_{i=1}^{N} \left| T_{W-L_i}r^L - T_{W-s_i}T_{s-i}T_{sp_i}r' \right|,$$

(2.6)

where $T_{W-L_i}r^L$ 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. Khamene & Sauer (2005) solve the alignment problem by attaching a rod to a position sensor, as shown in Figure 2.5. Both ends of the rod are pointer calibrated, their locations in space are therefore $T_W \leftarrow L r_1^L$ and $T_W \leftarrow L r_2^L$. 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 1.1, 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

$$\left| (T_{W\leftarrow L} r_2^L - T_{W\leftarrow L} r_1^L) \times (T_{W\leftarrow L} r_1^L - T_{W\leftarrow S} T_{S\leftarrow I} T_s p_i') \right| = 0.$$ 

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

$$f_{rod} = \sum_{i=1}^{N} \left| (T_{W\leftarrow L_i} r_2^L - T_{W\leftarrow L_i} r_1^L) \times (T_{W\leftarrow L_i} r_1^L - T_{W\leftarrow S_i} T_{S\leftarrow I} T_s p_i') \right|.$$ 

This is an equation with six unknowns—the six calibration parameters in $T_{S\leftarrow I}$.

![Figure 2.5: The geometry of a rod stylus.](image)

## 2.3 Three-wire Phantom

The three-wire phantom is solely used by Carr (1996) and Prager et al. (1998). Instead of mounting a pair of cross-wires in the solution, three mutually ortho-
2.4 Plane Phantom

ONAL wires are mounted. These three wires form the three principal axes of the phantom coordinate system as shown in Figure 2.6. 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

\[
\begin{pmatrix}
T_{F-w}W_{S-1}T_s \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{p^F_x}{p^F_y} = 0 \\ \frac{p^F_y}{p^F_x} = 0 \\ \frac{p^F_x}{p^F_y} = 0 \end{pmatrix}.
\]

(2.8)

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

\[
f_{3\text{-wire}} = \sum_{i=1}^{N_x} \left( \left( p^F_{ix} \right)^2 + \left( p^F_{iz} \right)^2 \right) + \sum_{i=N_x+1}^{N_x+N_y} \left( \left( p^F_{iy} \right)^2 + \left( p^F_{iz} \right)^2 \right) + \sum_{i=N_x+N_y+1}^{N_x+N_y+N_z} \left( \left( p^F_{ix} \right)^2 + \left( p^F_{iy} \right)^2 \right).
\]

(2.9)

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.

This technique does not require the user to align the scan plane with the phantom and potentially speeds up the calibration process. The wires can be scanned perpendicularly to the scan plane to ensure a good image of the wire is obtained as shown in Figure 2.7. However, 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 (Baumann et al., 2006;
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 2.8, then every point on the line in the image must satisfy Equation 1.1, i.e.

\[
\begin{pmatrix}
u \\
u
\end{pmatrix} = \begin{pmatrix}
p_x^F \\
p_y^F \\
p_z^F = 0
\end{pmatrix}.
\]

The equation for the \( z \) coordinate of the phantom coordinate system is the re-

---

**Figure 2.6:** The geometry of a three-wire phantom.

**Figure 2.7:** B-scan image of a three-wire phantom.
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

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

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 \( T_{F\leftarrow W} \). 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 the normal of the plane will be absent in the equation.

![Figure 2.8: The geometry of a plane phantom.](image)

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
a wavelet-based technique (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 & 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 2.9(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 2.9: The beam thickness problem and solution.](image)

It is possible to use a Cambridge phantom shown in Figure 2.9(b) to solve the beam thickness problem and to ensure that a strong echo is received when scanning at oblique positions. 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 the
brass bar. In order to assist the user to align the scan plane, a set of wedges (Figure 2.14(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 as shown in Figures 2.10(a) and (b). The user can scan the plane from different angles and still get a clear image. The beam thickness problem is solved since a thin bar is scanned. However, if a curvilinear probe is calibrated at a high depth setting, the reverberation due to the clamp may degrade the image so badly that the brass bar is undetectable as shown in Figure 2.10(c). In any case, 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. An incorrect calibration is often obtained because the phantom has not been scanned from a sufficiently diverse set of positions and orientations.

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 2.8. 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.

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 2.11, 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
2.5 Two-plane Phantom

satisfy:
\[
\begin{pmatrix}
p'_x^F \\
p'_y^F = 0 \\
p'_z^F
\end{pmatrix}
= T_{F\rightarrow W} T_{W\rightarrow S} T_{S\rightarrow I} T_s p_h^F.
\] (2.12)

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 parameters and the scales can be found by minimizing

\[
f_{\text{2-plane}} = \sum_{i=1}^{N} \left( \sum_{j=1}^{M_h} \left( p_{i,hj}^F \right)^2 + \sum_{j=1}^{M_v} \left( p_{i,vj}^F \right)^2 \right),
\] (2.13)

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

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
2.6 Two-dimensional Alignment Phantom

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 posi-
tioned 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) were the first to use such a phantom. They scanned a thin board
with three vertices as shown in Figure 2.12. 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 $f_{\text{point2}}$. If we
have captured $N$ images of a two-dimensional alignment phantom with $M$
fiducial points, calibration is found by minimizing

$$f_{2\text{D}} = \sum_{i=1}^{N} \sum_{j=1}^{M} \left| T_{W-S_i}^{-1} p_j^W - T_{S-j} S_i p_j^I' \right|. \quad (2.14)$$

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
2.6 Two-dimensional Alignment Phantom

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 to align the phantom with the point fiducials, which can be very difficult. An advantage is that only one frame ($N = 1$ in Equation 2.14) 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 2.13. 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 (Boctor et al., 2004; Gobbi et al., 1999; Pagoulatos et al., 2001; Zhang et al., 2004). It is possible to attach a position
sensor to the Z-phantom \cite{Chen2006, Lindseth2003b} and precision manufacture the wires relative to the position sensor \cite{Lindseth2003a}. This requires the coordinates of the position sensor to be known. The position sensor should be mounted close to the fiducials to minimize positioning errors. 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.

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:

$$z^W = w^W_3 + \frac{|z - w_3|}{|w_2 - w_3|} \left( w^W_2 - w^W_3 \right)$$

$$= w^W_3 + \frac{|z - b|}{|a - b|} \left( w^W_2 - w^W_3 \right), \quad (2.15)$$

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.

Figure 2.13: The geometry of a Z-phantom.
by minimizing

\[ f_{Z\text{-phantom}} = \sum_{i=1}^{N} \sum_{j=1}^{M} \left| T_{W-S_i}^{-1} z_{ij}^W - T_{S-I} T_{S-H} z_{ij}^H \right|, \]

where \( z_{ij} \) is the \( j \)th Z-fiducial in the \( i \)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.

It may be possible to segment the wires automatically. Chen et al. (2006) simplified their phantom to just two ‘Z’ wire configurations. The probe is also constrained by the phantom so that the phantom can only be scanned from one orientation. Their segmentation algorithm involves finding two sets of parallel 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 mounted onto the probe holder at a fixed location as well. The transformation \( T_{S-H} \) is therefore also fixed. In order to find this transformation, the probe holder is

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 2.14(a). The transformation \( T_{H-G} \) is therefore known by construction. The position sensor is mounted onto the probe holder at a fixed location as well. The transformation \( T_{S-H} \) 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. 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.

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 2.14(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 \( T_{R^{-I}} \) 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 \( \overrightarrow{p_2p_3} \).

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 2.14(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 as shown in Figure 2.15(a). Figure 2.15(b) shows a B-scan image of the wedges slightly misaligned. The surfaces 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.
2.7 Image Registration

Now, once the two calibrations are complete, the transformation that relates the wires to the gantry $T_{G\rightarrow R}$ 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 2.16. Calibration is therefore given by

$$T_{S\rightarrow I} = T_{S\rightarrow H}T_{H\rightarrow G}T_{G\rightarrow R}T_{R\rightarrow I}. \tag{2.17}$$

Figure 2.15: B-scans of a set of wedges.

Figure 2.16: 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 was acquired by a magnetic resonance (MR) scan. The phantom was scanned with a freehand 3D ultrasound system. A 3D image of the phantom was reconstructed by using an assumed calibration and Equation 1.1. An iterative optimisation algorithm was implemented to find the calibration and the image scales where the reconstructed image best fitted the MR model. The similarity measure between two 3D volumes $A$ and $B$ is given by their mutual information (Studholme et al., 1999):

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

(2.18)

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; Zitova & Flusser, 2003) is unknown.

### 2.8 3D Probe Calibration

Although it is not the main focus of this thesis 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 (Poon & Rohling, 2007; Sawada et al., 2004) 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 & Eulenstein (2002) and Hastenteufel et al. (2003) used an image registration technique. Poon & 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.
Chapter 3

Calibration Phantom Evaluation

3.1 Calibration Quality Assessment

In Chapter 1, we have discussed different error sources in a freehand 3D ultrasound system. Probe calibration is one of the main sources of error, 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.

3.1.1 Precision

One of the first measures used was formulated by Detmer et al. (1994) and used by various other research groups (Blackall et al., 2000; Brendel et al., 2004; Dandekar et al., 2005; Leotta et al., 1997; Meairs et al., 2000; Muratore & Galloway Jr., 2001; Prager et al., 1998). Now commonly named the reconstruction precision (RP), this measure is calculated by scanning a point phantom $p$ from different
3.1 Calibration Quality Assessment

positions and orientations. The point phantom is segmented in the B-scans and reconstructed in 3D space by using Equation 1.1. 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.

\[
\mu_{RP1} = \frac{1}{N} \sum_{i=1}^{N} \left| T_{W \leftarrow S_i} T_{S \leftarrow I_i} T_s p'_i - \overline{p'_i} \right|.
\]

(3.1)

This equation can be generalized to include multiple calibrations:

\[
\mu_{RP2} = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \left| T_{W \leftarrow S_i} T_{S \leftarrow I_j} T_s p'_i - \overline{p'_{ij}} \right|,
\]

(3.2)

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. 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 $p''$ on the B-scan. Calibration reproducibility is computed as follows:

$$
\mu_{CR} = \frac{1}{N} \sum_{i=1}^{N} \left| T_{S^{-i}i} T_s p'' - \overline{p_{S_i}} \right|.
$$  \hspace{1cm} (3.3)

During calibrations with a stylus (Section 2.2), the variation due to pointer calibration has an impact on probe calibration reproducibility. Calibration reproducibility should therefore incorporate pointer calibration imprecision. Suppose that $M$ pointer calibrations are performed, and $N$ probe calibrations are performed for every pointer calibration, then calibration reproducibility including pointer calibration imprecision is:

$$
\mu_{CR'} = \frac{1}{M N} \sum_{j=1}^{M} \sum_{i=1}^{N} \left| T_{S^{-i}i} T_s p''_j - \overline{p_{S_j}} \right|,
$$  \hspace{1cm} (3.4)

where $p_{S_i}^j$ denotes the point mapped by the $i^{th}$ probe calibration associated with the $j^{th}$ pointer calibration, and $\overline{p_{S_j}}$ is the mean of $(p_{S_1}^j, p_{S_2}^j, \cdots, p_{S_N}^j)$ for a given $j$.

Clearly, the measure for calibration reproducibility is not just dependent on the calibrations ($T_{S^{-i}i}$), but also on the point $p''$. When Prager et al. (1998) first used this measure, they chose $p''$ to be the centre of the image. Many research groups also gave the variation at the centre of the image (Lindseth et al., 2003b; Meairs et al., 2000). 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 at 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 (Gee et al., 2005; Treece et al., 2003). Leotta (2004) and the Cambridge group (Gee et al., 2005; Treece et al., 2003) also quoted the spread at 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 arising from 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 (Leotta, 2004; Pagoulatos et al., 2001), but 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.

3.1.2 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 (Ali & Logeswaran, 2007; Meairs *et al.*., 2000), 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 *in vivo* images, phantom images usually have a clear border due to a higher acoustic impedance difference between water and the phantom. As a consequence, 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 might 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.

### 3.1.2.1 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 became 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 & 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
3.1 Calibration Quality Assessment

and the stylus reading, i.e.

\[ \mu_{\text{PRA}} = p^W - T_{W \rightarrow S} T_{S \rightarrow I} T_s p'' . \]  

(3.5)

This is a measurement of system accuracy, and includes errors from every component of the system. One possible 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 can be difficult due to the thick beam width. It is important to align the point phantom with care to minimize misalignment errors during calibration accuracy assessment. There are other sources of error such as segmentation error and position sensor error, and 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 (Chen et al., 2006; Liu et al., 1998). This means that point reconstruction accuracy is not very appropriate to assess the 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 would remain unnoticed.

3.1.2.2 Distance Accuracy

Before the stylus was developed enabling the evaluation of point reconstruction accuracy, many groups assessed accuracy by measuring the distances between
3.1 Calibration Quality Assessment

objects (Blackall et al., 2000; Boctor et al., 2003; Dandekar et al., 2005; Krupa, 2006; Leotta, 2004; Leotta et al., 1997; Lindseth et al., 2003b; Prager et al., 1998). This technique is popular because the experiment is easy to set up. A phantom is manufactured with distinct landmarks. Even though the exact locations 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 distances are correct. The measure is

\[
\mu_{\text{Distance}} = \left| p_W^1 - p_W^2 \right| - \left| T_{W \leftarrow S_1} T_{S \leftarrow I} T_S p_1' - T_{W \leftarrow S_2} T_{S \leftarrow I} T_S p_2' \right|. \tag{3.6}
\]

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 3.1(a). Accuracy assessment performed in this way is incorrect. If the calibration is wrong in some way, 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 such a high distance measurement accuracy.

Figure 3.1: 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 3.1(b) and (c). This ensures that any calibration errors will map different
3.1 Calibration Quality Assessment

points on the line in a different direction. The line image will be a distorted curve for incorrect calibrations, and the distance between the two end points will be incorrect.

3.1.2.3 Volume Measurement

Some researchers produced a phantom with a known volume (Dandekar et al., 2005; Rousseau et al., 2005). The phantom is scanned and reconstructed in world space. The volume of the phantom can be calculated from the 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. 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. This is usually not the case and the user ends up with an invalid evaluation of their volume measurement algorithm.

3.1.3 Calibration Quality Comparison

Each of the three accuracy measures has their own importance in different clinical applications. In radiotherapy planning, where the 3D location of the tumour is required in a fixed coordinate system, point reconstruction accuracy should be used to determine the margin in the treatment dose volume. Distance accuracy is particularly important in Atherosclerosis diagnosis, where the intima-media thickness in the carotid artery is measured. Often in cancer diagnosis and treatment, the clinician is concerned about the volume of the tumour, rather than its absolute location. Hence errors in volume measurements become important in these situations.

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 the different research groups. However, even for calibration reproducibility, which does not contain user induced errors other than errors from calibration itself, it is difficult to compare
3.2 Other Phantom Evaluation Factors

Even though accuracy may be an important factor used to evaluate any calibration technique, there are several other factors that may be just as important. These include probe type, reliability, ease of use and calibration time.

3.2.1 Probe Type

There is a large difference between calibrating a linear and a curvilinear probe. 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 travels at a speed similar to soft tissue, the distortions will cause the plane to appear as a curve, and not a line. For a high curvature probe, the image needs to be rectified for accurate segmentation. If the Cambridge phantom is used to calibrate a high curvature probe at a high depth setting, the reverberation due to the clamp may degrade the image so badly that the brass bar is undetectable.
3.2 Other Phantom Evaluation Factors

3.2.2 Ease of Use

The difficulty of each calibration technique is dependent on user expertise. The 2D alignment phantoms require precise alignment of the whole phantom with the scan plane. This may be difficult to achieve. The user will probably take a very long time to complete the task. Similarly, a point or a plane phantom is very difficult to calibrate for an inexperienced user. It is crucial to scan the phantom from a wide variety of positions. The Cambridge phantom requires the user to mount the probe accurately, which is also a skilled task. However, the Cambridge phantom may be easy to use because a strong echo is always received (Prager et al., 1998). On the other hand, the Z-phantom does not need any alignment (Comeau et al., 2000). Not much skill or experience is required to calibrate a probe with this phantom.

3.2.3 Calibration Reliability

Calibration reliability is the trustworthiness of a given calibration, bounded by the precision and accuracy achievable with the calibration technique, when the user is satisfied that the probe has been calibrated successfully. For some calibration techniques, such as the plane phantom, it is important that the phantom has been scanned from a wide variety of positions to sufficiently constrain the calibration parameters. If the scanning protocol is not followed and a certain scan motion is neglected inadvertently, the calibration parameters may be under-constrained leading to an incorrect solution (Prager et al., 1998; Treece et al., 2003). The calibration obtained using such techniques may be unreliable.

Other phantoms, such as the point phantom with the aid of a stylus, place fiducial points directly in the sensor and the B-scan’s coordinate systems (Muratore & Galloway Jr., 2001). The only requirement is that three images of the phantom have been captured in non-collinear locations in the B-scans, while capturing more images makes the calibration more accurate. The calibrations produced are reliable, since the user does not need to follow a sophisticated protocol.

Calibration Reliability is slightly different to ease of use for a novice. A 2D alignment phantom is difficult to use for both experts and novices, because aligning the scan plane with a 2D object is difficult. However, once the user is
satisfied that the plane has been aligned, the resulting calibration is trustworthy and it is therefore reliable.

3.2.4 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 (Detmer et al., 1994; Huang et al., 2005). Nevertheless, automatic segmentation algorithms have been implemented (Gooding et al., 2005; Krupa, 2006). The automatic segmentation of the plane phantom makes it attractive to use, as calibration time is shortened considerably (Prager et al., 1998).

3.3 Calibration Technique Comparison

In the literature, the authors evaluated their calibration techniques with a heavy emphasis on precision and accuracy. This is understandable because a freehand 3D ultrasound system will be used in a clinical environment where accuracy is crucial. Gee et al. (2005) showed their mechanical instrument can perform probe calibration with an accuracy of 0.15mm, outperforming every other calibration technique. As a manufacturer of freehand 3D ultrasound systems, acquiring a mechanical instrument is economically feasible and is also the preferred approach. The manufacturer will calibrate the probe for a range of depth settings and the user will just load the correct calibration each time the system is used.

However, in many circumstances calibration is not the responsibility of the manufacturer, but rather the responsibility of the user. Many hospitals already have their own conventional 2D ultrasound machine and may simply purchase the freehand system as an add-on to their existing machine. Since the freehand system and the ultrasound machine are purchased from different suppliers, the manufacturer of the freehand system will not be able to calibrate the ultrasound probe.

If we rely on the user to produce an accurate calibration and when purchasing a mechanical instrument is uneconomical, the preferred calibration approach will be dependent on the factors described earlier in this chapter.
3.3 Calibration Technique Comparison

3.3.1 Accuracy

If accuracy remains the main concern, the Cambridge phantom variation of plane-based approaches can be used to achieve an accurate probe calibration (Treece et al., 2003). It was the most accurate calibration technique before been surpassed by the mechanical instrument (Gee et al., 2005). The plane-based calibration routine has been built into many freehand 3D systems (Prager et al., 1999; Rousseau et al., 2005; Treece et al., 2003; Varandas et al., 2004). The main drawback of the plane-based technique is that an expert is required to perform the calibration reliably. The current system cannot detect an unreliable calibration performed by a novice user. If the calibration is unreliable, the freehand 3D ultrasound system will produce plausible, but incorrect results. In Chapter 4, we will show how we can provide feedback during a plane-based calibration to ensure that the user has got a reliable calibration.

3.3.2 Time

In addition to the high accuracy achievable with plane-based techniques, they have gained popularity because they employ fully automatic segmentation allowing rapid calibration. Calibrating with a point phantom with manual segmentation usually takes 1–2 hours (Barratt et al., 2006; Huang et al., 2005). By employing an automatic segmentation algorithm, calibration time is shortened to a few minutes (Prager et al., 1998).

Although calibration time is shortened considerably with plane-based phantoms, it is still necessary to scan the plane from many positions and orientations, which takes a few minutes to complete. A more rapid calibration technique may be useful when calibration cannot be performed off-line by an expert. One possible application is in computer assisted surgery. The position sensor may need to be removed so that both the probe and the sensor can be sterilized. Probe calibration therefore needs to be performed on-the-spot in the operation room.

Calibrating with a two-dimensional alignment phantom only requires a single frame (Leotta, 2004; Sato et al., 1998). The main drawback of this approach is the difficulty in aligning the scan plane with the phantom. Although the Z-phantom variation (Comeau et al., 1998; Lindseth et al., 2003b) solves the
alignment problem, the absence of a reliable automatic segmentation algorithm prevents calibration from been performed in less than a few minutes. In Chapter 5 we will describe a variation of the Z-phantom that employs fully automatic segmentation enabling rapid calibration in a few seconds.

3.3.3 Ease of Use

Although the plane-based calibration is accurate, an expert is required to perform the calibration reliably (Treece et al., 2003). On the other hand, a Z-phantom is easy to use, but the accuracy is relatively poor (Lindseth et al., 2003b). In Chapter 6, we describe two novel phantoms that are easy to use and provide reliable and accurate calibrations.
Chapter 4

Reliable Calibration

One of the most accurate and widely used calibration techniques is the Cambridge phantom variant of the plane-based approach (Treece et al., 2003). Recall from Section 2.4 that a major difficulty is to scan the plane from a wide variety of positions. As the phantom is scanned from different positions and orientations, the phantom images are segmented and mapped to the phantom coordinate system. These points place constraints on the calibration parameters and the parameters are solved using an iterative optimisation algorithm. If the user neglects a particular motion when scanning the phantom, the calibration parameters become under-constrained and an incorrect minimum is obtained. Also, if sound travels at a different speed in the solution in which calibration is performed to soft tissue, the line image of the phantom may be distorted. Detection of such curves without modifying the segmentation algorithm may also lead to an incorrect calibration.

This chapter describes some refinements to make the process of single-wall calibration easier and more reliable for inexperienced users to perform (Hsu et al., 2006a). There are four main problems:

(1) the speed of sound in water at room temperature is slower than its speed in average soft tissue. The sound speed in water at 20 Celsius is only 1482m/s (Bilaniuk & Wong, 1993), while sound travels at 1540m/s in average soft tissue. For accurate calibration the water bath should be at 48 Celsius which gives a sound speed of 1540m/s. This approach was adopted by Doctor et al. (2003) and Treece et al. (2003), but the water temperature is difficult to maintain, and uncomfortable for the operator. A simple alternative is to add glycerol (Gobbi
et al., 2000) or ethanol (Martin & Spinks, 2001) to the water solution, so that sound travels at the required speed of 1540m/s in the solution. Lindseth et al. (2003b) modified the sound speed assumed by the ultrasound machine directly to correspond to the speed in cold water. This simple and easy approach may require a special arrangement with the manufacturer. Trobaugh et al. (1994) and Pagoulatos et al. (2001) made provisions by post-processing the B-scan images. Although the correction using this technique may be simple for linear probes, it is more complicated for curvilinear probes. Nevertheless, we have used an image processing technique in this chapter, since glycerol may not be readily available in every situation. This technique may be easier to use for novice users and still allows the user to use a glycerol / ethanol solution by pretending the temperature is 48 Celsius. If the calibration is performed in cold water without any adjustments, the sound speed difference would cause the phantom to appear further from the probe by 3.7% during calibration.

(2) Currently, the image scales are solved together with the rest of the calibration parameters during optimization. We will show how to find the image scales explicitly. Once these values are found, we can reduce the number of unknowns by two, making the optimization process more reliable.

(3) There is currently no feedback to the user to indicate when the plane has been scanned from a sufficiently diverse set of positions and angles to produce a well constrained solution to the resulting non-linear equations. The residual error from the optimisation is not an accurate measure of calibration quality, since it is possible to perform an incorrect calibration with a low residual error. An extreme example would be to repeatedly scan the plane from the same position. In this case, the residual error would be low and the user has no idea that the calibration is unreliable. In our experience supporting medical physicists and clinicians using our single wall calibration system (Treece et al., 2003), many inexperienced users have recorded under-constrained motion sequences leading to incorrect calibrations. In this chapter, we show how it is possible to provide feedback to the user on the reliability of the calibration.

(4) We do not know what motions are necessary for fully constrained spatial calibration. As a result, we scan the plane from a wide range of positions and angles to ensure that the calibration parameters are well-constrained. Prager
4.1 Improvement to Single Wall Calibrations

et al. (1998) and Treece et al. (2003) suggested two similar sequences of motions and both are sufficient for spatial calibration. However, we do not know which of the scanning motions are required and which are less important. We also do not know the impact of the inclusion of these redundant motions on constraining the calibration parameters.

In this chapter, we present the modifications we have made to current single wall calibration systems used by Prager et al. (1998) and Treece et al. (2003). We perform and analyze a number of calibrations in order to measure the reliability of a calibration using a metric based on the curvature of the objective function. These calibrations are used to compare with the results from Treece et al. (2003), in order to verify the effectiveness of the modifications we made to the single wall calibration protocol. We go on to simulate the calibration environment in Matlab and analyze each motion analytically.

4.1 Improvement to Single Wall Calibrations

We describe three modifications we have made to improve current plane-based calibration systems. First, we have followed the approach by Trobaugh et al. (1994) and Pagoulatos et al. (2001) to account for the difference between the speed of sound in water and soft tissue. The second is to determine the B-scan scales explicitly. Lastly, we provide feedback to the user on the reliability of the calibration.

4.1.1 Speed of Sound

The speed of sound problem is fixed by allowing the calibration to be performed at room temperature, but requiring the user to provide a measurement of the water temperature. The ultrasound images can then be corrected using one of the well-known polynomial models for the speed of ultrasound in water as a function of temperature (Bilaniuk & Wong, 1993).

Since sound travels slower in cold water than in average soft tissue, the image of the plane (dotted lines) appears further away than its actual position (solid
4.1 Improvement to Single Wall Calibrations

lines), as shown in Figure 4.1. Corrections can therefore be made by moving the image of the plane towards the face of the probe.

We used the same edge detection algorithm used by Prager et al. (1998) (Appendix D.3.1). Vertical samples are taken within the user defined region in the ultrasound image. A likely edge point is where the derivative of the smoothed signal first exceeds some threshold. The edge can then be found by applying the RANSAC line detection algorithm to these likely edge points (Fischler & Bolles, 1981). For a linear probe, the correction is a translation of each likely edge point toward the probe face; the axial coordinates $v$ of the points are multiplied by the temperature correction factor $t = \frac{\text{speed in cold water}}{\text{speed in average soft tissue}}$.

![Figure 4.1: Distortion due to sound speed errors. The dotted lines show the distorted plane due to a slower sound speed in cold water. The solid lines show the corrected plane. The diagram is exaggerated to emphasize the distortion.](image)

For a curvilinear probe, the plane appears as a curve rather than a line. We therefore first detect the probe shape automatically (Treece et al., 2002). The probe centre is then the intersection of the left and right edges. We further assume that the probe face is part of a circular arc with radius $R$ centred at the probe centre. We move each point on the curve (shown as a dotted line in Figure 4.1) towards the probe face, in the direction of the probe centre, by the temperature correction factor. The shifted points will lie on a straight line, and can be detected using the RANSAC line detection algorithm (Fischler & Bolles, 1981). Since this correction is mathematically sophisticated and extra steps need to be followed by the user to detect the probe shape reliably, an approximate
correction is considered for common abdominal curvilinear probes. Please refer to Appendix A.1 for a detailed discussion.

### 4.1.2 Explicit Scale Estimation

During calibration, the calibration parameters are solved using an iterative optimization algorithm. We increase the robustness of the calibration process by reducing the number of parameters involved. We can establish the image scales at the start of the procedure using a separate protocol, and hence remove the horizontal and vertical image scales from the optimisation. The result is that the non-linear optimisation only has to determine nine rather than eleven parameters, and this makes the system of equations easier to constrain.

The image scales are determined using the distance measurement tool that is available on most (if not all) clinical ultrasound machines. This tool provides the distance, in millimetres, between two points marked in the B-scan image. These points will appear in the grabbed B-scan on the computer. A simple graphical user interface has been implemented to enable the user to indicate these points on the captured B-scan and enter their separation. Since our ultrasound machines provide square pixels to a very good approximation, we have further simplified this procedure, with little loss of accuracy, by assuming that the horizontal and vertical image scales are the same. The scale is therefore estimated by

$$\frac{\text{distance (mm)}}{\sqrt{(\Delta u)^2 + (\Delta v)^2}}$$

where $\Delta u$ and $\Delta v$ are the horizontal and vertical distances, in pixels, between the two marked points on the B-scan. If the image scales are different, two distance measurements should be made. The scales can then be solved simultaneously.

### 4.1.3 Feedback on Calibration Reliability

During probe calibration we obtain a set of constraints involving nine variables, since the B-scan scales have been determined explicitly. The calibration parameters are found at the minimum of the objective function (Equation 2.11) by using a non-linear optimisation algorithm, and we can ensure that the global minimum is obtained (Appendix A.2). We wish to measure how well the calibration parameters are constrained. As the nine parameters are similarly scaled,
with distances in centimetres and angles in radians, we can get an indication of an under-constrained solution using the curvature of the objective function at the global minimum. If the calibration parameters are well-constrained, then the curvature of the objective function at the minimum should be high in every direction. If the curvature is low in some direction, then one or more parameters will be under-constrained. We can thus spot poor solutions and advise the user to scan the plane from some more angles to introduce further constraints.

We wish to use the minimum curvature at the global minimum as a measure of the quality of the calibration. We therefore measure the minimum curvature of the calibration by computing the Hessian, the matrix consisting of all the second derivatives of the objective function. The eigenvalues and eigenvectors of the Hessian are evaluated to give the magnitudes and directions of greatest and least curvature. If an optical tracking system is used, it is possible that the eigenvector corresponding to the minimum eigenvalue is found to be in the translation direction that defines the phantom in world space. This eigenvalue is expected to be lower than other eigenvalues, because this translation is typically an order of magnitude larger than the other calibration parameters. This means that the error in this direction is also expected to be an order of magnitude higher than other parameters. This does not bother us as the three parameters defining the phantom in world space are subsequently discarded, and are not used to judge how well a calibration is performed. In this case, we use the second minimum eigenvalue, whose corresponding eigenvector is in a direction relating to a combination of the six parameters that define the transformation from the mobile part of the position sensor to the corner of the B-scan image. For brevity, we will use the word eigenvalue or minimum eigenvalue in this thesis while we are in fact referring to the second minimum eigenvalue (minimum eigenvalue of interest). We note that if a magnetic position sensor is used, its stationary part can be placed near the plane phantom during calibration. Thus, in a good calibration there is no reason for the minimum eigenvalue to be low and the minimum eigenvalue should be used.
4.2 Calibration Experiments

In order to verify the effectiveness of our improvements, we performed a series of calibrations using three probes on two different ultrasound machines. The first is the Diasus (Dynamic Imaging Ltd., U.K.) 5–10MHz linear-array probe. The analog RF ultrasound data, after receive focusing and time-gain compensation but before log-compression and envelope detection, was digitized using a Gage CompuScope 14100 PCI 14-bit analog to digital converter (Gage Applied Technologies Inc., U.S.A.), and transferred at 10 frames per second to a Pentium(R) 4 2.80GHz PC running Microsoft Windows XP. Each frame consists of 127 vectors, each with 3827 data points sampled at 66.67MHz synchronized with the clock of the ultrasound machine. The RF data was converted to an analytic signal using matching Hilbert filters with a unity gain and 5–10MHz pass band. The amplitude of the bandpassed signal was suitably log-compressed so that the resulting images resemble those on the ultrasound machine. Finally the data was linearly interpolated to produce the image displayed on the computer based on the image scale, which we have set at 0.1mm/pixel. We also performed calibrations with the Toshiba (Toshiba Corporation, Japan) model SSA-270A/HG 7.5MHZ linear probe, and the 3.75MHz curvilinear probe. The B-scan images were digitized using a Brookitree (Conexant Systems Inc., U.S.A.) BT878 frame-grabber card and transferred to a 3.0GHz PC running Linux. Both probes were tracked using an AdapTrax (Traxtal Technologies, Canada) infrared LED target for the Polaris (Northern Digital Inc., Canada) optical tracking system. The probes and depth settings, as well as the cropped B-scan sizes in pixels used in the experiments, are listed in Table 4.1. One focus at approximately half the depth setting was used in each calibration. The scales in the Diasus calibrations were set manually to 0.1mm/pixel, since this is the scale at which the RF data, after log-compression, envelope detection and band-pass filtration, were displayed. The distance measurement tool was utilized to estimate the scales in the Toshiba calibrations. The scale estimation was highly repeatable by zooming into the images, so that the same scales were obtained in every measurement. Hence there is no variation in the scales in each calibration.
4.2 Calibration Experiments

Table 4.1: Probes and other settings used in the calibrations.

<table>
<thead>
<tr>
<th>Machine</th>
<th>Probe Type</th>
<th>Frequency</th>
<th>Depth</th>
<th>B-scan</th>
<th>Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diasus</td>
<td>Linear</td>
<td>5–10MHz</td>
<td>3cm</td>
<td>360 × 300</td>
<td>Preset</td>
</tr>
<tr>
<td>Diasus</td>
<td>Linear</td>
<td>5–10MHz</td>
<td>6cm</td>
<td>360 × 600</td>
<td>Preset</td>
</tr>
<tr>
<td>Toshiba</td>
<td>Linear</td>
<td>7.5MHz</td>
<td>6cm</td>
<td>222 × 416</td>
<td>Measurement</td>
</tr>
<tr>
<td>Toshiba</td>
<td>Curvilinear</td>
<td>3.75MHz</td>
<td>6cm</td>
<td>388 × 380</td>
<td>Measurement</td>
</tr>
</tbody>
</table>

To verify the effectiveness of our eigenvalue metric, we need to perform a series of valid calibrations that have different qualities of calibration. We anticipate that as we capture more images of the plane, the calibration parameters will be more constrained leading to calibrations with better qualities. We have performed four calibrations by capturing 30, 60, 90 and 120 images of the plane using the Cambridge phantom. The calibration consisting of 60 images was formed by capturing 30 images of the plane in addition to the 30 images from the previous calibration, and similarly for the calibrations with 90 and 120 frames. We then computed the eigenvalues and eigenvectors for each of our calibrations. Every calibration was repeated ten times. In all, a total of: 4 probes × 4 numbers of plane images × 10 repetitions = 160 calibrations were acquired.

It is important for plane-based calibration that the plane is scanned in a non-degenerate sequence to constrain the plane in space ([Prager et al., 1998; Treece et al., 2003]). We have ensured our calibrations all included the required motion sequences, namely: vertical movements, side to side rotations, front to back rotations, translation in the $x$ direction, translation in the $y$ direction and rotations about the $z$-axis, where $x$ and $y$ are the two axes in the plane phantom, aligned with the sides of the water tank, and $z$ is the corresponding axis perpendicularly out of the plane, as shown in Figure 4.2.

The quality of each calibration was assessed by measuring its precision and accuracy. Precision is measured by the consistency of the locations of the four corners and the middle of the image when located in 3D space using the calibration parameters, i.e. calibration reproducibility defined in Section 3.1.1. These five errors were then averaged.
4.2 Calibration Experiments

Figure 4.2: The sequence of probe movement required during calibration to constrain the 11 calibration parameters (before scales estimation). The diagram is taken from Treece et al. (2003).
4.2 Calibration Experiments

The precision measurements indicate the variation in the image position due to spatial calibration errors. The repeatability of the calibrations is not an indication of accuracy, as the calibrations could be highly repeatable but biased. We assess the accuracy of our calibrations by scanning a tissue mimicking phantom consisting of coplanar spheres with 2mm diameter (Kofler Jr. & Madsen, 2001). These spheres lie on a grid with known dimensions (Appendix A.3), and they show up as dark circles in the B-scans. We scanned the phantom with the Diasus probe at 3cm and 6cm depth, using the three patterns in Figure 4.3. These patterns are used to assess system accuracy and are not necessary during clinical use of the freehand 3D ultrasound system. In order to maximize the visibility of the spheres, six focal points were used in the middle region of the image. The images were recorded at approximately 7–15 frames per second. Although many foci were used during image acquisition, some of the spheres away from the focal points could not be seen clearly and segmented properly, so they were not included in the accuracy assessment. The phantom was scanned freehand, but as slowly as possible so that each sphere appears in a number of consecutive images. A typical recording consisted of approximately 400 scans down the 10cm length of the scanning window. The images of the spheres were then semi-automatically segmented, reconstructed in space using the parameters from our calibrations. The centres of the reconstructed spheres are fitted to the predefined grid using least squares minimization (Treece et al., 2003). We then measure the error as the 3D distances between the centres of the reconstructed spheres and the corresponding points on the grid. This is distance measurement accuracy detailed in Section 3.1.2.2. Since the alignment of the ultrasonic data and the grid is optimised, any offset between the reconstructed spheres and the grid would remain undetected. Nevertheless, it does measure the accuracy of locating features relative to each other within the data set.

4.2.1 Calibration Comparison

We verify the effectiveness of the improvements we have made to the single wall calibration method by comparing our calibrations based on 120 images with the
4.2 Calibration Experiments

(a) Translation  (b) Horizontal Rotation  (c) Vertical Rotation

Figure 4.3: Scanning patterns used when scanning the phantom. Approximately 400 scans were recorded in each case.

results from Treece et al. (2003), where the highest definition system using plane-based calibrations was reported. Figure 4.4 shows that the calibrations in this thesis are more precise than those achieved by Treece et al. (2003). The error for the Diasus 3cm probe has been reduced by more than a half, and the error for the 6cm depth setting has been reduced by more than a third.

Figure 4.4: Comparison of the errors due to spatial calibration alone achieved in this thesis and in Treece et al. (2003). The errors shown are the mean 3D errors at the four corners and the centre of the B-scan.

The probe motion during calibration is restricted when the depth is set low at 3cm. This can result in a weakly constrained optimisation. By removing the scale factors from this optimisation we reduce the dimensionality of the search and greatly improve the precision of the result.

The higher errors in the calibrations performed by Treece et al. (2003) are most likely caused by the variation in temperature during calibration. The hot 48 Celsius water will cool down rapidly in a standard room environment. By
correcting for the temperature difference, calibration in cold water is possible.

Figure 4.5 shows a comparison of our system accuracy with the results from Treece et al. (2003). It shows that our adjustments have improved the system accuracy slightly.

The errors shown are the mean 3D distance errors between the centres of the segmented spheres and their corresponding points on the regular grid.

4.2.2 Curvature of Valid Calibrations

We have performed a series of calibrations with an increasing number of scans. Figure 4.6 shows the precision as the number of scans is increased in our calibrations. The decreasing graphs show that more consistent calibrations were achieved at a higher number of scans. We have thus obtained four calibrations with different calibration quality for each probe. The linear behaviour of the Diasus probe at 6cm depth may appear peculiar at first glance, but in fact it saturates between 120 and 150 scans. This was confirmed by an additional experiment and further data analysis. This linear behaviour is also evident in other probes, most noticeable in the Diasus probe at 3cm depth before it saturates after 90 scans.

Precision on its own is not a sufficient measure for calibration quality as the calibrations can be repeatable but biased. Figure 4.7 shows the orthographic projection of the sphere centres’ distribution used in accuracy assessment. The
4.2 Calibration Experiments

Figure 4.6: The variation of the four corners and the centre of the cropped B-scan due to spatial calibration for the different types of probes.

Small grey circles are the actual positions and size of the spheres forming the predefined grid. In order to make the errors more visible, the small grey circles are magnified by a factor of four and shown as larger black circles. The crosses show the centre of each reconstructed sphere with the error magnified by a factor of four. The diagrams in the left column show the reconstruction errors for the Diasus probe at 3cm depth; the errors at 6cm depth are shown in the right column.

Figure 4.8 shows the average accuracy as a function of the number of scans. The graphs are decreasing, but relatively flat. This is most likely because the dominant errors are not caused by spatial calibration. Possible sources include the optical tracker, temporal misalignments, image capture and segmentation of the spheres.

Figures 4.6 and 4.8 verify that our calibrations are more precise and accurate as more images of the plane are captured. Nevertheless, we will use precision as the measure for calibration quality. This is because precision is more sensitive to calibration errors than accuracy, since precision measures the variations caused by spatial calibration only, and not a combination of calibration and other sources of errors. We have also verified that this choice is viable by the high accuracy in each calibration, since the errors are found to be less than 0.35mm, the accuracy
4.2 Calibration Experiments

Figure 4.7: Orthographic projection of the phantom and the reconstructed spheres. The small grey circles show the true size and location of the spheres in space. The larger black circles show these spheres magnified by a factor of four. The distributions of the reconstructed spheres are shown as crosses, with the error magnified by a factor of four. The consistent magnification means that the crosses and the large black circles are directly comparable.
of our tracking system. The even distribution of located spheres in Figure 4.7 further confirms that the calibrations are unbiased and accurate.

We now compute the eigenvalues and eigenvectors of each calibration as described before. Figure 4.9 shows the graph of the second minimum eigenvalue (minimum eigenvalue of interest) of the Hessian at the calibration solution as the number of scans is increased. A clear linear relationship can be seen (correlation coefficient $r > 0.99$). This shows that the minimum curvature at the solution is related to the number of scans of the plane used in the calibration. Please refer to Appendix A.4 for a mathematical analysis on the linearity of the graphs.

We have shown in Figure 4.6 that higher calibration precision is obtained as the number of images of the plane is increased. We have further verified in Figure 4.9 that calibration quality is reflected in our eigenvalue metric. This enables us to plot precision against eigenvalue, as shown in Figure 4.10. This shows that we can use the size of the eigenvalue as an indicator for the quality of the solution. A minimum eigenvalue above 3 indicates a reliable calibration with the Diasus probe at 6cm depth and a minimum eigenvalue above 1 indicates a reliable calibration with the same probe at 3cm depth setting. Although the threshold may vary with probe and depth setting, a high eigenvalue will ensure a well-constrained calibration. In our studies, an eigenvalue above 3 indicates a reliable calibration for the three probes. We can therefore advise the user that calibrations with our probes are reliable when an eigenvalue above 3 is observed.
4.2 Calibration Experiments

Figure 4.9: The graphs show the second minimum eigenvalue (minimum eigenvalue of interest) of the Hessian calculated at the solution of the objective function as the number of scans is increased.

A similar analysis can be performed with other probes to determine the reliability of the associated probe calibrations.

Figure 4.10: The variation of the four corners and the centre of the B-scan as the calibration becomes more constrained.
4.2.3 Incomplete Scanning Patterns

We have shown in the previous section that we can use the eigenvalue to determine whether the solution is well-constrained, and hence the reliability of the calibration. This assumes that the user has performed all the required scanning patterns in Figure 4.2. Inexperienced users often neglect some of these motions inadvertently. We performed calibrations consisting of 30 images with one of the motions missing with the Diasus probe at 6cm depth to investigate whether incomplete calibrations do result in an under-constrained solution, reflected by a low eigenvalue.

The minimum eigenvalues for these incomplete calibrations are shown in Figure 4.11. In the two cases where side to side and front to back rotations are missing, the under-constrained optimisation resulting from incomplete scanning patterns is reflected by a near zero minimum eigenvalue. For the case where vertical movements are missing, the minimum eigenvalue also dropped by more than a factor of two, indicating an insufficiently constrained solution. In the other three cases the eigenvalue did not drop as expected. This suggests that although these motions are absent, the solution is well-constrained. We verify this by performing a series of calibrations without these motions and show that they are just as precise as calibrations with all motions exercised.

![Figure 4.11: The minimum eigenvalues for the calibrations where a particular motion is missing.](image)
4.2 Calibration Experiments

We performed an additional 10 calibrations with minimal rotations and translations in the aforementioned directions, but still containing the remaining three types of motions. Figure 4.12 shows the graph of average error against the number of scans. The graph for good calibrations where all motions are exercised using the same probe (Figure 4.6) is duplicated here for comparison.

![Graph showing comparison between calibrations.](image)

Figure 4.12: Comparison between calibrations with minimal rotation about the z-axis and minimal translation in the xy-plane, and calibrations where all six types of motion are exercised.

From the figure it can be seen that the apparent paradoxical high eigenvalue for calibrations where rotations about the z-axis and translations in the xy-plane are missing is in fact correct. The average errors differ by less than 0.1mm from a low to a high number of scans. This means that these motions are not necessary to constrain the optimisation sufficiently for accurate calibration. This is because we have reduced the number of parameters to be optimized from eleven to nine, hence reducing the degrees of freedom which the calibration needs to constrain. This observation can be confirmed by simulating the calibration process mathematically. We have found that successful calibrations can be simulated excluding the aforementioned motions.
4.3 Calibration Protocol Simulation

We have devised a metric that will detect whether the system of equations is under-constrained and found certain scanning motions to be unnecessary. However, the exact scanning patterns that are required to constrain the calibration parameters is still unknown. We therefore simulate the calibration procedure in Matlab. In the simulation, the plane is scanned from different positions and we verify whether these scans are sufficient to constrain the calibration parameters. Since only the position data is simulated, excluding the speckle in B-scan images or position sensor errors, the simulation results are only limited by the precision of Matlab. A pseudo-code of the simulation is given in Appendix D. We will give a theoretical configuration that is sufficient to constrain the calibration parameters in this section and discuss the effect of capturing redundant images during calibration.

Let us recall that for plane-based calibrations, the user scans the plane from different positions. A point \( p^I \) on the image of the plane in the B-scan can be mapped to the phantom’s coordinate system by Equation 1.1. Assuming the scales are known, the equation becomes

\[
p^F = T_{F\rightarrow W} T_{W\rightarrow S} T_{S\rightarrow I} p^I. \tag{4.1}
\]

Since the phantom’s coordinate system is placed so that the \( xy \)-plane coincides with the plane phantom, the \( z \)-coordinate of \( p^F \) needs to be zero, i.e.

\[
p^F_z = 0. \tag{4.2}
\]

This places a constraint on the nine calibration parameters—six parameters from \( T_{S\rightarrow I} \) and two rotations and a translation from \( T_{F\rightarrow W} \).

In order to find what is a sufficient sequence of images that needs to be acquired, we simulate the calibration process hypothetically in Matlab. Let us assume that \( T_{S\rightarrow I} \) and \( T_{F\rightarrow W} \) have rotations \( a, b, c \) and \( \alpha, \beta, \gamma \) respectively and translations \( x_s, y_s, z_s \) and \( x_w, y_w, z_w \) respectively. We note that the calibration parameters are \( a, b, c, x_s, y_s, z_s, \beta, \gamma, z_w \), but the others are required to set up the environment. 

64
4.3 Calibration Protocol Simulation

In our scenario, we will assume that the mobile part of the position sensor is placed to coincide with the B-scan and that the stationary counterpart of the position sensor is placed to coincide with the plane phantom, i.e. \( a = b = c = x_s = y_s = z_s = \alpha = \beta = \gamma = x_w = y_w = z_w = 0 \). This assumption is only necessary so that we can obtain the reading \( T_{W\rightarrow S} \). Note that this calibration setting is not possible in practice.

We now start to scan the plane by placing the probe at hypothetical locations. For each image, two points are used to set up two constraints. Without loss of generality, we will assume the two points are at \( p_x^I = 0 \) and \( p_x^I = 1 \).

We first scan the plane with the scan plane held perpendicularly to the plane phantom, so that the probe and phantom axes are parallel to each other with the \( z \)-axes in opposite directions as shown in Figure 4.13. The origin of the probe is at \((1, 1, 1)^t\) and its \( x \) and \( y \)-axes in the directions \((2, 1, 1)^t\) and \((1, 1, 0)^t\). The probe is then translated to the right by 1 unit to \((2, 1, 1)^t\).

We now rotate the probe about its origin in the plane by 45° to the left and right, and then rotate the probe about the lateral axis through the origin by 45° to the front. We note in passing that given a front rotation has been performed, a back rotation is not necessary.

Lastly, we translate the front and back rotations by 1 unit away from the plane in the direction of the rotated axial axis.

Following the above protocol, the plane has been scanned 7 times. We computed the 14 constraints \( C_i(x) : \mathbb{R}^9 \rightarrow \mathbb{R} \) analytically in Matlab. The function \( f(x) = \sum_{i=1}^{14} C_i(x)^2 \) is differentiated analytically with respect to the nine calibration parameters. The Hessian is formed analytically from the corresponding second derivatives. The Hessian is then evaluated at the solution, i.e. by substituting \( a = b = c = x_s = y_s = z_s = \beta = \gamma = z_w = 0 \). Now, all of the eigenvalues of the Hessian are found to be positive, the Hessian is therefore positive definite and the set of constraints is sufficient to constrain the calibration parameters. If we try to constrain the parameters with any six of the seven scans, the minimum eigenvalue of the Hessian becomes zero. This shows that this set of scans is minimal in the sense that any six of the seven scans are not sufficient to constrain the calibration parameters.
4.3 Calibration Protocol Simulation

There are 14 constraints resulting from the 7 scans. Since there are only 9 calibration parameters, not all of these constraints are active. It is possible to remove 5 of the 14 constraints so that the calibration parameters remain constrained. The required constraints are shown as a circle in Figure 4.13.

Figure 4.13: Minimum set of scanning patterns required for a plane-based calibration. The dark circles show the active constraints in the optimization process.

We repeat the above simulation using values in a typical calibration setting. The mobile part of the position sensor is placed at 1cm offsets in each direction.
of the three principal axes and at an inclination of 1 radian about each axis. The stationary part of the position sensor is placed about 150cm to the left, 10cm to the back and 100cm above the plane phantom, with an elevational inclination of 30°, and approximately 10° rotation about the other two axes. Specifically, $a = 1$, $b = 1$, $c = 1$, $x_s = 1$, $y_s = 1$, $z_s = 1$ and $\alpha = -\frac{9\pi}{20}$, $\beta = \frac{19\pi}{20}$, $\gamma = \frac{\pi}{6}$, $x_w = -150$, $y_w = 10$, $z_w = 100$. The probe is scanned at an initial position $(10, 10, 10)^t$, i.e. 10cm above the plane phantom. The same protocol is followed and we confirm that the same protocol is sufficient to constrain the calibration parameters. Plane-based calibrations should therefore comprise repetitions of the above protocol. Repetition of valid sets of constraints is necessary to minimize noise and will be reflected in our eigenvalue metric.

4.3.1 Redundant Images

The next question that comes to mind is whether we need to remove the five inactive constraints from our seven scans. We would like to find out whether this extra information will bias our solution. This problem can be generalized to a typical scenario when an inexperienced user is trying to perform a plane-based calibration. The user may perform a valid and well-constrained calibration, but at the same time capture a series of redundant images, and we would like to know whether the result will be equally valid as the result where no redundant images were acquired. Note that repeating a set of valid scans further confirms the validity of these scans, and therefore adds information to the optimisation process. The repeated scans are not considered redundant.

Let us suppose the user has recorded $N$ constraints in total, none of which are redundant. We form the function $f_1(x) = \sum_{i=1}^{N} C_i(x)^2$, where $C_i$ are the constraints. Now, if a series of $M$ redundant constraints $R_i(x)$ are acquired, the corresponding function will be $f_2(x) = f_1(x) + \sum_{i=1}^{M} R_i(x)^2$. It is important to note that since the redundant images do not add extra information to the existing constraints, the eigenvalues of the two Hessians should be identical. Hence the eigenvalue metric on its own will not detect redundant constraints.

Now, in an ideal situation where the images of the plane can be segmented without error and every component of the equipment is noise-free, the minimum
of $f_1$ and $f_2$ will be attained at the correct calibration. Suppose the correct calibration is $x_c \in \mathbb{R}^9$, then $C_i(x_c) = 0$ and $R_i(x_c) = 0 \Rightarrow f_1(x_c) = 0$ and $f_2(x_c) = 0$. Thus the inclusion of these redundant constraints would not affect the solution.

In practice, however, $C_i(x_c)$ and $R_i(x_c)$ are seldom zero. This is largely due to two reasons. The first is due to the noise of the system and second due to some systematic error. Now, the calibration parameters are solved by minimizing the functions $f_1$ and $f_2$. If $f_1$ attains its minimum at $x_c + a$, then $x_c + a$ would be believed as the correct calibration, with an error of $a$. This is the accuracy of plane-based calibrations.

Suppose that, for some $i$, the error of $R_i$ is due to unbiased noise in the system, such as position sensor noise, then the minimum of $R_i(x)^2$ is $x_c + \epsilon_{ni}$, where $\epsilon_{ni} \in \mathbb{R}^9$ is non-zero and has an expected value of zero. This means that the expected minimum of $R_i(x)^2$ is the correct calibration $x_c$. Note that $R_i(x)^2$ is a function of nine variables and has multiple minima. Nevertheless, since we assumed we have sufficiently constrained the calibration parameters, we can disregard the other minima. Now, the minimum of $f_1(x)$ occurs at $x_c + a$ and the expected minimum of $R_i(x)^2$ is at $x_c$. Hence the expected minimum of $f_1(x) + R_i(x)^2$ will be closer to $x_c$ than $x_c + a$. Although the inclusion of $R_i$ is not guaranteed to improve the calibration, we do expect it to improve calibration by including a multiple of such scans to remove noise. In fact, this is why it is necessary to repeat the aforementioned calibration protocol several times. Such repetition is not redundant and is reflected in our eigenvalue metric.

Now, if the error of $R_i$ is due to some systematic error, then the minimum of $R_i(x)^2$ will be at $x_c + \epsilon_{si}$, where $\epsilon_{si}$ is non-zero. This systematic error may be segmentation error, beam-width error causing the plane to appear at the wrong depth, or human errors such as mounting the probe incorrectly on the probe holder for a Cambridge phantom. An example of such a scenario is when a user performs a calibration with a different number of oblique scans, say more front rotations than back rotations. Note that a well-constrained calibration is obtained if all other motions are exercised. The extra front rotations give rise to redundant constraints of this type. The effect of inclusion of such constraints
is that the minimum $x_c + a$ is shifted towards $x_c + \epsilon_{si}$, probably worsening the accuracy.

In order to minimize the above systematically biased constraints, we have to rely on the user’s expertise to not include any of them. Segmentation algorithms are automatic and reliable. The user needs to ensure that the correct line has been detected, this is not difficult. The user is also required to mount the probe accurately for a Cambridge phantom and scan at equal oblique angles on both sides of the normal. These are standard requirements of a plane-based calibration.

We note that the minimal sequence of motions is not symmetrical, since it contains a front rotation without the corresponding back rotation. This would cause a bias if a roughened plane is used. The correct sequence should therefore include the corresponding back rotation as well. It is also not necessary to remove the five inactive constraints, since these constraints are not biased.

4.4 Conclusion

We have proposed some changes to the plane-based 3D ultrasound calibration protocol that increase its reliability and ease of use, while maintaining its state-of-the-art performance in terms of speed and accuracy. The compensation for the temperature of the water and the simple technique for explicit entry of scale values are both easy to use. In particular, fewer degrees of freedom need to be constrained by spatial calibration, and so rotations about the $z$-axis and translations in the $xy$-plane are no longer needed to achieve a good calibration.

We have also produced a foolproof indicator of the calibration quality by calculating the eigenvalues of the Hessian at the solution of the objective function. If a motion necessary for calibration is missing, which is a common scenario for inexperienced users, the incomplete calibration is indicated by a low eigenvalue. We have also found that certain motions that were previously thought to be necessary for probe calibration, are in fact redundant. The eigenvalue criterion has been evaluated and shown to be effective on three different probes and at two different depth settings.

We have identified the critical motions that are required to constrain the calibration parameters sufficiently. An accurate calibration should therefore consist
of repetitions of the identified motion sequence. In order to perform a reliable plane-based calibration, the probe should be placed at the upright position perpendicular to the plane, and then rotated about the lateral and the elevational axes.
Chapter 5

Rapid Calibration

Segmentation of isolated points in ultrasound images is traditionally a slow process. The lack of a reliable automatic algorithm requires users to segment the images manually. We show in this chapter how we can segment wires fully automatically, allowing a very rapid calibration. Note that Chen et al. (2006) provided automatic segmentation and rapid calibration with a simplified Z-phantom by restricting probe movements shortly after the publication of the work presented in this chapter in September 2006 (Hsu et al., 2006b).

When calibrating with a 2D alignment phantom, the user aligns the scan plane with the phantom, which is placed in a known location in 3D space. Specific points on the phantom are marked in the B-scans. Since these points are known in the sensor’s coordinate system, calibration is found by finding the transformation that best transforms the points in the B-scans to their corresponding points in the sensor’s coordinate system. The details are given in Section 2.6. This type of phantom potentially allows for rapid calibration as calibration only requires one image of the phantom.

A major disadvantage of using this phantom is that aligning the scan plane with a 2D phantom is very difficult. The Z-fiducial phantom was designed to solve the alignment problem. The intersection of the ‘Z’ shaped wires with the scan plane forms a virtual 2D phantom that can be used to define the position of the scan plane in space. The main drawback of using this type of phantom is the difficulty associated with segmenting isolated points in each B-scan. Furthermore, each wire does not appear as a single dot or disc in the B-scans, but rather as a
5.1 Spatial Calibration

5.1.1 The Calibration Phantom and the Principles Behind its Use

Figure 5.1(a) shows the calibration phantom. The phantom consists of two parallel polyacetal blocks 5cm apart. We have drilled a number of holes with 0.5mm diameter in these blocks in a predefined configuration. These holes serve as the locations for the ‘Z’ shaped fiducials, which are placed in 5 rows with 4 fiducials in each row, except the first row which has 3 fiducials, as shown in Figure 5.1(b). A 1mm thick translucent 40° Shore A silicone rubber membrane is clamped under tension on top of the phantom. There are 3 cone-shaped divots in the two blocks to fit a 3mm ball-pointed 3D localizer. These divots serve as the principal axes of the phantom and are used to determine the phantom position in 3D space.
5.1 Spatial Calibration

All dimensions are precision manufactured by our workshop with a tolerance of ±0.1mm. We then interwove an ordinary off-the-shelf monofilament fishing line with a diameter of approximately 0.5mm through the holes in the two blocks to form the ‘Z’ shaped fiducials. The fishing line has a breaking strain of 25lb, this allows us to fasten the wire tightly onto the phantom so that it remains taut under water during probe calibration.

![Isometric View of the Phantom](image1)

(a) Isometric View of the Phantom

![Fiducial Positions](image2)

(b) Fiducial Positions

Figure 5.1: The phantom consists of two rigid plastic plates that are held 75.6mm apart by bolts at the four corners. A rubber membrane is clamped across the top and wires are threaded through holes in the plates to form the required pattern of Z shapes. The figure shows a 3D sketch of the positions of the wires and the membrane.

In order to perform calibrations in real-time, we need to track and segment the wires fully automatically. This is achieved by the introduction of the membrane, which is treated as a planar phantom, and is segmented automatically (Prager et al., 1998). The positions of the wires are at known depths beneath the membrane by construction. If the image scales are known, the search region for the wires can be limited to five narrow strips, one at each depth, as shown in Figure 5.2. This allows reliable and automatic wire segmentation.

In general, spatial calibration involves finding the horizontal and vertical scales in the B-scan images and the 6 parameters that define the phantom in space in
5.1 Spatial Calibration

Figure 5.2: Wire segmentation of a typical B-scan image. The search regions are limited to five narrow strips at known depths beneath the rubber membrane.

addition to the 6 parameters that define the B-scan relative to the sensor’s position. The values of the first 8 variables can be found using a separate protocol prior to phantom scanning. We first compute the scale values using the distance measurement tool provided by the ultrasound machine (Section 4.1.2). A calibrated pointer is then used to locate the phantom in space by its divots. This leaves only the 6 parameters that define the rigid body transformation from the scan plane to the position sensor to be found by calibration.

In our phantom design, the wires are interwoven through the holes $H_1$, $H_2$, $H_3$, $H_4$, $H_5$ and $H_6$ in a ‘Z’ shape, as shown in Figure 5.3. These holes are precision manufactured relative to the divots and are therefore known in space. The 3D positions of $A$, $B$, $C$ and $D$ can be calculated using trivial coordinate geometry. When the probe is placed over the wire configuration, the scan plane intersects the wires at $M$, $Z$ and $N$. These points can be located in the B-scan image. Since the image scales have been determined, the distances $|MZ|$ and $|MN|$ can be measured off the B-scan image. The location of $Z$ can thus be computed: $Z = B + \frac{|HZ|}{|BC|}(C - B) = B + \frac{|MZ|}{|MN|}(C - B)$, since $\triangle BMZ$ and $\triangle CNZ$ are similar. The manufacturing tolerance of $H_i$ is 0.1mm, the tolerance of $B$ is approximately 0.1mm as well, and the maximum error of $|C - B|$ is thus 0.2mm. Since $\frac{|MZ|}{|MN|} < 1$, the tolerance of $Z$ is less than 0.3mm.
5.1 Spatial Calibration

For each ‘Z’ shape wire configuration, the Z-fiducial $p_i$ (labelled $Z$ in Figure 5.3) is located in space using the aforementioned equation. This can be transferred to the sensor’s coordinate system using the inverse of the position sensor’s readings: $T_{W^{-1}}^{-1}S_i p_i^W$. By segmenting each Z-fiducial in the B-scans, we can locate the fiducials $p_i^I$ in the B-scan’s and the sensor’s coordinate systems.

Assuming we have located at least 3 non-collinear fiducials, we can compute the corresponding calibration $T_{S^{-1}}^{-1}$ that best fits the two data sets (Arun et al., 1987), by minimizing $|T_{W^{-1}}^{-1}S_i p_i^W - T_{S^{-1}}^{-1}p_i^I|$.

5.1.2 Image Segmentation and Processing

Since the wires are at known depths beneath the rubber membrane, we can limit the search region to five narrow strips. Furthermore, since the wires are mounted parallel to the rubber membrane, the segmentation algorithm remains robust even if the probe is at an inclination to the phantom. We have used a fairly simple algorithm to segment the smeared blobs. Since the smeared image of the wire appears to be approximately 1–3mm in diameter, the average intensities of each 1mm $\times$ 1mm block in the search region are computed. A user defined threshold is placed on the average intensities. This divides the search region into a number of connected regions. Potentially, each region represents a smeared image of the wire. The maximum average intensity of each connected region is located. The centroid of a small rectangular region (approximately 0.5mm $\times$ 0.25mm) centred at this location is computed and stored as a potential wire location. Since we have
limited the search region to a narrow strip, our simple segmentation algorithm is robust over a wide range of thresholds, which can be manually adjusted until the maximum number of wires is reliably segmented. The secondary echo from the membrane together with the generally poor ultrasound image quality gives rise to a few falsely detected potential wire locations. If we fit the detected wires to the known geometry of the wire configuration, we can easily reject any falsely detected wires that lie outside the expected region (Appendix B.1). An overview of the segmentation algorithm is shown in Appendix D.3.2.

It is not clear whether the segmented echo arises from the edge or the centre of the wire. However, since the spatial resolution of curvilinear probes is typically about 1mm and the wire diameter is only 0.5mm, the transducers will not be able to distinguish the difference between the two echoes.

Now, because sound travels at different speeds in soft tissue, rubber (Folds, 1974) and water (Bilaniuk & Wong, 1993), we measure the water temperature, compute the speed of sound in water and rubber at this temperature and post-process the segmented wires in the B-scan images by shifting the wires towards the probe by the factor \( \frac{\text{computed speed of sound}}{1540} \). This removes any error due to sound speed differences between water and soft tissue, furthermore it enables us to perform calibrations in water at any temperature. The speed of sound is approximately 920m/s in silicone rubber at 48 Celsius (Folds, 1974). When a curved probe is used, the ultrasonic beam would pass through the rubber membrane at an angle. For conventional curvilinear probes, where ultrasound waves travel at a maximum angle of approximately 30° to the normal, the uncorrected error is approximately \( \tan 30° \left( 1 - \frac{920}{1540} \right) \cdot 1\text{mm} \approx 0.2\text{mm} \) in the lateral direction.

### 5.1.3 Precision and Accuracy

We need to examine the quality of the calibration that can be produced by using this phantom. We measure the quality by assessing its precision and accuracy. Precision is measured by variation of the four corners and centre of the B-scan in the sensor’s coordinate system, i.e. calibration reproducibility as defined in Section 3.1.1. The mean 2D error is also computed. This is the error in the scan plane alone, neglecting the error in the elevational direction.
5.1 Spatial Calibration

Precision measures the reproducibility of the calibration. This does not directly measure the accuracy of the calibration, since the calibrations can be reproducible but biased. One technique for assessing the accuracy of the system is to scan an object, reconstruct its image in 3D space and compare distances or volume measurements from the reconstructed image with the known dimensions (Lindseth et al., 2003b; Rousseau et al., 2006). We use two metrics to assess the accuracy of our phantom—internal accuracy and point reconstruction accuracy. These accuracy measures are chosen to compare the results with the literature (Lindseth et al., 2003b; Pagoulatos et al., 2001). The metrics are different to the distance measurement used in Chapter 4, where the purpose is to compare with another plane-based system (Treece et al., 2003). Internal accuracy measures the consistency with which we can locate fiducial points on the same Z-phantom that was used in the calibration. We independently scanned our phantom, segmented the wires automatically, and compared the 3D locations of the mapped fiducials $p_i^W$ with the locations computed from the phantom construction $p_i'^W$. We measure point reconstruction accuracy by scanning a point phantom and reconstructing its image $p_i^W$ in 3D space. The location $p_i^W$ of this point phantom is also obtained by using an external source—an independent pointer. Point reconstruction accuracy serves as a verification process since a different phantom is used, and it measures the accuracy of the whole 3D ultrasound system. In both cases, accuracy is the mean error of the located fiducials: $\frac{1}{N} \sum |p_i^W - p_i'^W|$, where $N$ is the number of fiducials located.

During point reconstruction accuracy assessment, we scan the flat end of a 1.5mm thick metal wire from 10 different positions by manually aligning the B-scan with the tip of the wire near the centre of the B-scan. The 3D location of the wire tip is found by using a calibrated pointer that is not used during spatial calibration. Since the error in the location reported by the pointer can be as large as 1mm, we capture 10 readings of the wire tip with the pointer. We segment the wire in each image manually, reconstruct its image in 3D space and compare these locations with the mean of the positions determined by the independent pointer. Since the wire has been scanned from one angle near the centre of the B-scan, PRA near the corners of the image may be worse. This sub-optimal accuracy assessment was performed when the author’s knowledge
Spatial calibration was performed by submerging the phantom in water at 48 Celsius to match the speed of sound in soft tissue. Nevertheless, the water temperature is measured and any discrepancy can be adjusted by post-processing the images as described in the previous section. The water was initially heated to 55 Celsius and allowed to cool during the experiment. A thermometer was used to monitor the water temperature continuously and the water temperature did not drop below 43 Celsius. Water cooling introduces a maximum error of \( \frac{7}{1540} \cdot 8\text{cm} \approx 0.35\text{mm} \) when scanning at 8cm depth. Since our phantom is only 55mm high, the error is generally less than 0.35mm. The phantom was submerged into the water-bath with care to eliminate any air bubbles clinging onto the phantom. A calibrated pointer was then used to localize the phantom at its divots. Since the phantom remained stationary in the water-bath throughout the duration of the experiment, the phantom was not re-localized with the pointer.

In order to demonstrate the robustness of our phantom to poor image qualities, we chose to calibrate our two-generation old Toshiba (Toshiba Corporation, Japan) model SSA-270A/HG 3.75MHz curvilinear probe at a depth of 8cm and 15cm. The probe has served many years in a clinical environment, followed by much abuse in our laboratory, has some dead crystals and produces very poor images. One focus was used at the middle of the image in order to produce B-scans as fast as possible at 18Hz. The B-scan images were digitized using a Brooktree (Conexant Systems Inc., U.S.A.) BT878 frame-grabber card and transferred to a 3.0GHz PC running Linux. An area of 494 \times 420 \text{ pixels} of the B-scan images was cropped for calibration. The probe was tracked using an AdapTrax (Traxtal Technologies, U.S.A.) infrared LED target for the Polaris (Northern Digital Inc., Canada) optical tracking system.

In order to assess the calibration quality achievable with this phantom, we scan the phantom 100 times. The probe was held perpendicularly over the phantom about 1–2cm above the rubber membrane. The B-scan produced is shown in
Figure 5.2. The probe is removed and reinserted into the water-bath between each scan. The probe position is random during each image acquisition, but we have ensured that the whole phantom is visible within each B-scan. There is thus a small amount of translation and rotation in each axis between each acquisition. During each repetition, one frame is saved with the corresponding sensor information, phantom location and the calibration results for accuracy assessment.

We divided our 100 frames into 10 groups with 10 frames each. One calibration was computed from each group, forming the 10 repetitions that we used to evaluate calibration consistency. Single frame calibrations were computed by taking the first frame from each group. Similarly, $n$ frame calibrations were computed by finding the parameters that best fit all the fiducials segmented from the first $n$ frames of each group.

Despite the overall noisy image, the segmentation algorithm has segmented on average 14 of the 15 fiducials that appears on the B-scan images when scanning at 8cm, and segmented 18 out of the total 19 fiducials when calibrating at 15cm. Any undetected fiducial point is simply excluded while solving for the calibration parameters. The residual RMS error of our calibrations varied between 1.5mm and 2.5mm.

We then performed an additional calibration in water at room temperature at a 15cm depth setting. This calibration was used to verify that our temperature correction routine is correct.

5.3 Results

Figure 5.4 shows the internal accuracy and precision of our calibrations as more frames are included in each calibration. The error for the precision graph is the mean 3D error of the four corners and the centre of the B-scan. The error saturates if more than 5 frames are used to compute the calibration parameters.

The RMS error of the reconstructed point during external accuracy assessment is 2.5mm and 3.0mm from the mean location of the wire tip when using the calibrations at 8cm and 15cm respectively. The standard deviation of the reconstructed points from their centroid is 1.2mm at both depths.
5.4 Discussions

We compare our precision and internal accuracies based on 10 frame calibrations with Pagoulatos et al. (2001), where a similar number of frames and fiducials were used. We can deduce from Figure 5.4 that our accuracy saturates just below 1.5mm, which is better than the 2.8mm error reported by Pagoulatos et al. (2001). Lindseth et al. (2003b) attached a position sensor to their phantom and reported an accuracy of 1.2mm. The precision of our system has a mean 3D error of 0.8mm and 1.5mm when calibrating at 8cm and 15cm respectively. This error is slightly better than the errors reported by Pagoulatos et al. (2001), which ranged mostly between 1mm and 2mm depending on the location in the B-scan. Our precision at the centre of the B-scan is 0.5mm when calibrating at 8cm.
Table 5.1: Variation in the four corners and centre of the B-scan due to calibration imprecision. The dimensions are in millimetres.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Position in the B-scan</th>
<th>Mean Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2D</td>
</tr>
<tr>
<td>8cm</td>
<td>Top Left</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Top Right</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>Bottom Left</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Bottom Right</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>Centre</td>
<td>0.37</td>
</tr>
<tr>
<td>15cm</td>
<td>Top Left</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Top Right</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Bottom Left</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Bottom Right</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Centre</td>
<td>0.53</td>
</tr>
</tbody>
</table>

depth. This also compares favourably with the results of Lindseth et al. (2003b), who reported an error of 0.6mm near the centre of the image. This means that our system can produce calibrations at least as good to those in the literature by including at least 6 frames in each calibration, as suggested in Figure 5.4. At least six frames are necessary to ensure a precise calibration, so that a large number of fiducial points can be used to constrain the calibration parameters. Using more frames also minimizes segmentation and position sensor errors. Since we are able to segment the fiducials automatically, a large number of frames can be acquired and processed in a few seconds.

Our external accuracies are slightly worse than the 2.4mm and 2.0mm obtained by Pagoulatos et al. (2001) at 9cm and 16cm depth, as well as the 1.15mm quoted by Lindseth et al. (2003b). However, care must be taken when comparing these results since the verification process is based on different verification phantoms and ultrasonic systems. A possible source of error is in the verification process itself. Our probe is notorious for producing poor images. This adds
5.4 Discussions

difficulty in aligning the scan plane with the wire and introduces segmentation errors.

Our external accuracy for the calibration in cold water matches the accuracy for calibrations performed in water at 48 Celsius. This verifies that our calibration accuracy is not influenced by the water temperature. In either case, there are problems associated with both approaches. Since the ultrasound beam former assumes that the sound speed is 1540m/s, calibrating in a medium where sound travels at a different speed causes the beam to broaden, resulting in a loss of spatial resolution (Anderson et al., 2000; Dudley et al., 2002). On the other hand, it may be uncomfortable for the user to handle the phantom at 48 Celsius. It is also difficult to maintain the water temperature at 48 Celsius. The phantom may also expand due to an increase in the temperature, but this error is small given that the thermal expansion coefficients of the various materials are less than $2 \times 10^{-6}/\text{K}$.

From Table 5.1 it can be deduced that the 3D error is much larger than the 2D error. This means that the main component of the error is in the elevational direction. The reason is because the fiducial’s position in the elevational direction is estimated from the distance between the wires in the B-scan images, and a slight error of a few pixels during wire segmentation will translate into an error of several millimetres in the elevational direction.

Another main source of error is positioning our phantom with a pointer. This error can be minimized by averaging a larger number of readings and repeating many pointer calibrations, e.g. averaging over 10 readings instead of 1 will reduce the variance by a factor of 10. However, high repetition is impractical due to the time duration required. From our experience, the point positioning accuracy of our pointer is approximately 1mm. This error lies in the Polaris itself and from inaccuracies when calibrating the pointer. Any error introduced at this stage will result in a bias in the calibrations. Other sources of error to a lesser degree, approximately several tenth of a millimetre, include uncertainties in the position sensor, phantom movements and phantom construction errors. Phantom movement during calibration is dependent on the environment and the skill of the user. This error can be removed by attaching another position sensor to the phantom, but this would require two targets to be tracked simultaneously.
Phantom construction error is dependent on the skill of the manufacturer: in our case, the phantom is produced with a tolerance of 0.1mm.

5.5 Improving the Elevational Error

The relatively poor estimation of the calibration in the elevational direction is a characteristic weakness of the Z-fiducial phantom. This is evident, although not emphasized, from the calibrations performed in the literature with the Z-phantom (Pagoulatos et al., 2001). We therefore attempt to improve the estimation of the elevational calibration parameter.

We first capture a series of images of the planar membrane that is mounted on the phantom from an oblique angle, as shown in Figure 5.5. It is important that the probe be tilted at approximately the same oblique angle on both sides when scanning the membrane, otherwise the beam-width effect will cause the membrane to appear at an incorrect depth (Prager et al., 1998). Since the membrane is mounted at a precise location on the phantom, the top of the membrane is known in space.

![Figure 5.5: The motions used to scan the membrane to improve the elevational error.](image)

As we scan the phantom as previously described, we obtain a preliminary calibration where the elevational offset is inaccurate. We will assume that the other 3 rotations and 2 translations are correct and independent of the elevational offset.
This assumption, as we shall see, is true if the elevational error is small. We form a new calibration by replacing the elevational offset with a variable, while holding the other parameters invariant. Now, points on the membrane can be mapped to the sensor’s coordinate system by the transformation with the unknown elevational offset, with each coordinate expressed as a linear function of the unknown elevational offset. This point can be mapped to the world coordinate system by the sensor’s readings. Again, each of its coordinates is a linear function of the unknown elevational offset. This point must lie on the plane, whose equation is known. We have therefore obtained a linear equation with one unknown for each point on the plane (Appendix B.2). Every line image of the plane will give rise to two independent equations, hence \( n \) images of the plane will result in a system of \( 2n \) linear equations with one unknown, and the elevational offset is then solved by least square minimization.

We captured 16 images of the plane and replaced the elevational offset with the value calculated using the above method. Figure 5.6 shows that both 3D precision and accuracy are improved slightly, with the error saturating after 8 images. The precision and accuracy of the calibrations performed at 15cm are improved to just below 1.2mm, and the accuracy of the calibrations performed at 8cm is improved to approximately 1mm. The relatively constant precision of the calibrations performed at 8cm depth suggests that the error has saturated and our correction did not worsen this precision.

We now compare our precision with previous calibrations using other types of phantoms with the same ultrasound machine. Our precision is approximately 0.3mm less accurate than the plane-based technique using a Cambridge phantom (Chapter 4), and 0.5mm less accurate than the mechanical device designed by Gee et al. (2005). These comparison are given in Table 5.2. Recall that our precision and accuracy matches the results by Pagoulatos et al. (2001) and Lindseth et al. (2003b). Their results are given in the same table.

The results indicate that we are able to produce calibrations in real-time, while maintaining the accuracy achievable with Z-fiducial phantoms. Calibrating in real-time potentially allows the clinician to change the probe, crop, depth and zoom settings and re-calibration will only take a few seconds. The accuracy
5.5 Improving the Elevational Error

Figure 5.6: Precision and accuracy are improved due to a more accurate estimate of the elevational offset as more frames of the membrane are used.

Table 5.2: Comparison of calibration precision and accuracy. Figures are given in millimetres.

<table>
<thead>
<tr>
<th>Phantom</th>
<th>Depth</th>
<th>Precision</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Centre</td>
<td>Mean</td>
</tr>
<tr>
<td>This chapter</td>
<td>8cm</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>This chapter</td>
<td>15cm</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Pagoulatos et al. (2001)</td>
<td>9cm</td>
<td>—</td>
<td>1–2</td>
</tr>
<tr>
<td>Pagoulatos et al. (2001)</td>
<td>16cm</td>
<td>—</td>
<td>1–4</td>
</tr>
<tr>
<td>Lindseth et al. (2003b)</td>
<td>8cm</td>
<td>0.6</td>
<td>—</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>6cm</td>
<td>—</td>
<td>0.4</td>
</tr>
<tr>
<td>Gee et al. (2005)</td>
<td>12cm</td>
<td>0.2</td>
<td>0.5</td>
</tr>
</tbody>
</table>

† These figures are based on a different phantom, and are not directly comparable.
nevertheless remains poorer than the Cambridge phantom (Chapter 4) and the mechanical device designed by Gee et al. (2005).

5.6 Conclusion

Wire segmentation in ultrasound images has always been a difficult problem due to the amount of noise in the B-scans. With the simple addition of a rubber membrane, we have shown how it is possible to solve this difficult segmentation problem. The wires can be segmented fully automatically and we are able to produce calibrations in real-time.

We have noted that the estimation of the elevational offset is relatively poor when the calibration is performed using a Z-phantom. This estimation can be improved by capturing 8 images of the planar membrane. These images can be captured before calibration, and therefore the calibrations can still be produced in real-time.

In order to perform accurate spatial calibration using our Z-phantom, the user needs to first capture 8 images of the plane from an oblique angle. The phantom is then located in space by using a pointer. This process takes approximately 2 minutes. The phantom is then scanned and calibrations are produced in a few seconds.

Although the accuracies achievable with our Z-phantom remain poorer than the accuracies achievable with some of the existing classes of phantoms, we have demonstrated that very rapid calibration is achievable with our Z-phantom, with accuracies comparable to other Z-phantoms.

In situations where accuracy is crucial, such as radiotherapy planning, calibration using other state-of-the-art phantoms is recommended. However, in situations where sub-millimetre accuracies are not necessary, our Z-phantom allows the user to perform scanning on-the-fly with different depth and zoom settings. Re-calibration due to depth or zoom changes takes only a few seconds to complete.
Chapter 6

Accurate Calibration

One of the ways to solve spatial calibration is to locate points in both the sensor and the B-scan’s coordinate system. A stylus is usually required in the process. Nevertheless, if locating the points in both coordinate systems is possible, spatial calibration can be solved in a closed form.

Muratore & Galloway Jr. (2001) calibrated their probe by locating points directly in the scan plane with a stylus. Assuming at least 3 non-collinear points have been located and that the scales are known, the calibration parameters are solved in a closed-form. This technique has several drawbacks. One of the more subtle disadvantages is that the technique requires two targets to be tracked simultaneously. The major difficulty is to align the tip of the stylus with the scan plane accurately. The finite thickness of the ultrasound beam makes the target visible in the B-scans even if the target is not exactly at the elevational centre of the scan plane.

Khamene & Sauer (2005) improved on this technique by imaging a rod transversely. Both ends of the rod are pointer calibrated to define the location and orientation of the rod in space. Each image of the rod sets up a constraint on the calibration parameters. Probe calibration can then be found using optimisation techniques. However, as we shall see in this chapter, this technique is very imprecise and inaccurate.

In this chapter, we have designed a novel phantom (cone phantom) that can be used to locate points in both the sensor and the B-scan coordinate systems (Hsu et al., 2007). We also present calibration results of the Cambridge stylus
6.1 Materials and Methods

6.1.1 The Calibration Phantoms

Figure 6.1 shows the cone phantom and the Cambridge stylus. During probe calibration, we will scan a point \( p \) that is on the phantom. This point \( p' \) can be segmented in the ultrasound image and changed to metric units by an appropriate scaling factor \( T_s \) to give \( p' = T_s p'' \). If the probe calibration \( T_{S→I} \) is known, then the segmented point can be mapped to world space by:

\[
p^W = T_{W→S} T_{S→I} p',
\]

(6.1)

where \( T_{W→S} \) is given by the Polaris readings.

(a) Cone Phantom  (b) Cambridge Sty-  (c) Cone Phantom  (d) Cambridge  (a) Cone Phantom  (b) Cambridge Sty-  (c) Cone Phantom  (d) Cambridge  (a) Cone Phantom  (b) Cambridge Sty-  (c) Cone Phantom  (d) Cambridge

Figure 6.1: The calibration phantoms.

6.1.1.1 Cone Phantom

The cone phantom (Figure 6.1(a)) is primarily comprised of a polypropylene block machined into the shape of two hollow cones with 1mm wall thickness. The centre of the circle where the two cones join is the fiducial point \( p \) that we will use for
probe calibration. The cones have outer diameters varying between 12mm and 52mm. They are supported by an aluminium frame structure consisting of two 135cm × 135cm square plates held together by four bolts. There are 3 cone-shaped divots \(d_i\) in the two blocks to fit a 3mm ball-pointed Polaris stylus. The positions of the divots, \(d_i^W\), are therefore known in 3D space. These divots serve as the principal axes of the phantom and are used to determine the phantom position in 3D space. The location of \(p\) is measured with a Mitutoyo CMM (Mitutoyo Corporation, Japan) relative to the divots, i.e. \(p^A = f_{\text{CMM}}(d_i^A)\), where the function \(f_{\text{CMM}}\) is determined by the coordinate measuring machine and is valid for arbitrary coordinate system \(A\). In particular, \(p^W = f_{\text{CMM}}(d_i^W)\) in world space. Every component of the phantom is precision manufactured by our workshop with a tolerance of ±0.1mm. The mechanical CMM has an accuracy of 0.01mm.

If we align the scan plane with the circle where the two cones join as shown in Figure 6.1(c), we get a circle with a 12mm diameter. This circle can be segmented with its centre \(p''\) automatically. As we shall see later, the image scales are necessary for reliable segmentation. Let us assume for the moment that the scales \(T_s\) are known, hence \(p'^i = T_sp''\) can be computed.

Now, recall that we have already found the coordinates of this centre in 3D space. We can transform this point to the sensor’s coordinate system by using the inverse of the position sensor’s readings:

\[
p^S = T^{-1}_{W \rightarrow S} p^W = T^{-1}_{W \rightarrow S} f_{\text{CMM}}(d_i^W).
\]

For probe calibration, we need to find the single transformation \(T_{S \rightarrow I}\) that best transforms \(\{p'^i\}\) to \(\{p^S_i\} = \{T^{-1}_{W \rightarrow S_i} f_{\text{CMM}}(d_j^W)\}\). This can be found (Arun et al., 1987) by minimizing

\[
f_{\text{cone}} = \sum_i \left| T^{-1}_{W \rightarrow S_i} f_{\text{CMM}}(d_j^W) - T_{S \rightarrow I} p'^i \right|.
\]

### 6.1.1.2 Cambridge Stylus

Figure 6.1(b) shows a photograph of our new Cambridge stylus. It consists of a stainless steel shaft, 120mm in length and 13mm in diameter. On one end of
the shaft, a platform is ground to fit a PassTrax (Traxtal Technologies, Canada) position sensor in such a way that the $z$-axis of the sensor is parallel to the shaft. The other end of the shaft is sharpened for accurate pointer calibration. As above, this means that the point $r^L$ is found by a pointer calibration. If the position sensor cannot be placed with its $z$-axis parallel to the shaft, the same technique can still be used provided that both ends of the shaft are pointer calibrated so that the orientation of the shaft is known.

The main feature of this stylus is that part is thinned in the shape of two cones, to meet at the critical point $p$. This point is precisely $20\text{mm}$ above the stylus’s tip. Here, the diameter is $1.5\text{mm}$. The stylus was precision manufactured by our workshop with a tolerance of $\pm 0.1\text{mm}$.

From the known geometric model of the stylus, $p^L = r^L - (0, 0, 20)^t\text{mm}$. We now place fiducial marks by aligning this critical point with the scan plane. From here, the calibration process is identical to that for a Polaris stylus, the only difference being that we are aligning the critical point, and not the stylus’s tip. The calibration parameters are found by minimizing

$$f_{\text{Cambridge}} = \sum_i \left| T_{W-S_i} T_{S-I} T_{S-p_i'} - T_{W-L_i} (r^L - (0, 0, 20)^t) \right|. \quad (6.5)$$

### 6.1.2 Segmentation

Figure 6.2 shows typical ultrasound images of the two phantoms when they are properly aligned. Figure 6.3 shows images of the same phantoms when they are slightly misaligned by approximately $1\text{mm}$. The images of the cone phantom (Figure 6.2(a) and Figure 6.3(a)) appear very similar. Thus, the cone phantom is aligned visually by moving the probe along its length until we get a circle with the smallest radius. The Cambridge stylus is aligned when the blob from the cones (Figure 6.3(b)) turns into a strong reflection as shown in Figures 6.2(b). This happens when the stylus’s shaft is perpendicular to the scan plane and a strong reflection is obtained.

An overview of the segmentation algorithms can be found in Appendices D.3.3 and D.3.4. For the cone phantom, we first place a user defined threshold on the B-scan to remove any unwanted reflections, resulting in a binary image consisting
6.1 Materials and Methods

(a) Cone Phantom  (b) Cambridge Stylus

Figure 6.2: Typical images of the different phantoms with the segmentation result imposed.

(a) Cone Phantom  (b) Cambridge Stylus

Figure 6.3: Typical images of the different phantoms when they are slightly misaligned.

mainly of the circle that is to be segmented. We now shift each pixel upwards by the temperature correction factor. Edge detection is performed by applying a Sobel filter to overlapping $3 \times 3$ blocks of the image. If the image scales are known, we can apply the Hough transform (Hough, 1959) on the resulting edges to detect a circle with a 12mm diameter.

When segmenting the image of the Cambridge stylus, we first search through the entire image to find the pixel with maximum intensity. This pixel serves as a starting point for the segmentation algorithm. We are anticipating a strong reflection and an image of the point spread function, which will be symmetrical about the direction of propagation of the ultrasound waves. This means that for a linear probe, there should be a vertical line of symmetry. We iterate this line through all possible angles, and across a horizontal distance about 10% of the image width, near the pixel with maximum intensity. The angles are incremented
in steps of $1^\circ$. During each iteration, the correlation of the image on either side of the line is computed. Once we have found the line of symmetry corresponding to the maximum symmetric auto-correlation, the required point is assumed to be on this line. We sum the intensity across each depth on this line and find the depth that has the highest intensity. The point on the line of symmetry corresponding to the brightest depth is segmented. This algorithm is used to segment the image of each stylus, and the results are imposed on Figures 6.2(b). The line of symmetry is shown as well. The near-vertical line of symmetry verifies that the point has been segmented correctly. Since the top, rather than the centre, of the surface is segmented, a correction is performed in the optimisation process to shift the segmented point downwards by half the stylus thickness. The thickness of the Cambridge stylus is 1.5mm.

In order to calibrate our linear probes in cold water, we now post-process the segmented points by shifting them upwards towards the probe face. Assuming that the user will measure the water temperature, the speed of sound in water at this temperature can be computed \cite{Bilaniuk93}. We then shift the pixels upwards by the temperature correction factor $\frac{\text{sound speed in water}}{1540}$.\n
\section{6.2 Results}

In order to measure the calibration quality of the different phantoms, we calibrated a Diasus (Dynamic Imaging Ltd., U.K.) 5–10MHz linear-array probe. The analog radio-frequency (RF) ultrasound data, after receive focusing and time-gain compensation but before log-compression and envelope detection, was digitized using a Gage CompuScope (Gage Applied Technologies Inc., U.S.A.) 14200 PCI 14-bit analog to digital converter, and transferred at 25 frames per second to a Pentium(R) 4 2.80GHz PC running Microsoft Windows XP. Since we have access to the RF data, the image scales could be set at our discretion. The B-scans were displayed at 0.1mm/pixel, with a cropped size of 30.0mm $\times$ 38.1mm. The probe was tracked using a PassTrax (Traxtal Technologies, Canada) target for the Polaris (Northern Digital Inc., Canada) optical tracking system.

In additional to the cone phantom and the Cambridge stylus, we also calibrated the same probe with phantoms shown in Figure 6.4. These phantoms
also perform calibration by locating fiducial points directly in the scan plane. Figure 6.4(a) and (b) show two standard Polaris styli, one with the a sharp tip and the other a spherical tip. We have measured the spherical tip of the Polaris stylus to be 3.0mm in diameter with a digital micrometer. They are similar to the approach used by Muratore & Galloway Jr. (2001), where a stylus was used to locate points in the scan plane. Figure 6.4(c) shows the rod stylus similar to the one used by Khamene & Sauer (2005). This technique was proposed as an improvement on the approach by Muratore & Galloway Jr. (2001). We have attached a 15cm long, 1.5mm thick rod to a Polaris active marker. Both ends of the rod are sharpened for accurate pointer calibration. The detail of calibrating with these phantoms has been outlined in Section 2.2. Figure 6.5 shows various images of the three styli. Figures 6.5(a)–(d) show images of the sharp, spherical and rod styli. Figures 6.5(e) and (f) show images of the sharp and spherical styli when they are misaligned with the scan plane by approximately 1mm.

![Figure 6.4](image_url)

(a) Sharp Stylus    (b) Spherical Stylus    (c) Rod Stylus

Figure 6.4: Calibration styli to be compared.

In order to compute the precision and accuracies achievable with each phantom, we performed a series of calibrations using each phantom. A pointer calibration was initially performed on each stylus. This was then followed by five
6.2 Results

Figure 6.5: B-scan images of the various styli. Figures (a) and (b) show images of the sharp and spherical styli when they are aligned with the scan plane. Figures (c) and (d) show images of the rod stylus when it is perpendicular and at an angle to the scan plane. In (e) and (f), misaligned images of the sharp and spherical styli are shown.

probe calibrations. The five probe calibrations are used to evaluate the repeatability of the calibration. During each probe calibration, 20 points were located throughout the image with the stylus. In order to evaluate the impact of the pointer calibration on probe calibration, this whole routine was again repeated five times, each time with a different pointer calibration, as shown in Figure 6.6.

For probe calibrations with the cone phantom, a pointer calibrated stylus (Figure 6.1 (b)) was used to locate the phantom in space. Ten readings were taken at each divot, and the mean was used for probe calibration. Five calibrations, each with 20 images of the cone spread throughout the B-scan, were performed. Again, this process was repeated five times, each time re-calibrating the stylus and relocating the phantom by its divots. Thus, a total of 25 calibrations was computed for each phantom.
6.2 Results

Figure 6.6: The calibration protocol. Five pointer calibrations were performed. For each pointer calibration, five probe calibrations were performed.

6.2.1 Precision

One way to assess the calibration quality is by computing its precision. Precision is measured by Equation 3.3:

\[ \mu_{CR} = \frac{1}{N} \sum_{i=1}^{N} \left( p_{Si} - \overline{p_{Si}} \right). \] (6.6)

It is obvious that the above measure is dependent on the point chosen in the B-scan. We will therefore measure the variation of the four corners and the centre of the B-scan.

In order to investigate whether the pointer calibration is sufficiently precise and measure its impact on the calibration result, we compute both calibration reproducibility measures outlined in Section 3.1.1. First, the precision for each group of the five calibrations specific to the same pointer calibration is calculated. These five precisions for the five groups of calibrations are then averaged, specifically by Equation 3.4:

\[ \mu_{CR'} = \frac{1}{25} \sum_{j=1}^{5} \sum_{i=1}^{5} \left( p_{jSi} - \overline{p_{jSi}} \right). \] (6.7)

The results are shown in Table C.1. Table C.2 shows the precision without
differentiating the 25 calibrations, i.e. by Equation 3.3:

\[
\mu_{CR} = \frac{1}{25} \sum_{i=1}^{25} \left( p_{S_i} - p_{S_i}^* \right).
\] (6.8)

The mean variations of the four corners and the centre of the B-scan are given in Figure 6.7. Each of these variations can be broken down into their corresponding components in the lateral, axial and elevational direction. Figure 6.8 shows the mean variation of the four corners and the centre of the B-scan, including pointer calibration variations, in these three principal axes.

![Figure 6.7: Probe calibration precisions.](image)

### 6.2.2 Accuracy

While precision measures calibration reproducibility, this does not reflect the accuracy of the calibration, since there may be a systematic error. We assess the accuracy of our calibrations by measuring the point reconstruction accuracy of a point target. We scan the tip of a wire that is 1.5mm thick in water at room temperature. A typical image of the wire is shown in Figure 6.9. The image of the wire tip \( w \) is reconstructed in space by using the obtained calibrations. The 3D location of the wire tip is found by using an independent stylus. In fact, we have used the Cambridge Stylus, but pointer calibrated by its tip, rather than the critical point. The RMS error of the pointer calibration is 0.8mm. Ten readings of the wire tip were obtained to eliminate errors from the stylus’s readings. The true
Figure 6.8: Mean variation of the four corners and the centre of the B-scan in the three principal axes of the scan plane.

**location** of the wire tip is taken to be the mean of the stylus’s readings. Accuracy (PRA) is measured by the amount of mismatch between the reconstructed image and the true location (Equation 3.5):

\[
\mu_{\text{PRA}} = \left| T_{W \leftarrow L_i} w^L - T_{W \leftarrow S} T_{S \leftarrow I} T_{I} w^R \right|.
\]  

(6.9)

Figure 6.9: B-scan of the wire used for point reconstruction accuracy assessment.

Point reconstruction accuracy may be dependent on the position of the probe and the position of the wire tip in the B-scan. We therefore capture images of the wire tip at five different locations in the B-scans—near the four corners and the centre of the B-scan. The wire tip is aligned with the scan plane by hand, and segmented from the B-scans manually. For each location, the probe is held...
perpendicular to the wire and rotated through a full revolution about the axial axis at six different angles. Furthermore, five images of the wire are taken at each angle and at each location in the B-scans. A total of 6 angles × 5 locations × 5 repetitions = 150 images of the wire tip is captured. The point reconstruction accuracy at each region in the B-scan is shown in Table C.3. The mean point reconstruction accuracy of the five regions is shown in Figure 6.10.

![Different estimates of the probe calibration accuracy.](image)

There are two main sources of error in our point reconstruction accuracy valuation. First, the wire tip may not be aligned perfectly with the scan plane. This is why we scan the tip of the wire five times for each probe position and region in the B-scan. We calculate the centroid of the wire tip and evaluate the mean point reconstruction error from the centroid to the true location. This point reconstruction accuracy, where alignment errors are eliminated, is shown in Figure 6.10.

Secondly, point reconstruction accuracies may contain an error from locating the wire tip by using a stylus. Reconstruction precision (Section 3.1.1) measures the variation of the reconstructed points about their centroid. This measure does not include errors from locating the wire tip and can be used to verify the validity of our point reconstruction accuracies. We also give the reconstruction precision in Figure 6.10.
Point reconstruction accuracy consists of a bias and a spread. The bias consists of calibration error and other systematic errors, such as systematic misalignment, systematic segmentation, sound speed errors and an incorrect position of the true location. Since we have positioned our probe in six different angles, any systematic error will cause the reconstructed points to be clustered into six groups. An exaggerated representation of the situation is shown in Figure 6.11. For each probe angle, we compute the centroid of the points belonging to that cluster. The bias is measured to be the distance from the centroid to the true location. The mean of the six biases are given in Figure 6.10. The spread is measured by the RMS distance from each reconstructed point to the corresponding centroid, i.e. the standard deviation of the points within each cluster. Table C.4 shows the biases and spreads.

Figure 6.11: Calculation of bias and spread of the point reconstruction accuracy.

6.3 Discussion

Figure 6.7 shows the precision of the probe calibrations with a fixed pointer calibration. This precision can be achieved if the pointer calibration is supplied by the manufacturer of the stylus. Probe calibration precision that includes variations due to pointer calibration imprecision is also shown in the same figure. The slight increase in error of the sharp and Cambridge stylus is caused by pointer calibration imprecision. The precision of the spherical stylus improved slightly.
This means that the pointer calibration of the spherical stylus is very precise and the main source of error is caused by probe calibration imprecision. The precision of the cone phantom remains unchanged, thus phantom positioning in 3D space with the spherical stylus is repeatable.

From the two precision measures, we see that the rod stylus performs considerably worse than the other styli/phantom. There is also a discrepancy between the two precision measures. This means that although for a particular pointer calibration, a probe calibration precision of about 2mm can be achieved, a significant worse precision of 5mm should be expected when given an arbitrary probe calibration.

The variations in the three principal axes due to calibration imprecision is given in Figure 6.8. If the calibration is correct, then variations in the lateral direction will be primarily segmentation errors. The variation in the axial direction will be caused by a combination of segmentation and sound speed errors, and the elevational direction variations caused by alignment errors. In all calibrations, the variation is greatest in the elevational direction. This agrees with our knowledge of such techniques, where aligning the scan plane with the phantom is difficult. The cone phantom gives the best precision in the lateral direction. This is because a circle inherently contains much more information than an isolated point. The image of the cone is nevertheless blurred slightly in the lateral direction due to the fact that the ultrasound axial resolution is approximately three times better than the lateral resolution, which causes a lateral variation in the segmentation algorithm. However, from the high precision observed in the lateral direction, we can conclude that segmentation error is low compared to other sources of errors. The relatively high elevational error signifies that aligning the cones phantom is harder than the spherical and the Cambridge stylus. This is because aligning such styli is aided by the strong reflection when the styli are perfectly aligned. The rod stylus performs poorly as noted in the previous paragraph.

The biases are only slightly less than the point reconstruction accuracy shown in Figure 6.10. The spreads of the point reconstruction errors are much less than the bias. This means that the majority of the point reconstruction error lies in systematic errors, such as calibration, systematic misalignment, systematic segmentation and sound speed errors. The spreads given in Figure 6.10 (Table C.4)
are slightly larger, but follow closely the precision given in Figure 6.7 (Table C.2). This spread is a combination of variations due to calibration, position sensor, segmentation and phantom misalignment.

We have calculated two PRAs, one of which eliminates wire misalignment errors. The difference between the two PRA measures is found to be less than 0.1mm. If we assume that the wire can be aligned without a systematic bias, then the result shows that the wire tip is well aligned during accuracy evaluation.

It may be possible that a significant portion of the point reconstruction error is caused by the incorrect location of the wire tip. We have therefore computed reconstruction precision. This measures the variation of the reconstructed points from their centroid. The difference between the two measures is attributed to errors in the true location of the wire and calibration errors. Since the reconstruction precisions are only slightly less than the point reconstruction accuracies, this verifies that our point reconstruction accuracies are not biased by the error in locating the wire tip.

The relatively high error in the point reconstruction accuracy using the calibrations from the rod stylus further suggests that such a phantom is unreliable for probe calibration. Both the sharp and spherical stylus have similar point reconstruction accuracies, being slightly worse than the cone phantom and Cambridge stylus. This difference is mainly attributed to the fact that it is difficult to align the scan plane with the tip of the sharp and spherical styli. Even though the spherical stylus produces a better reflection, which can suggest whether we have aligned its tip correctly, this information is not found to be helpful as the calibration accuracies have not been improved. In fact, the accuracy of a spherical stylus is worse than the sharp stylus. This is due to the 3mm thick spherical tip of the stylus. Even with good visual feedback, it is still difficult to align the tip of the spherical stylus with the scan plane. The Cambridge stylus produces better accuracies than other styli, because any misalignment of the stylus with the scan plane is readily visible in the B-scans. The cone phantom produces the best accuracy.

The poor performance of the rod stylus is due to its design and calibration routine, which requires us to poke the scan plane in different orientations with the rod. If the scan plane is poked with the rod parallel to the normal of the scan
plane, then the scan plane can be anywhere along the length of the rod. Thus ideally the scan plane should be poked from different directions, each as oblique as possible to the normal of the scan plane. However, the image of the rod degrades as its inclination with the normal increases, resulting in segmentation errors. So, for a good calibration, we need both a good image of the rod and a high inclination of the rod relative to the normal of the scan plane. These are two incompatible conditions. As a result, an imprecise and inaccurate calibration is obtained.

Although precision and accuracy may be the most important measures to judge a calibration technique, there are other factors that should be noted. Firstly, the image scales need to be known in order to segment the circle for the cone phantom. We have this information because we have access to the RF data. This may not be the case for other users, in which case the scales may need to be determined by another technique, such as using the ultrasound machine’s distance measurement tool (Chapter 4). The ultrasonic settings such as transmitter and receiver gain and time gain compensator need to be set manually by the user for optimal segmentation. The user also needs to ensure that the correct circle has been segmented by the algorithm, rejecting and repeating any obvious incorrect segmentations.

The stylus has the advantage that segmentation is not highly dependent on the ultrasound machine settings. The image scales can also form part of the optimisation process, when functions $f_{\text{stylus}}$ (Equation 2.6) and $f_{\text{cone}}$ (Equation 6.4) are minimized. However, it appears that function $f_{\text{rod}}$ (Equation 2.7) has multiple local minima. In our calibrations, it is necessary to fix the image scales and choose a starting point sufficiently near the solution for the optimisation to converge to the correct minimum. Nevertheless, since the phantom does not need to be aligned with the scan plane, calibration is easy and rapid to perform.

6.4 Conclusion

We have presented two phantoms that are both easy to use and produce accurate calibrations. The rod stylus is very simple and quick to use, but produces very poor precision and accuracy. This phantom may be useful to obtain a quick
estimate of the calibration, although this may be needed in the first place for
the optimisation when using such a phantom. The Cambridge stylus is clearly an
improvement on both the sharp and the spherical stylus with a small modification
to the design. Better accuracies are achieved, while maintaining its ease of use and
calibration simplicity. The cone phantom produces the best calibration accuracy
at the expense of a more sophisticated phantom, segmentation and calibration
protocol. Another advantage of the cone phantom over every other stylus is that
calibration does not require two targets to be tracked simultaneously.
Chapter 7

Phantom Comparison

In the last three chapters, we have solved some of the key problems facing probe calibration today. In this chapter, we will answer the question “how should one choose a calibration phantom 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.

7.1 Accuracy

Table 7.1 shows the precision (CR) and accuracy (PRA) achievable with the different phantoms. These results are from this thesis or a paper from our research group, using the same precision and accuracy measures, on the same accuracy assessment phantom and similar ultrasound settings when performing the calibrations. Hence the figures in this table are directly comparable.

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 7.2 shows the quoted precision reported by the various research groups. We have not included accuracy measurements for reasons outlined in Section 3.1.2.
Table 7.1: Precision and accuracy of calibrations performed by our research group. The top half of the table shows the calibrations on a linear probe, and the bottom half shows the calibrations on a curvilinear probe. The precision and accuracy measurements are in millimetres.

<table>
<thead>
<tr>
<th>Phantom</th>
<th>Depth</th>
<th>Precision (CR)</th>
<th>Accuracy (PRA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Centre</td>
<td>Mean</td>
</tr>
<tr>
<td>Point (Cone)</td>
<td>3cm</td>
<td>0.27</td>
<td>0.59</td>
</tr>
<tr>
<td>Stylus (Spherical)</td>
<td>3cm</td>
<td>0.31</td>
<td>0.44</td>
</tr>
<tr>
<td>Stylus (Cambridge)</td>
<td>3cm</td>
<td>0.45</td>
<td>0.61</td>
</tr>
<tr>
<td>Plane</td>
<td>3cm</td>
<td>0.39</td>
<td>0.57</td>
</tr>
<tr>
<td>Cambridge Phantom</td>
<td>3cm</td>
<td>0.83</td>
<td>0.88</td>
</tr>
<tr>
<td>Mechanical Instrument</td>
<td>6cm</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>(Gee et al., 2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanical Instrument</td>
<td>12cm</td>
<td>0.24</td>
<td>0.44</td>
</tr>
<tr>
<td>(Gee et al., 2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z-phantom</td>
<td>8cm</td>
<td>0.47</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>15cm</td>
<td>1.07</td>
<td>1.54</td>
</tr>
</tbody>
</table>

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 are of the same order. In general, precision increases as the depth setting becomes shallower. This is exactly what we would expect, since it is more difficult to constrain a larger image. The exception is the plane-based technique used at a very shallow depth (3cm). The small B-scan area limits capturing images of the plane from different orientations, leading to a low precision.

### 7.2 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
Table 7.2: Precision of calibrations performed by the various research groups. The measurements are in millimetres.

<table>
<thead>
<tr>
<th>Phantom</th>
<th>Probe</th>
<th>Depth</th>
<th>Precision (CR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Centre</td>
</tr>
<tr>
<td>Point (Meairs et al., 2000)</td>
<td>Linear</td>
<td>3.5cm</td>
<td>1.81</td>
</tr>
<tr>
<td>Point (Lindseth et al., 2003b)</td>
<td>Linear</td>
<td>8cm</td>
<td>0.62</td>
</tr>
<tr>
<td>Point (Sphere) (Chapter 6)</td>
<td>Linear</td>
<td>3cm</td>
<td>0.31</td>
</tr>
<tr>
<td>Point (Cone) (Chapter 6)</td>
<td>Linear</td>
<td>3cm</td>
<td>0.27</td>
</tr>
<tr>
<td>Plane (Rousseau et al., 2003)</td>
<td>Linear</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Plane (Rousseau et al., 2005)</td>
<td>Sector</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Plane (Rousseau et al., 2003)</td>
<td>Linear</td>
<td>3cm</td>
<td>0.39</td>
</tr>
<tr>
<td>Cambridge phantom</td>
<td>Linear</td>
<td>3cm</td>
<td>0.83</td>
</tr>
<tr>
<td>2D alignment (Lindseth et al., 2003b)</td>
<td>Linear</td>
<td>8cm</td>
<td>0.44</td>
</tr>
<tr>
<td>2D alignment (Leotta, 2004)</td>
<td>Sector</td>
<td>10cm</td>
<td>0.67</td>
</tr>
<tr>
<td>Z-phantom (Lindseth et al., 2003b)</td>
<td>Linear</td>
<td>8cm</td>
<td>0.63</td>
</tr>
<tr>
<td>Z-phantom (Chapter 5)</td>
<td>Curvilinear</td>
<td>8cm</td>
<td>0.47</td>
</tr>
<tr>
<td>Z-phantom (Chapter 5)</td>
<td>Curvilinear</td>
<td>15cm</td>
<td>1.07</td>
</tr>
<tr>
<td>Image registration (Blackall et al., 2000)</td>
<td>Linear</td>
<td>4cm</td>
<td>—</td>
</tr>
</tbody>
</table>

into account when choosing such a phantom. Based on these factors, we will now compare some of the most widely used phantoms, namely: point phantom, stylus, plane phantom, Z-phantom and their variants.

The two variants of the point phantom to be compared are the cone phantom and the spherical stylus discussed in Chapter 6. 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 thesis. 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 of the spherical stylus over a typical point phantom is that the spherical stylus can be moved around. Nevertheless, this 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 comparison, we will disregard some phantoms used by individual groups. 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 very 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 7.3 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.

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
7.2 Phantom Comparison

Table 7.3: Probe calibration factors for the different phantoms.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Point</th>
<th>Stylus</th>
<th>Plane</th>
<th>2D alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SS</td>
<td>CP</td>
<td>CS</td>
<td>PP</td>
</tr>
<tr>
<td>Precision</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>6↑</td>
</tr>
<tr>
<td>Accuracy</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Easy to use (Novice)</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Easy to use (Expert)</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Segmentation</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Speed</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Reliability</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Phantom simplicity</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Linear probe</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Curvilinear probe</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

SS — Spherical stylus  PP — Plexiglass plate
CP — Cone phantom  C — Cambridge phantom
CS — Cambridge stylus  Z — Z-fiducial phantom

↑ This is based on the precision at 3cm. The precision is better when calibrating at a higher depth.
‡ It is possible, but not advisable, to use these phantoms to calibrate the corresponding probe.

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 very 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 with its extremely simple design. 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 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.
Chapter 8

Conclusion

In this thesis, 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. We have analyzed the different measures used to assess the calibration quality and quantified the accuracy of each phantom.

In freehand 3D ultrasound calibration, we would like to answer three questions: 1. How reliable is the calibration? 2. How can we perform calibration as quickly as possible? 3. How can we perform calibration as accurately as possible?

We have shown in Chapter 4 that it is possible to use an eigenvalue metric to determine whether the calibration parameters are sufficiently constrained leading to a reliable plane-based calibration. If the calibration parameters are inadequately constrained, the minimum eigenvalue of the Hessian of the objective function when evaluated at the global minimum will be close to zero. This information can be used to show whether or not a reliable calibration has been obtained.

We have also identified the critical motions that are required to constrain the calibration parameters. A good calibration should therefore consist of repetitions of these motions. We have shown that including symmetrical redundant constraints does not bias the solution and can be included in the optimisation process.

In Chapter 5 we have shown how to segment isolated points in a Z-phantom automatically by the additional of the membrane. Since such a phantom only requires one frame for probe calibration and does not require any alignment of
the probe with the phantom, calibration can be performed within seconds. This is a major improvement from any previous calibration attempts, where a few minutes were necessary. The phantom is also very easy to use.

In Chapter 6 we presented two phantoms that can produce accurate calibrations. The cone phantom produces accuracies matching the Cambridge phantom, while at the same time it is much easier to use for a novice user. It also outperforms the Cambridge phantom in terms of precision and calibration time. The Cambridge stylus is small in size, easy to use and an accuracy matching a plane phantom (plexiglass) can be achieved.

In the last chapter, we have pointed out the situations where a particular phantom may be more suitable than others. 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 in between these extremes, offering moderate accuracy, ease of use and phantom complexity.
Appendix A

Plane-based Phantom

A.1 The Curvilinear Probe

It is clear that the temperature correction routine is more sophisticated for the curvilinear probe than for the linear probe. This includes the detection of the probe shape and curvature in addition to the mathematical adjustments, which is an extra step for the user during calibration. We therefore investigate what is the expected error, if we use the temperature correction routine for a linear probe to approximate the exact correction.

Figure A.1 (a) shows an exaggerated diagram when the plane is scanned in such a way that the plane appears as a horizontal line, or in fact a slight curve. The exact correction is shown with arrows pointing in the direction of the shift, the result is a horizontal straight line, shown as a dashed line in the diagram. This is the true position of the plane in space. The approximate correction is shown as arrows ending with a circle. From the diagram it can be deduced that the maximum error occurs at the middle of the image, where the point is over corrected. If this central point appeared at a depth $d$ in the B-scan, and $t$ represents the temperature correction factor, then the point is corrected to position $td$. For an exact correction, the point should be corrected to $t(d - c) + c$, where $c$ is the distance from the middle of the probe face to the top of the image. The overcorrection is thus the difference $(1 - t)c$. The line detection algorithm is likely to estimate the line to be around the middle of the two extreme points as the best fit line (shown as a dashed line), and so the expected error $e$ should be
about half the overcorrection error, i.e. $e \approx \frac{1}{2}(1 - t)c$. Since $t = 0.963$ for water at 20 Celsius, and $c$ is less than 10mm for common abdominal curvilinear probes, this error ($e \approx 0.1$–0.2mm) is negligible compared with other errors of several millimetres in a freehand 3D ultrasound system.

Figure A.1: The difference between the corrected planes when applying the correct and approximate correction routines to the displayed plane. The approximate correction is shown with arrows ending with a circle. It is assumed that a best fit line will be fitted to the non-linear corrected points. Both diagrams are exaggerated to show the correction error more clearly.

Figure A.1 (b) shows an exaggerated diagram when the plane is scanned from an oblique direction. The difference between the two correction routines in this case is mainly a rotational error, rather than a distance error. The incorrectly corrected line will appear not as steep as the real position of the plane, while the distance at the centre will remain roughly at the same place as before. This error is difficult to quantify, as it is dependent on the depth, curvature and width of the probe, slant of the plane, and even the thresholds in the line segmentation algorithm. As an estimate of the expected error, we compute the error using the dimensions of our Toshiba curvilinear probe, as shown in Figure A.2. Assuming the depth is set at 17cm. The maximum error occurs when the plane is scanned most obliquely. In this case, the error can be computed by simple trigonometry to be $0.6^\circ$. In practise the error will be less than this value. Since this type of
motion only occurs in approximately 20% of the calibration, its impact on the calibration error is limited.

Figure A.2: An estimate of the error caused by temperature correction differences. The dimensions are shown in centimetres for our Toshiba curvilinear probe.

We verify that using the simple temperature correction is a good approximation to the exact correction by performing an additional 10 calibrations with the Toshiba curvilinear 3.75MHz probe using the simple temperature correction routine, and compare this result with the previous calibrations with the same probe, but with the exact temperature correction routine applied. Since the probe shape no longer needs to be detected, it is therefore not necessary to have the sides of the probe visible within the images. This allows the B-scans to be cropped and zoomed, the more common setting for the curvilinear probe in clinical environments. The difference in settings, including the cropped B-scan sizes, for the two sets of calibrations is shown in Table A.1.

The repeatability of the curvilinear probe shape detection is shown in Table A.2, with the values in pixels. It is clear that the centre and radius of the probe are accurate to within 2 pixels (0.3mm).

Figure A.3 shows the mean 3D errors of the four corners and the centre of the B-scan for the curvilinear probe when the two temperature correction routines are applied. The low errors in Figure A.3(b) indicates an accurate temperature
A.2 Solution of the Criterion Function

Table A.1: Setting differences for the Toshiba curvilinear 3.75MHz probe used for determining the effect of the different temperature correction routines. The depth is set at 6cm.

<table>
<thead>
<tr>
<th>Correction</th>
<th>Cropped B-scan</th>
<th>Refresh Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>388 × 380</td>
<td>36Hz</td>
</tr>
<tr>
<td>Approximate</td>
<td>496 × 416</td>
<td>18Hz</td>
</tr>
</tbody>
</table>

Table A.2: Repeatability of the estimation for curvilinear probe centre position and radius. The unit of the values in the table is pixels. The scales in each B-scan are determined explicitly to be 0.143mm/pixel.

<table>
<thead>
<tr>
<th>Centre</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>Mean</td>
<td>193.1</td>
</tr>
<tr>
<td>RMS</td>
<td>0.115</td>
</tr>
<tr>
<td>95%</td>
<td>±0.237</td>
</tr>
</tbody>
</table>

correction is not required for high precision calibrations. In fact, a slightly better precision is observed when an approximation to the temperature correction routine is used. This is because there is no uncertainty in detecting the probe shape. In any case, the variation due to probe calibration imprecision is small given that uncertainty in the position sensor alone is 0.035cm. An exact temperature correction routine for common abdominal curvilinear probes is therefore not necessary to achieve high precision.

A.2 Solution of the Criterion Function

We have utilized the non-linear Levenberg-Marquardt optimisation algorithm (More, 1977) to solve for the solution to the system of non-linear equations involving nine parameters. The calibration parameters are solved in three stages:
A.3 Verification of the sphere distances

Although the location of each sphere was precisely manufactured, the phantom has shrunk over the years by losing water. This has caused the spheres to be closer than the manufactured dimensions, typically near the scanning window at the top away from the rigid walls. We first measured the distances between the spheres. This was done by capturing multiple ultrasound images in the direction

Figure A.3: The effect of temperature correction routines on the consistencies of the four corners and the centre of the B-scan.

the angles are first solved, then the distances, and then a final global optimisation to fine-tune the solution (Prager et al., 1998). As is the case for all optimisation algorithms, a starting point is necessary. This is chosen at random, but within a reasonable estimate of what the parameters can be. This means that the solution given by the optimisation algorithm is dependent on the starting point. Hence, we need to ensure that the optimisation converges to the global minimum, rather than any other local minimum. Based on trial runs of the optimisation, the convergence to a non-global minimum occurs less than 50% of the time. We therefore repeat the optimisation 50 times at different starting points and accept the solution with the minimum error. This process takes less than a second on our 3.0GHz PC. The probability that the optimisation still converges to the incorrect minimum is 1 in $10^{15}$.

A.3 Verification of the sphere distances
of the coplanar spheres with the Diasus 5–10MHz probe with multiple foci down its depth. The horizontal and vertical distances between the spheres were measured using the distance measuring tool provided by the ultrasound machine in a single image. The measurements are shown in Table A.3. The vertical distances between spheres have shrunk by approximately 5%, while the horizontal distances remained roughly unchanged. We modelled the locations of the spheres to lie on a rectangular grid, with the measured vertical distance and the manufactured horizontal distance. We have kept the manufactured horizontal distance as we do not anticipate any shrinking or expanding in the horizontal direction as the phantom is surrounded by rigid walls on its sides. The very slight difference in the measured distances is within measurement error.

Table A.3: The measured and original manufactured distances between two or three spheres, depending on the number of spheres that could be seen in each B-scan. All distances are in millimetres, displayed in the form mean ± standard deviation.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Measured Distances</th>
<th>Original Distances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Horizontal</td>
<td>Vertical</td>
</tr>
<tr>
<td>Diasus 3cm</td>
<td>30.21 ± 0.087</td>
<td>9.47 ± 0.12</td>
</tr>
<tr>
<td>Diasus 6cm</td>
<td>30.05 ± 0.181</td>
<td>19.14 ± 0.13</td>
</tr>
</tbody>
</table>

### A.4 Eigenvalues of Valid Calibrations

Suppose that the user captured a series of scans during a plane-based calibration leading to an objective function with \( n \) constraints: 

\[
f(x) = \sum_{i=1}^{n} C_i(x)^2.
\]

Let us assume the global minimum of \( f \) is at \( x = x_c \), and let \( H(x) \) denotes the Hessian of \( f(x) \). During the eigenvalue analysis, \( H(x_c) \) is evaluated. Suppose that \( v \) is an eigenvector of \( H(x_c) \) with an associated eigenvalue \( \lambda \), then \( H(x_c)v = \lambda v \).

If the user repeated the above procedure and captured \( m \) sets of the same of images, then the objective function is 

\[
f_2(x) = mf(x) = m \sum_{i=1}^{n} C_i(x)^2,
\]

with the
corresponding Hessian $H_2(x_c) = mH(x_c)$. Now,

\[
\begin{align*}
H(x_c)v &= \lambda v \\
\Rightarrow mH(x_c)v &= m\lambda v \\
\Rightarrow H_2(x_c)v &= (m\lambda)v.
\end{align*}
\]

The eigenvectors therefore remain the same, with the corresponding eigenvalues increasing by a factor or $m$. 
Appendix B

Z-fiducial Phantom

B.1 Rejection Criterion for the Z-phantom

After we have detected the wires in each row, we try to map each wire to the corresponding wire on the phantom. We know the distance between the wires that are perpendicular to the polyacetal blocks. If we assume that the probe was held roughly parallel to the blocks, the distances between the parallel wires should correspond roughly to the distance on the B-scans. We therefore first search for wires that were apart by the required distances. This automatically rejects any outliers. After these wires are successfully detected, there should be one and only one wire between each pair of parallel wires. If only one wire is detected between the pair of parallel wires, this location is accepted as a fiducial. If there is more than one positive detection, the location with the highest intensity is accepted. Should at any stage a wire fail to be detected, the corresponding fiducial is automatically rejected.

B.2 Correcting the Elevational Offset

After performing spatial calibration with the Z-phantom, we obtain a preliminary transformation

\[ T_{S \rightarrow I} = \begin{pmatrix}
    a_{11} & a_{12} & a_{13} & a_{14} \\
    a_{21} & a_{22} & a_{23} & a_{24} \\
    a_{31} & a_{32} & a_{33} & a_{34} \\
    0 & 0 & 0 & 1
\end{pmatrix}. \]
B.2 Correcting the Elevational Offset

Since the elevational offset $a_{34}$ is inaccurate, we replace it with a variable $t_z$. The above transformation becomes

$$T_{S^{-1}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we scan the membrane, we obtain points $p_i^I$ on the planar membrane together with their corresponding sensor readings $T_{W-S_i}$. We can therefore map each point on the planar membrane to the world coordinate system by

$$p_i^W = \begin{pmatrix} p_{ix}(t_z) \\ p_{iy}(t_z) \\ p_{iz}(t_z) \\ 1 \end{pmatrix} = T_{W-S_i} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} p_i^I.$$

We note that $p_{ix}(t_z), p_{iy}(t_z)$ and $p_{iz}(t_z)$ are linear. Since the planar membrane is positioned at a known location relative to the divots, we can compute its equation, $Ax + By + Cz + 1 = 0$, in the world coordinate system. Now, each $p_i^W$ lies on this planar membrane, hence we obtain a linear equation with one unknown for each $i$,

$$a_it_z + b_i = Ap_{ix}(t_z) + Bp_{iy}(t_z) + Cp_{iz}(t_z) + 1 = 0.$$

The elevational offset $t_z$ is calculated by least-squares optimization that best fits the system of linear equations.
Appendix C

Cambridge Stylus and Cone Phantom

Table C.1 shows the mean of the precisions of the probe calibrations specific to each pointer calibration, and Table C.2 shows the precision of the same calibrations without differentiating the calibrations. Table C.3 shows the point reconstruction accuracies in the different regions of the B-scan. Table C.4 shows the biases and spreads, as explained in Figure 6.11, of the point reconstruction accuracies.

Table C.1: Precision of the probe calibrations specific to each stylus pointer calibration in millimetres.

<table>
<thead>
<tr>
<th>Point</th>
<th>Sharp</th>
<th>Spherical</th>
<th>Rod</th>
<th>Cone</th>
<th>Cambridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top left</td>
<td>0.52</td>
<td>0.53</td>
<td>4.35</td>
<td>0.33</td>
<td>0.50</td>
</tr>
<tr>
<td>Top right</td>
<td>0.53</td>
<td>0.51</td>
<td>1.31</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>Bottom left</td>
<td>0.53</td>
<td>0.54</td>
<td>1.61</td>
<td>1.08</td>
<td>0.55</td>
</tr>
<tr>
<td>Bottom right</td>
<td>0.47</td>
<td>0.47</td>
<td>1.55</td>
<td>0.71</td>
<td>0.47</td>
</tr>
<tr>
<td>Centre</td>
<td>0.51</td>
<td>0.51</td>
<td>1.46</td>
<td>0.30</td>
<td>0.51</td>
</tr>
<tr>
<td>Mean</td>
<td>0.51</td>
<td>0.51</td>
<td>2.05</td>
<td>0.59</td>
<td>0.51</td>
</tr>
</tbody>
</table>
### Table C.2: Precision of undifferentiated probe calibrations in millimetres.

<table>
<thead>
<tr>
<th>Point</th>
<th>Sharp</th>
<th>Spherical</th>
<th>Rod</th>
<th>Cone</th>
<th>Cambridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top left</td>
<td>0.64</td>
<td>0.57</td>
<td>4.36</td>
<td>0.39</td>
<td>0.63</td>
</tr>
<tr>
<td>Top right</td>
<td>0.64</td>
<td>0.45</td>
<td>4.38</td>
<td>1.14</td>
<td>0.71</td>
</tr>
<tr>
<td>Bottom left</td>
<td>0.55</td>
<td>0.40</td>
<td>5.69</td>
<td>0.82</td>
<td>0.71</td>
</tr>
<tr>
<td>Bottom right</td>
<td>0.72</td>
<td>0.45</td>
<td>5.77</td>
<td>0.33</td>
<td>0.58</td>
</tr>
<tr>
<td>Centre</td>
<td>0.50</td>
<td>0.31</td>
<td>5.02</td>
<td>0.27</td>
<td>0.45</td>
</tr>
<tr>
<td>Mean</td>
<td>0.61</td>
<td>0.44</td>
<td>5.04</td>
<td>0.59</td>
<td>0.61</td>
</tr>
</tbody>
</table>

### Table C.3: Point reconstruction accuracies in different regions of the B-scan in millimetres.

<table>
<thead>
<tr>
<th>Point</th>
<th>Sharp</th>
<th>Spherical</th>
<th>Rod</th>
<th>Cone</th>
<th>Cambridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top left</td>
<td>4.09</td>
<td>4.26</td>
<td>8.22</td>
<td>1.53</td>
<td>2.78</td>
</tr>
<tr>
<td>Top right</td>
<td>3.68</td>
<td>3.72</td>
<td>7.51</td>
<td>1.83</td>
<td>2.43</td>
</tr>
<tr>
<td>Bottom left</td>
<td>2.77</td>
<td>3.44</td>
<td>7.85</td>
<td>1.78</td>
<td>1.94</td>
</tr>
<tr>
<td>Bottom right</td>
<td>2.86</td>
<td>3.68</td>
<td>7.89</td>
<td>1.84</td>
<td>2.21</td>
</tr>
<tr>
<td>Centre</td>
<td>2.65</td>
<td>3.07</td>
<td>7.25</td>
<td>1.86</td>
<td>1.52</td>
</tr>
<tr>
<td>Mean</td>
<td>3.21</td>
<td>3.63</td>
<td>7.75</td>
<td>1.77</td>
<td>2.18</td>
</tr>
</tbody>
</table>

### Table C.4: Bias and spread of the point reconstruction accuracies. The figures are quoted as (bias, spread) in millimetres.

<table>
<thead>
<tr>
<th>Point</th>
<th>Sharp</th>
<th>Spherical</th>
<th>Rod</th>
<th>Cone</th>
<th>Cambridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top left</td>
<td>4.03, 0.96</td>
<td>4.20, 0.89</td>
<td>7.75, 4.88</td>
<td>1.48, 0.58</td>
<td>2.68, 0.94</td>
</tr>
<tr>
<td>Top right</td>
<td>3.57, 1.05</td>
<td>3.62, 0.94</td>
<td>6.93, 4.88</td>
<td>1.69, 0.96</td>
<td>2.28, 1.02</td>
</tr>
<tr>
<td>Bottom left</td>
<td>2.71, 0.70</td>
<td>3.40, 0.58</td>
<td>7.32, 5.67</td>
<td>1.70, 0.63</td>
<td>1.81, 0.80</td>
</tr>
<tr>
<td>Bottom right</td>
<td>2.78, 0.80</td>
<td>3.65, 0.56</td>
<td>7.27, 5.67</td>
<td>1.80, 0.45</td>
<td>2.12, 0.69</td>
</tr>
<tr>
<td>Centre</td>
<td>2.58, 0.80</td>
<td>3.02, 0.68</td>
<td>6.85, 5.17</td>
<td>1.83, 0.50</td>
<td>1.37, 0.77</td>
</tr>
<tr>
<td>Mean</td>
<td>3.14, 0.86</td>
<td>3.58, 0.73</td>
<td>7.22, 5.26</td>
<td>1.70, 0.62</td>
<td>2.05, 0.84</td>
</tr>
</tbody>
</table>
Appendix D

Algorithms and Source Code

D.1 Mathematical Functions

D.1.1 Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm (More, 1977) is described and implemented by Prager et al. (1998) to find the solution $\phi$ of the function $f(\theta, \phi) = 0$, where $\theta$ is the measurement. If we expand $f$ by its Taylor expansion, we can find the update $\Delta \phi$ to the current estimate $\phi_i$:

$$0 = f(\theta, \phi) \approx f(\theta, \phi_i) + \frac{\delta f(\theta, \phi_i)}{\delta \phi} (\phi - \phi_i)$$

$$\Rightarrow \Delta f = J(\phi - \phi_i) = J\Delta \phi,$$

where $\Delta f = -f(\theta, \phi_i)$ is the error and $J = \frac{\delta f(\theta, \phi_i)}{\delta \phi}$ is the Jacobian. The updated solution is given by:

$$\phi_{i+1} = \phi_i + (J^T J + \varepsilon I)^{-1} J^T \Delta f,$$

where $\varepsilon$ is a damping term chosen at each iteration for convergence stability and $I$ is the identity matrix. This iterative algorithm terminates when the correction $\Delta \phi$ becomes sufficiently small.

D.1.2 Open Source Libraries

Below is a list of open source libraries that implement the mathematical functions used in this thesis.
D.2 Plane-based Calibration Simulation

Eigenvalue / Eigenvector Template Numerical Toolkit, National Institute of Standards and Technology.
http://math.nist.gov/tnt/index.html

Levenberg-Marquardt Algorithm Minpack, University of Tennessee, Knoxville and the Oak Ridge National Laboratory.
http://www.netlib.org/minpack

D.2 Plane-based Calibration Simulation

1. Assume a calibration and a position of the tracking system relative to the phantom’s coordinate system.

2. Place the probe at different positions and orientations and find the position of the plane in the B-scans.

   (a) For each set of positions, find the objective function (Equation 2.11) algebraically.

   (b) Find the Hessian algebraically by differentiating the objective function.

   (c) Evaluate the Hessian numerically.

   (d) Evaluate the minimum eigenvalue and check whether it is positive.

D.3 Segmentation Algorithms

D.3.1 Plane-based Phantom

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 & Bolles, 1981). This segmentation algorithm has been implemented in Stradwin\(^1\).

\(^1\)Stradwin: http://mi.eng.cam.ac.uk/~rwp/stradwin/
D.3 Segmentation Algorithms

D.3.2 Z-phantom

1. Find the image scales (Section 4.1.2).
2. Detect the membrane as a plane (Appendix D.3.1).
3. Define the search region relative to the membrane.
4. Detect potential wire locations within the search region.
5. Reject incorrectly detected wires (Appendix B.1).

D.3.3 Cone Phantom

1. Find the image scales (Section 4.1.2).
2. Apply a threshold to the image.
3. Shift each pixel upwards by the temperature correction factor.
4. Apply a Sobel filter to overlapping $3 \times 3$ blocks of the image.
5. Apply a Hough transform on the detected edges to find a circle with a 12mm diameter.

D.3.4 Cambridge Stylus

The segmentation algorithm has been implemented in Stradwin.
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


LINDSETH, F., TANGEN, G.A., LANGØ, T. & BANG, J. (2003b). Probe calibration for freehand 3-D ultrasound. Ultrasound in Medicine & Biology, 29, 1607–1623. 25, 27, 35, 37, 39, 44, 45, 47, 72, 77, 80, 81, 84, 85, 106


134


SAUER, F., KHAMENE, A., BASCLE, B., SCHINUNANG, L., WENZEL, F. &
VOGT, S. (2001). Augmented reality visualization of ultrasound images: sys-
tem description, calibration, and features. In Proceedings of the IEEE and ACM
International Symposium on Augmented Reality, 30–39, IEEE. 16

SAWADA, A., YODA, K., KOKUBO, M., KUNIEDA, T., NAGATA, Y. & HI-
treatment system based on a real time 3D ultrasound image correlation: a
phantom study. Medical Physics, 31, 245–250. 32

volumetric imaging system. I. Transducer design and beam steering. IEEE
Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 38, 100–
108. 1

STATE, A., CHEN, D.T., TECTOR, C., BRANDT, A., CHEN, H., OHBUCHI,
within a pregnant patient. In Proceedings of the IEEE Conference on Visual-
ization, 364–368, IEEE. 12

entropy measure of 3D medical image alignment. Pattern Recognition, 32, 71–
86. 31

ultrasound measurement of large organ volume. Medical Image Analysis, 5,
41–54. 2

TREECE, G.M., PRAGER, R.W., GEE, A.H. & BERMAN, L.H. (2002). Cor-
rection of probe pressure artifacts in freehand 3D ultrasound. Medical Image
Analysis, 6, 199–214. 7, 37, 49

TREECE, G.M., GEE, A.H., PRAGER, R.W., CASH, C.J.C. & BERMAN,
& Biology, 29, 529–546. 5, 6, 8, 35, 37, 40, 42, 44, 45, 46, 47, 48, 53, 54, 55,
56, 57, 77
REFERENCES


REFERENCES

