



Nilesh Hampiholi

Robot Guided Ultrasound Calibration and Volume Scanning

Institute of Medical Technology and Intelligent
Systems
Building E | 21073 Hamburg
www.tuhh.de/mtec



A project paper written at the Institute of Medical Technology and Intelligent Systems
and submitted in partial fulfillment of the requirements for the degree Master of Science.

Author: Nilesh Hampiholi

Title: Robot Guided Ultrasound Calibration and Volume Scanning

Date: April 30, 2022

Supervisors: Johanna Springer (MSc.)
Stefan Gerlach (MSc.)
Prof. Dr.-Ing. Alexander Schlaefer

Referees: Prof. Dr.-Ing. Alexander Schlaefer
Prof. Dr. Zwi Zweitgutachter

Contents

| | |
|---|-----------|
| Abstract | ix |
| 1 Introduction | 1 |
| 1.1 Motivation | 1 |
| 1.2 Project Environment | 2 |
| 1.3 Project Outline | 3 |
| 2 State-of-the-art | 5 |
| 2.1 Ultrasound Imaging | 5 |
| 2.2 Review of Ultrasound Calibration Techniques | 6 |
| 2.3 The Hand Eye Problem | 8 |
| 2.4 Calibration Phantom | 9 |
| 3 Methods and Material | 13 |
| 3.1 Experimental Setup | 14 |
| 3.2 Design of Calibration Phantom | 15 |
| 3.3 Data Acquisition | 17 |
| 3.4 Semantic Segmentation | 19 |
| 3.5 Spatial Calibration | 23 |
| 3.6 Hand Eye Calibration | 28 |
| 3.7 3D US Volume Reconstruction | 30 |
| 4 Results and Discussion | 33 |
| 4.1 US Scanning and Image Segmentation | 33 |
| 4.2 Registration | 35 |
| 4.3 Hand-Eye Calibration | 37 |
| 4.4 Surface Reconstruction using 2D US Images | 38 |
| 5 Conclusion | 41 |
| 5.1 Conclusion | 41 |
| 5.2 Future Work | 42 |
| Bibliography | 43 |
| List of Abbreviations | 47 |
| Declaration | 49 |

List of Figures

| | | |
|------|---|----|
| 1.1 | Project Environment | 3 |
| 2.1 | Types of wire phantom | 9 |
| 2.2 | N-Wire Phantom | 10 |
| 3.1 | Overview | 13 |
| 3.2 | Ultrasound Calibration Steps | 14 |
| 3.3 | Experimental Setup | 14 |
| 3.4 | Constructed N-Wire Phantom | 16 |
| 3.5 | Virtual Point Cloud Phantom | 17 |
| 3.6 | Artifacts | 17 |
| 3.7 | Initial Segmentation Results | 20 |
| 3.9 | Stages of US image segmentation | 21 |
| 3.8 | Flow diagram of segmentation procedure. | 22 |
| 3.10 | Scanning Plane | 25 |
| 3.11 | Localization Algorithm | 26 |
| 3.12 | Coordinate Transformations | 28 |
| 3.13 | Scanned Phantom | 30 |
| 3.14 | Scanned Plane of 3D Phantom | 31 |
| 4.1 | Data Acquisition Samples | 33 |
| 4.2 | Segmentation Results | 34 |
| 4.3 | Segmentation Results | 35 |
| 4.4 | Localization Results | 36 |
| 4.5 | Validation | 38 |
| 4.6 | 3D Phantom | 39 |
| 4.7 | Surface Reconstruction Isometric View | 39 |
| 4.8 | Surface Reconstruction | 40 |

Abstract

Ultrasound is utilized as medical imaging technology across multiple clinical diagnostic and therapeutic contexts primarily due to its availability and non-radioactive nature. Although 3D ultrasound probes have become increasingly accessible, most systems still employ 2D probes as they are highly efficient and affordable. Also, the 2D ultrasound images can be used to reconstruct the 3D ultrasound volume. One of the applications of ultrasound is to enable the practitioner to have a 3D visualization of the tissue, like a hologram displayed in real-time as an overlay over the practitioner's field of view of reality. To accomplish the same, an ultrasound probe is swept across the area of interest, and the trajectory is recorded by the robot. The challenge is that the system records the end-effector pose and not the pose of the scan plane. Through calibration, the transformation matrix is established between the robot end-effector coordinate frame and the ultrasound probe frame. Current ultrasound calibration procedures usually demand freehand motions demanding human intervention, which restricts the application and integration with current robotic systems. This work presents a method for automatic robot-guided ultrasound probe calibration and volume scanning that is also efficient. The proposed method employs a wire phantom with four layers of N-shaped structure stacked one below the other. A UR3 lightweight robot and a Cepasonics ultrasound system are employed to perform the calibration. The overall calibration is accomplished in 500 seconds. An automatic segmentation algorithm identifies the markers with an accuracy of 0.3812 mm and 0.5670 mm in x and z directions. For spatial calibration, a combination of geometric method and iterative closest point algorithm is applied in coordination with a virtual point cloud phantom achieving a mean 3D localization error of 0.7049 mm. Through this procedure, sensor-to-robot calibration is accomplished with a mean translation and rotation error of 4.081 mm and 12.4308°. Validation of the ultrasound probe calibration is achieved by reconstructing the surface of a 3D printed model phantom using 2D ultrasound images. Results demonstrated that the robot-assisted wire phantom calibration significantly improved the calibration repeatability and shorter data acquisition time.

1 Introduction

1.1 Motivation

Visualization in image-guided interventions can be viewed as the interface between the surgeon and the patient. The goal of visualization is to represent the organ and surgical environment in a clear and precise way, supporting the surgeon in navigating through the treatment objective. For this purpose, the visualization should incorporate many signals, particularly tracking and imaging data, into a single representational interface.

Most treatment methods and therapeutic interventions have historically needed direct access to the interior anatomical structures through direct visual assessment to treat the underlying condition. To assist the surgeon, several medical imaging modalities have been employed to identify disorders, design the treatment, and monitor the patient throughout surgery. Significant efforts have been made over the past two decades to decrease the invasiveness of surgical operations, through developments in medical imaging, surgical navigation, visual analytics, and display technologies [1].

Surgical navigation is particularly valuable and helpful in minimally invasive interventions. In contrast to open surgery, the surgeon does not have direct access to the surgical area in minimally invasive interventions. Therefore, image-guided navigation may certainly be regarded as important to accurately approach and treat the target area. Over the last two decades, extensive research led to a boost in the cost-effectiveness and overall treatment outcome of image-guided treatments. In endoscopic and laparoscopic surgeries, Ultrasound (US) is one of the excellent possibilities for imaging. Many studies have focused on the use of US as a real-time guiding tool for spine needle treatments [2]. US provides both a visualization of the anatomy and real-time needle tracking. It is beneficial as there is no harmful radiation exposure to either the patient or the medical practitioner with a relatively short procedure time. Also, US systems are less expensive and more readily available.

The introduction of these procedures into the clinical domain, on the other hand, demands the modification of the limited working settings [3]. Due to the inadequate view of the surgical site, surgeons must maintain precise tool manipulation while relying on imaging data and, in most circumstances, tracked navigation devices. In manual US imaging, the doctor must have the agility and sensitivity to position the US probe for imaging. As a result, improved technology solutions are necessary to automate the imaging procedure while ensuring the safety of the patient. The imaging system and robotic manipulators can be co-registered, achieving improved accuracy and simplifying highly intricate processes when compared to human performance. Much effort has been directed towards the development of actuated manipulators and robotic-assisted imaging to improve the imaging process and offer innovative intervention options [4].

Robots have been utilized in image-guided interventions for accurately placing needles or other tools for biopsy and providing treatment owing to their precision and controlled motion capabilities [3]. Several Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) guided robots have been reported to reduce radiation exposure, and

1 Introduction

steer the instrument in image space. However, in the case of US imaging, a robot might replace the surgeon since the US probe can be installed on the robot arm and traversed along a path to provide scanning motion for 3-D imaging. Typically, US technology produces 2-D images that the surgeon must interpret [5]. The robot not only moves the probe to sweep the region of interest but also records the location to tag each image slice with its position data for 3-D reconstruction. Robots offer controlled motion, tracking, and probe locking features, which gives further benefits over manual handling. Locking, for example, may maintain the probe on the target while a surgeon inserts the needle and conducts a biopsy. Controlled motion leads to more consistent scanning in a given direction and with uniformly dispersed slices. Other prospective advantages are precise revisiting of a particular anatomical region and continuous monitoring of organ movements and deformations. The robot also helps in reducing the line-of-sight difficulties associated with optical tracking and the noise associated with electromagnetic equipment [5].

The US probe alone is not sufficient for the positioning of each US volume in a single reference frame. The calibration of the probe is essential in the significant majority of medical applications in order to record all data in the same mathematical space [6]. In order to precisely reconstruct the 3D US volume and visualize, a homogeneous transformation matrix to transform the location of individual pixels from the US image frame to the robot base frame is needed. This requires the spatial relation between the robot end-effector and the US probe frame, which can be estimated using the hand-eye calibration technique. Calibration is, therefore, a key step in a robot-guided US imaging system and inaccurate calibration leads to failure of the system. There may be serious challenges with 3D reconstruction if there is an inaccurate calibration of the system. To establish the transformation between the robot and image frames in the US calibration, a phantom with a known configuration is frequently necessary [7]. There are different kinds of phantoms or models used for US calibration which include wall, cross-wire, and N-fiducial phantoms. Out of these phantoms, N-phantom is one of the most often used since it is easy to construct and generate several intersecting spots in a single scan. Also, owing to the special properties of its geometry, it is convenient to localize the pose [8].

1.2 Project Environment

This section discusses the project environment concisely and emphasizes the need for calibration. Also, a brief review of Augmented Reality (AR) based interventions is discussed. This research is the module of a project aimed to achieve puncture support using AR with US imaging. The modules of the project are shown in Figure 1.1 and the module concerned with this research is highlighted in blue.

AR is progressively gaining effect in medicine, primarily for instructional applications, since its inception in the 1980s. A commonly proposed use case of AR-head-mounted displays is to visualize data assisting in image-guided interventions. AR headsets, calculate visual information (holograms) and are shown in real-time as an overlay on the user's field of view of reality. The intended benefits of AR-based treatments are reduced workload, fewer problems, and shorter procedures. These advantages are immediately applicable in clinical practice, although they are difficult to validate. Case difficulty and practitioner skill disparities have a major influence on outcomes. Furthermore, the anticipated advantages cannot be deduced directly from AR characteristics. Instead,

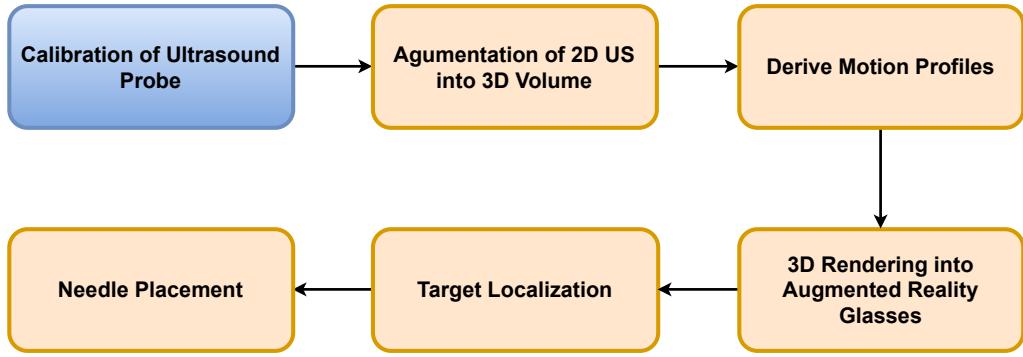


Fig. 1.1: The figure gives a graphical representation of the various modules involved in the project aimed at achieving puncture support using AR with US. Also, the module explored in this research is highlighted.

they are generated by intermediary mechanisms such as accurate eye-hand calibration, increased task attention, or a better spatial understanding of images. Farshad-Amacker et al. [9] experimented whether US with AR technology would accomplish an AR with US view with the US image as hologram. To produce a hologram at a precise spot it is important to provide an accurate specific calibration.

In order to conduct the guided intervention, the region of interest must be precisely represented. This may be accomplished by converting the recorded data from the ultrasound coordinate frame to the robot's coordinate system and from the robot to the AR co-ordinate frame. Ultrasound probe calibration is critical to achieving this transformation accurately.

1.3 Project Outline

This work is structured as follows. Chapter 2 provides an overview of the various imaging modalities available and discusses them briefly. It also highlights advances in ultrasonic calibration procedures. A comparison of several calibration phantoms is presented, as well as a brief overview of hand-eye calibration techniques is discussed.

In Chapter 3 the experimental setup used in this experiment is discussed and it guides through the various steps of ultrasound calibration which include data acquisition, semantic segmentation, registration, solving the hand-eye problem, and reconstruction of 3D ultrasound volume.

In Chapter 4 the different findings achieved during the studies are briefly examined. The suggested methodologies are analyzed with respect to the state of the art. The three vital factors used for comparison include: mean euclidean error, precision, and Fiducial Registration Error (FRE).

And finally, Chapter 5.2 provides a summary of the project results and a discussion about future improvements.

2 State-of-the-art

2.1 Ultrasound Imaging

Imaging is crucial since the surgeon cannot see the anatomical or pathological target directly during minimally invasive procedures. There are various medical imaging modalities in use today, each based on a distinct technology. The image capture technique and image properties change depending on the modality since the physical notion varies. As a consequence, multiple modalities are employed in combination with one another, and there is no ideal modality for image-guided therapies. The most widely utilized modalities include CT, MRI, US and X-Ray.

US imaging utilizes high-frequency sound waves to generate an image. Since US is relatively safe and noninvasive, it has become a useful diagnostic tool in medicine. US is defined as acoustic waves with frequencies beyond perception by the human ear, from about 20 kHz to several hundred MHz. The ultrasound spectrum used in medical devices ranges from 1 MHz to 10 MHz [1]. Higher-frequency ultrasound waves create higher-resolution images, but their shorter wavelength makes them unable to penetrate deeper tissues. Lower-frequency waves have better penetrating power, but because of their longer wavelengths, their resolution is lower. The US images are generated using two basic principles, the piezoelectric effect, and the pulse-echo principle. The transducer consists of ceramic crystals and when an electric current is passed to the transducer, the crystals deform and vibrate. This vibration produces the ultrasound beam. The frequency of the US waves produced is predetermined by the crystals in the transducer. The waves are produced in pulses, not continuously, because the same crystals are used to generate and receive sound waves, and they cannot do both at the same time. In the time between the pulses, the ultrasound beam enters the patient and is bounced or reflected to the transducer. These reflected sound waves, or echoes, cause the crystals in the transducer to deform again and produce an electrical signal that is then converted into an image [10].

The ultrasound transducer can take many different shapes depending on the field of view and arrangements of elements, but the most common is a linear array of acoustic elements capable of acquiring 2D images. The main disadvantage of ultrasound over CT or MRI is its limited field of view. Because ultrasound is a regional imaging modality, it cannot typically provide complete cross-sectional slices of the body. Furthermore, the signal-to-noise ratio is lower than in CT and MRI, and image quality may be operator-dependent. Nonetheless, because ultrasound probes are small and portable, they provide unparalleled access to patients. In terms of the limited field of view, the operator can easily and quickly adjust it by repositioning the probe. In addition, by tracking the US probe, 3D image volumes with a wider field of view can be reconstructed. Because ultrasound provides real-time radiation-free imaging, there is a significant benefit to using it in image-guided interventions.

Brightness Mode Imaging

Brightness Mode (B-Mode) is a two-dimensional ultrasound visual display made up of bright dots that indicate ultrasound echoes. The amplitude of the returning echo signal determines the brightness of each dot. The returning echo's location along the axis is based on the amount of time it takes for the ultrasound wave to be transmitted from the transducer and reflected. Echoes arising from structures in the near field (close to the transducer) take less time than those coming from the far-field (farther away from the transducer) because they travel a shorter distance. In real-time B mode, the transducer sweeps the ultrasound beam through the patient many times a second. With each pass of the ultrasound beam, multiple lines of dots are generated, producing a complete image. With several beam sweeps performed per second, as a result, a real-time image is generated.

2.2 Review of Ultrasound Calibration Techniques

This section presents an overview of the advancements in US calibration approaches. The US probe calibration procedure is crucial in the design of a robot-assisted 3-D ultrasound system enabling image-guided interventions. In all image-guided surgical procedures, registration is a critical step because, distinct data streams, patient images, and tracked instruments are gathered relative to different reference frames, they must be mapped in the reference system. In the case of US, registration is the method used to ascertain the spatial relationship between the images and the calibration phantom. To achieve an accurate hand-eye calibration between the US probe and robot, the accuracy of registration plays a vital role. To achieve registration, different methods are used which include feature or landmark-based and image-based registration. Feature-based approaches use previously derived homologous information, such as point sets from the phantom or surfaces. The rigid connection between the two-point sets can be calculated if a set of matching points is found. The use of feature-based approaches is justified since they are unaffected by the properties of the registered objects. The data must have three or more non co-linear points to find a unique solution. The point-paired registration has the benefit of having many least-squares closed form solutions. The method of registration varies with the phantoms used in the calibration process because it is the computation of the spatial relationship between the image coordinate system and the used phantom coordinate system.

Many researchers have carried out numerous experiments to achieve the particular calibration. A study conducted at Queens University Canada by Thomas Chena and Adrian D. [11] evaluated the calibration system on a series of 10,000 pictures, taken from the Double-N phantom in freehand motion. The images were taken from a broad variety of angles and depths ensuring that the US probe is physically dipped into the phantom. Data were randomly split into 50 groups of 200 photos in each creating 50 separate trials for testing. To prevent bias, the data for validation were isolated from those for calibration. For calibration, images were chosen from the other 120 photos in each set and then given to the calibration system to compute the calibration parameters and the FRE was calculated from the remainder 80 photos in each set for validation. The approach incorporated the known geometry of the N wire phantom to determine the coordinates of the middle wire. However, the approach has some severe drawbacks. The calibration transformation calculated by the approach is not always orthogonal.

Skewed (non-orthogonal) calibration matrices were acquired, as the input images are poor quality and fiducial line detection is unreliable. Also, the FRE is allowed to be isotropic, however, this assumption is not correct, and hence, the resultant calibration matrix is not optimum. The study obtained a mean registration error of sub-millimeter in an axial and lateral direction of 0.14 mm and 0.19 mm respectively.

Guillermo Carbajal and Andras Lasso [12] used the above mention approach and built on it introducing an optimization method that does not assume isotropic FRE and the resulting calibration matrix can be constrained to be orthogonal and to have isotropic spacing. The initial result is optimized by minimizing a cost function based on in-plane error. Another study conducted a comprehensive comparison experiment on the arbitrary wire phantom in 20 independent trials and evaluated the precision, accuracy, and robustness of the calibration procedure [13]. The results of shallow probe studies show that the proposed method improves average calibration precision to 0.896 mm and accuracy to 1.022 mm, compared to the closed-form solution of 0.938 mm and 1.140 mm when N-wire is used.

Another work provides a unique probe spatial calibration technique [14] focusing on the construction of a calibration phantom in combination with a tracking stylus. The calibration phantom is particularly designed to make it simpler to align the stylus tip with the B-scan (Section 2.1) image plane. The spatial position of the stylus tip is monitored in real-time, and the pixel location of the 2-D image is recovered and collected at the same time. The mean calibration error is 0.71 mm, with a standard deviation of 0.21 mm. To determine the distribution of the generated ultrasound data, a research [15] applied a polynomial equation. The greatest value of the polynomial function is applied to assess the phantom's coordinate position. The average error of polynomial equation probe calibration is $17 \times 10^{-5}\%$. In another study by Koo T.K. and Silvia [16] the objective is to employ point-based phantoms and compare an actuator-assisted approach to allow freehand 3D ultrasonic calibration. The actuator-assisted approach was used to calibrate a frequently used cross-wire phantom as well as two custom-made point-based phantoms, each comprising seven co-linear point targets, and the results were compared to the normal freehand cross-wire phantom calibration. The results showed that the actuator-assisted single cross-wire phantom calibration significantly improved calibration reproducibility while providing comparable point reconstruction precision, point reconstruction accuracy, distance reconstruction accuracy, and data acquisition time to freehand cross-wire phantom calibration. Although the accuracy in the cross-wire phantom is 0.84 mm and the special phantom is 1.5 mm, the acquisition time is lowered by 57%.

Maria Chatrasingh and Jackrit Suthakorn [17] presented a unique design for the N-fiducial US calibration phantom that enable automated feature extraction with accuracy equivalent to multiple frame calibration. The Random Sample Consensus (RANSAC) approach is used for feature extraction, and it is capable of estimating both 2D and 3D models. The RANSAC feature extraction method was paired with a closed-form calibration approach to perform automated calibration of 5 N-fiducials stacked in the vertical direction, providing an accuracy of 0.94 mm and a precision of 2.67 mm. Hui Zhang and Filip Banovac and Amy White [18] presented a needle fiducial-based electromagnetic localization approach for calibrating freehand 3D ultrasound as a prerequisite for constructing an intraoperative navigation system. Images of the needle tip are captured, and the locations are recognized either manually or automatically. For calibration and validation, three group locations are recorded, each with nine frames and the largest

2 State-of-the-art

Root Mean Square (RMS) error for calibration with six or more frames were 1.22 mm.

In another research, Ralf Bruder and Florian Gries [19] studied numerous phantoms, point matching, and sub-pixel approaches to offer highly accurate tracking results in 3D US volumes as to the foundation for hand-eye calibration. Without operator participation, automatic calibration of 100 recorded points is completed for the single target. Again, for complex phantom, measurements were obtained at three unique sites, with 50 designated targets. The cables were visible in the ultrasonic volume and traced using a repeated maximum intensity search with an additional distance criterion, so that features in close proximity to previously identified spots are not recognized. Alternatively, a template matching technique is employed with several, established patterns for wires at varying angles to the beam. The collected feature locations are matched against a specified virtual phantom or the initially captured phantom using the Iterative Closest Point (ICP) method. The accuracy increase for complicated phantom calibration owing to enhanced target localization. With standard resolution, both calibration approaches yield acceptable results, with RMS errors of less than 2 mm.

2.3 The Hand Eye Problem

Many robot-assisted medical applications that require accurate computation of the relationship between the robot's coordinate system and the coordinate system of the imaging system. When it comes to a machine communicating with objects in its environment either observed by a camera, tracking system, or other localization devices, hand-eye, eye-in-hand, and robot to world calibration are all valid topics to discuss. In general, robot to world calibration involves solving a matrix equation of the form,

$$MX = YN \quad (2.1)$$

Where the matrices M and N are known, but the matrices X and Y are unknown. M is typically a robot's pose matrix and can be computed from the encoder values and N is the position and orientation of a calibration object with respect to a tracking device which can be estimated by various registration techniques. As a result, the matrix X is the end-effector to sensor transformation, and the matrix Y is the robot to world transformation. The methods to solve the equation of the form 2.1 there are three approaches: separable closed-form solutions, simultaneous closed-form solutions, and iterative closed-form solutions. Solving the orientation component separately from the positional component is the separable closed-form approach, the simultaneous closed-form solutions arise from simultaneously solving the orientation component and the positional component, while the iterative solutions arise from solving both the orientation component and positional component iteratively using optimization techniques. The study by Mili Shah and Roger D. Eastman [20] offers an in-depth comparison of the three methods. Typically, this problem is approached by obtaining measurements at various stations $(M, N)_i$ and eliminating one of the two unknown matrices to arrive at the simplified equation [2.2].

$$M_j^{-1} M_i X = X N_j^{-1} M_i \quad (2.2)$$

The first solutions to this problem were independently proposed by Shiu and Ahmad [21] and Tsai and Lenz [22]. The techniques make use of matrix algebra and the particular

features of homogeneous matrices to determine the rotational and translational elements of the matrix X independently. A unique solution is achieved given that the axis of rotation of M_i are neither parallel nor antiparallel to one another. Other works that compute these components individually are based on quaternion algebra, screw motion analysis, euclidean group properties, and the solution of nonlinear equations. A research [23] presents a promising linear decomposition strategy that incorporates Singular Value Decomposition (SVD) followed by estimation of the optimal rotation matrix, and the least-squares equation to solve for the rotation matrix of X. Zhiqiang Zhang and Lin Zhang [24] demonstrates a computationally efficient iterative technique uses a dual quaternion to represent the rigid transformation, and a two-step iterative approach is developed to obtain the real and dual portions of the dual quaternion simultaneously, and hence the estimate of rotation and translation of the transformation.

There are various additional techniques to solving this problem. In this project, the tool or flange calibration method [25], which employs QR24 factorization to solve the equations, is employed. This technique was developed targeting medical applications. The approach uses a simultaneous computation tool to flange and a robot to world transformations by calculating a solution to the matrix equation 2.1. It is computed using a least-squares approach. As real robots and localization are all afflicted by errors, the least-squares approach allows for non-orthogonal matrices, partially compensating for imperfect calibration of the robot or localization device [25].

2.4 Calibration Phantom

To make the calibration procedure easier, methods for ultrasonic probe calibration have been studied extensively. Only a few research groups have completed probe calibration without the use of calibration phantoms, whereas the majority of researchers create different phantoms with known geometry constraints to serve as fiducials for probe calibration. Numerous studies have been done to develop and analyze distinctive forms of phantoms as the accuracy of anatomical localization is influenced by the calibration system (Figure 2.1). Not only do the kinds of phantom influence system accuracy, but so do closed-form solutions or calibrating techniques. The phantoms are characterized in the following ways: single point targets, cross-wire phantoms, multiple point targets, 2D form alignment phantoms, three-wire phantoms, fiducial phantoms, and wall phantoms [26].

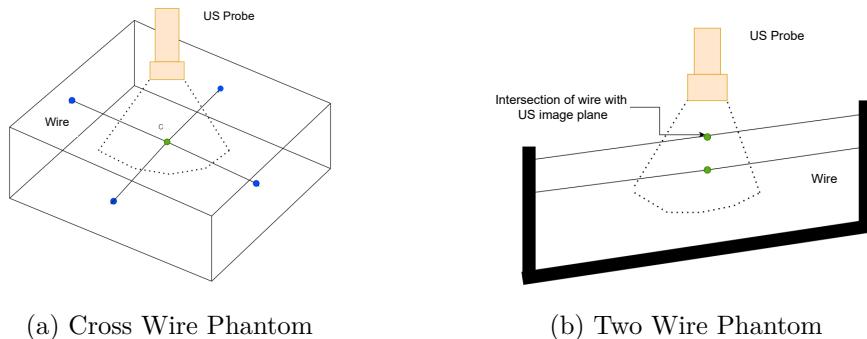


Fig. 2.1: The figure illustrates the different types of phantom used for US probe calibration.

2 State-of-the-art

One of the earlier approaches for probe calibration was to use a point phantom. A spherical bead-like structure or a pair of cross wire phantoms can be used to construct it [13] [27]. While constructing a point phantom is simple, images generated by scanning a point phantom would have to be manually segmented, making the calibration procedure time-consuming and laborious. Figure 2.1a depicts a traditional calibration phantom that may be positioned inside any tank and used to scan from the top. The cross-wire system scans the junction of two crossing wires or a tiny item such as a bead or pinhead. The target is suspended in a water bath, scanned from distinct viewpoints, and a crossing point (or pinhead) is manually identified in each image. The accuracy of this approach is heavily reliant on the manual detection accuracy of the images. Each image has only one data point, which is incredibly difficult to manually determine. Since it is difficult to distinguish a single target from imaging artifacts, the procedure is challenging to automate [28].

In another study, [15], the phantom was a horizontally positioned 1 mm copper wire which connected the two sides of an acrylic frame. The two parallel wires are spaced 1 cm and 5 cm apart in the x and y direction respectively. The wire phantom was then suspended 5 cm below the surface of the water. The intensity of the pixels is then used to segment the acquired images. Each image's peak point is computed using a polynomial equation and the peak position coordinate is derived.

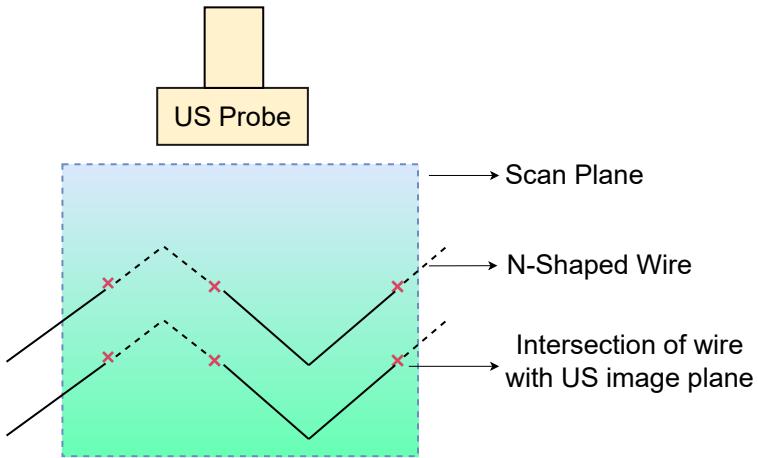


Fig. 2.2: The figure illustrates an N-fiducial and scanning plane.

The three-wire phantom scans three wires that are immersed in water and interlocked in such a manner that they construct an orthogonal coordinate system. The data points, like those in the cross-wire technique, must be manually labeled. The precision of this procedure is also determined by whether the wires are properly orthogonal and straight. The benefit of this approach over the Cross-wire method is that scanning the length of the wire is easier than keeping the image focused at a single place while scanning from different directions [27]. Probe calibration can alternatively be accomplished using a plane phantom. As the name suggests, it features a planar surface immersed in the water bath. This plane is viewed from various directions in the US images, each of which contains lines that lay in the plane. One problem with this method is that

specular reflection causes low returning intensity at oblique scan angles and the plane reflects a large amount of ultrasonic energy away, making automated segmentation problematic [28].

The N-wire phantom is among the few phantom designs that can achieve complete transformation in a single scan. It is constructed by connecting three wires in a plane to obtain an N or Z shape. One 3D position is determined by the intersection of the N-fiducial with the scan plane creating three high-intensity pixel clusters in the ultrasound image. The pose of a scanned plane in space could be established whenever at least three non co-linear points are obtained [8, 17]. Although determining a plane requires at least three non co-linear points, multiple N-fiducials are positioned in an un-patterned multitude of designs as shown in Figure 2.2. In comparison to the point or plane phantom, which requires at least 550 photographs to compute satisfactory accuracy, the N-wire phantom only requires 6–30 images [29]. In the N-wire phantom, just the centerline is utilized to compute the calibration transform. However, more precision can be obtained if all lines are employed. The primary downside of adopting this phantom is the difficulty in segmenting pixels that indicate wire-plane intersection in US images.

3 Methods and Material

This chapter includes an in-depth summary of the experimental setup involved throughout the US calibration process. The methodologies adopted for probe calibration, image segmentation, and fiducial registration methods are also widely researched. Figure 3.1 illustrates all the transformation computed in the calibration. The main goal is to achieve the transformation from robot end-effector to the US probe coordinate frames. This can be computed by solving the hand-eye problem which in turn requires the transformation from phantom to the US probe frames. This is obtained by target registration, using the acquired images and the known geometry of the phantom.

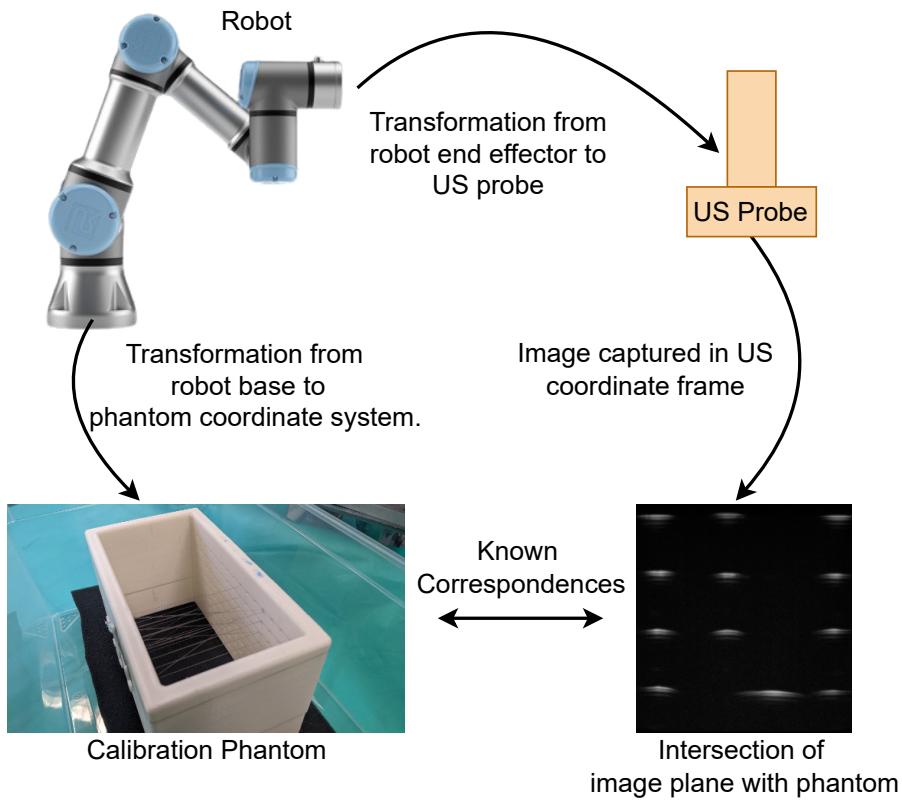


Fig. 3.1: Illustration of the transformations computed in the calibration process.

The calibration procedure consists of the following steps: design of the phantom, data acquisition, semantic segmentation, registration, and solving the hand-eye problem. Figure 3.2 shows a flow diagram of the various steps involved. Each step is explored in detail in this chapter and the results are presented in Chapter 4.

3 Methods and Material

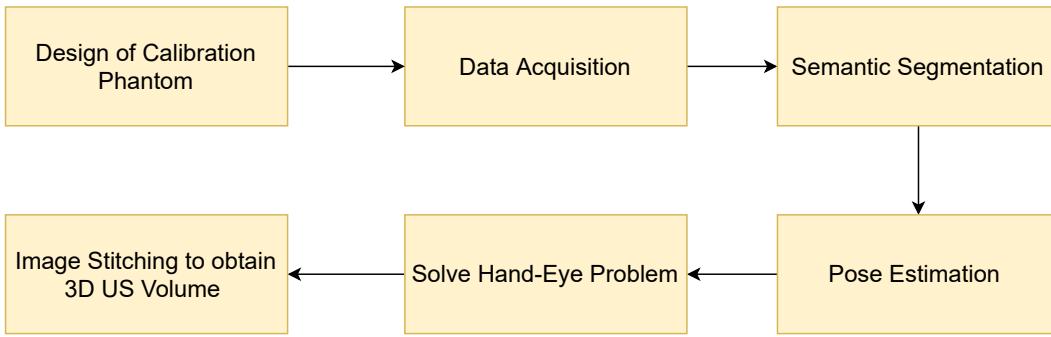


Fig. 3.2: The figure illustrates a flow diagram of the various steps involved in the US probe calibration.

3.1 Experimental Setup

This section describes the experimental setup used in the US probe calibration procedure. The experimental setup consists of a US probe (linear probe L5-14, Cephasomics) that is mounted at the end-effector of the UR3 collaborative robot using a specially designed mount. The setup is visualized in the Figure 3.3. There are other varieties of calibration phantoms available as discussed in Section 2.4, however, the N-wire phantom is used for calibration here. The wire phantom is specially designed for this purpose and placed inside a plastic container with a foam base. The construction of the phantom is further discussed in Section 3.2.

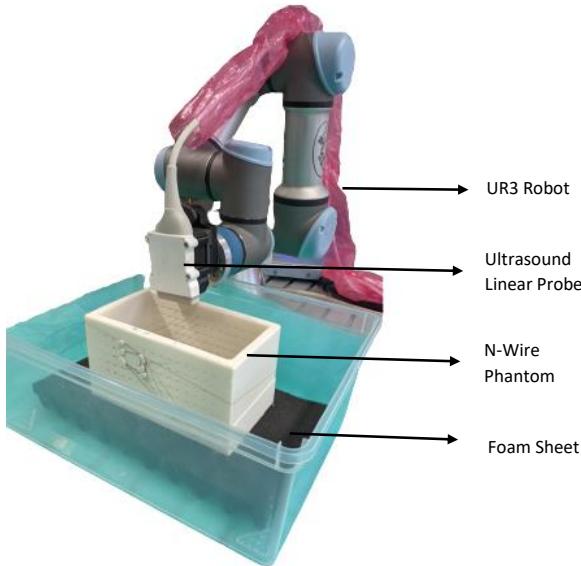


Fig. 3.3: The figure shows the experimental setup used in US probe calibration.

The UR3 collaborative robot is a lightweight table-top 6-DOF robot that is ideal for light assembly chores and automated workstation applications. The robot weighs about

11 kg and has a load capacity of 3 kg. The robot has position repeatability of roughly ± 0.03 mm making it an adequate choice for medical applications. The cephasonics US platform is high-performance hardware that provides easy US application development environment through its compact design and comprehensive user interface with built-in libraries for faster development.

The size and form of a probe's face are often used to define it. The probe selection is of great importance for obtaining acceptable images. There are occasions when more than one probe is appropriate for a given inspection. The three major types of probes are: linear, curvilinear, and phased array probes. Linear probes have a high frequency and are superior for imaging superficial tissues and arteries. Curvilinear probes have a larger footprint and lower frequency, or they can have a smaller footprint (wider field of view) and higher frequency. A phased array probe produces a visual from an electronically manipulated beam in a close array, yielding an image that comes from a point and is ideal for reaching between ribs in cardiac ultrasound.

The linear array probe is intended for high-resolution imaging. Within a flat head, the crystals are positioned in a linear pattern to create sound waves in a straight path. The resulting image is rectangular in form. This probe operates at higher frequencies (5–14 MHz), resulting in superior resolution and less penetration. As a result, the probe features make it an adequate choice for performing US-guided treatments. The probe used in this project has 128 piezoelectric crystals placed at a distance from 0.3 mm from the centerline of each element and operates in the given frequency range. The elements are positioned such that acoustics waves generated meet at the focus point. The characteristics of the probe are given in Table 3.1.

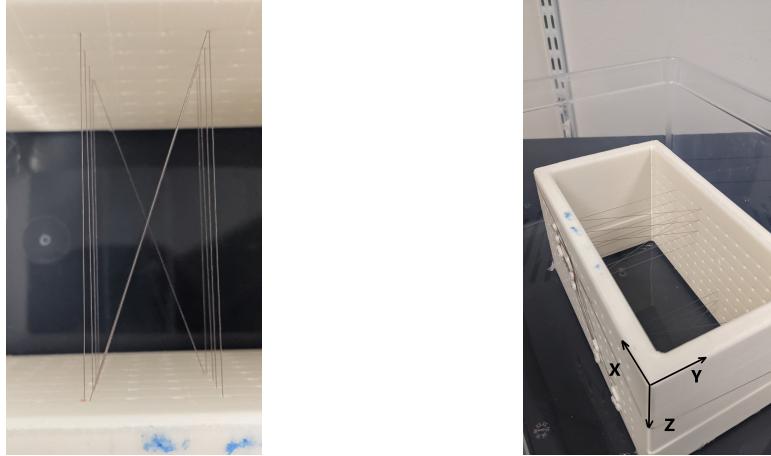
Tab. 3.1: Characteristics of US probe

| Feature | Value |
|------------------|--------------------|
| Type | Linear Array Probe |
| Center Frequency | 5-14 MHz |
| Elements | 128 |
| Pitch | 0.3 mm |
| Focus | 16 mm |

3.2 Design of Calibration Phantom

The background of the various kinds of phantoms and their importance is discussed in Section 2.4. The N-wire phantom is among the few phantom designs that can achieve complete transformation in a single scan. It is very efficient and enables a simultaneous target localization as it generates three co-linear marker points. This section gives a detailed construction process of the wire phantom used in this research.

3 Methods and Material



(a) Arrangement of 4 N-wire fiducials (b) Phantom printed using 3D Printer

Fig. 3.4: Figures show the 3D printed phantom and N-fiducials stacked one below other.

The N-wire phantom is designed in Fusion 360 and printed using a 3D printer. It is a void cuboid with measurements $15\text{ cm} \times 10\text{ cm} \times 10\text{ cm}$. The front of the phantom is cut with 2 mm diameter holes extending to the depth of the phantom. The holes are 10 mm apart in the x and z directions as shown in Figure 3.4b. A single strand of long fishing wire is inserted through the perforations in the 3D printed construction to create an N or Z-shaped fiducial of the size 8 cm and 3 cm along length and width respectively. The traditional design consisted of two levels, with the N-shaped structures stacked one on top of the other. To attain the best results, we employed four layers of fiducials. Three layers are oriented in the same way, while the fourth is oriented in the opposite direction. This aids in estimating rotation along the z -axis with respect to the phantom coordinate frame. The phantom is then immersed in a water bath with a foam base ensuring that all of the fiducials are entirely submerged in water, effectively allowing scanning to be performed in circumstances similar to those seen in the human body. This is discussed in detail in Section 3.3. The 3D printed phantom is shown in Figure 3.4. Along with the printed phantom, a replica representing the N-fiducial is generated using MATLAB. This virtual phantom is merely a point cloud that is used to easily obtain the matching correspondences and compute the registration. The virtual phantom is depicted in Figure 3.5.

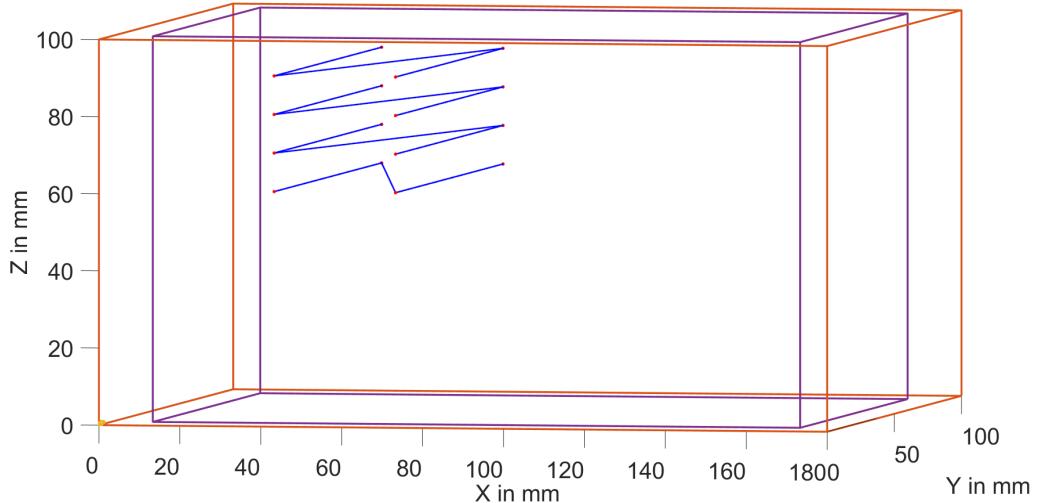


Fig. 3.5: Figure shows a virtual point cloud phantom. It is a replica of the constructed phantom.

3.3 Data Acquisition

Data acquisition is a critical aspect of the experiment since the data has to be devoid of artifacts. The segmentation of US images is challenging due to aberrations such as speckle, attenuation, and shadows. While traditional segmentation approaches such as region growing and threshold-based algorithms have the benefit of being simple and fast, they are often insufficient to segment US images, thus effort must be taken to minimize the mentioned disruptions. The results of the initial scanning are depicted in Figure 3.6.

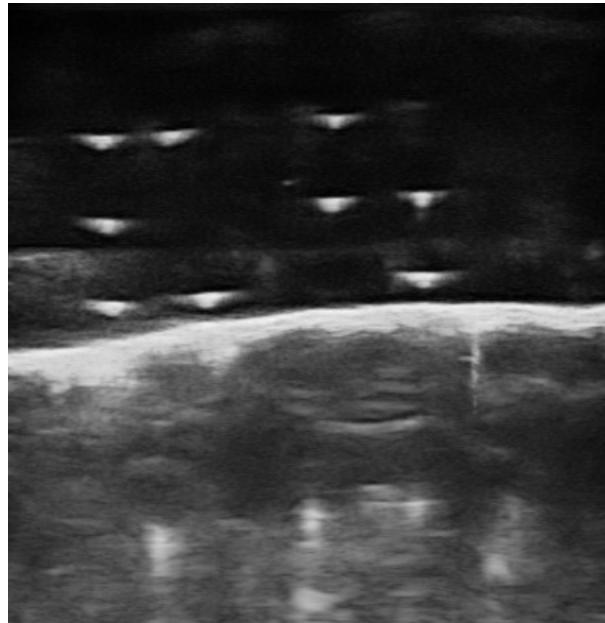


Fig. 3.6: Figure depicts the various artifacts generated during data acquisition.

3 Methods and Material

Artifacts

As it can be seen in Figure 3.6 there is large number of artifacts. These are caused due to small suspended particles in the water, gas or air bubbles, and highly reflective surfaces present in the structure. An ultrasonic beam causes reverberation by repeatedly bouncing back and forth between two highly reflecting contacts or between the transducer and a powerful reflector. Each time sound wave pulse returns to the transducer after hitting a reflecting surface such as gas, bone, or metal, particularly if this interface is near the transducer. The probe partially captures this echo, resulting in a hyperechoic line.

Comet tail is another sort of distortion. Comet tail artifacts can be seen with gas bubbles in the intestinal loop (in our case, air bubbles inside the water bath), and form thin layers separated by liquid. The waves rebound between the layers, resulting in many echoes that return to the probe at irregular intervals, leaving a trail of closely spaced, discrete, very bright, small echoes that resemble a comet's tail. Metallic pellets, surgical clips, or a biopsy needle might also generate this distortion. Shadowing is another type of artifact that is caused by virtually total absorption or reflection of the sound beam by a high attenuation structure. When sound is reflected (as in the case of a soft tissue-gas interface), the region underneath the structure appears inhomogeneous due to repeated reflections or reverberations. When a large percentage of the ultrasonic beam is absorbed and reverberations are missing, this results in weakly echoic or anechoic shadowing.

All of the artifacts mentioned may be seen forming in Figure 3.6. The little spots near the top of the picture are caused by suspended particles in the water. We can also witness the creation of a hyperechoic line following the second layer of N-fiducial. Also, reflection from the phantom's base is generating a white beam in the center of the image. There are also a few bright white specks in the image caused mostly by air bubbles.

Acquisition Procedure

The phantom was initially placed in a container filled with lukewarm water as a coupling media during actuator-assisted calibration, as shown in Figure 3.3. Using a specially developed mounting tool, the ultrasound probe is attached to the robot's gripper. The robot's starting position is chosen so that the probe is exactly on top of the calibration phantom, about 10 mm from the fiducials. The probe is set up in such a way that the image plane is perpendicular to the target locations. It is also necessary to ensure that the transducer is submerged in water during the scanning operation. To guarantee image quality, the probe is coated with ultrasonic gel before attachment.

Several procedures are performed to reduce the artifacts. It is crucial to use clean water so that it contains fewer suspended particles. Air bubbles in the water can be eliminated manually. A foam sheet is placed beneath the phantom to conceal the water bath's translucent bottom face. This is because foam has properties to absorb unwanted sounds [30]. This covers the reflective surface and rejects any external noise entering the system. As a result, shadowing and repetition artifacts are reduced.

When the setup is finished, the robot is programmed to position the probe such that all the wires are visible in the US image. Data is collected from various angles by rotating the probe along the x , y , and z axis with respect to the phantom coordinate frame. At each step, data is collected using B-mode imaging (Section 2.1), and the pose of the robot is recorded. Table 3.2 shows the motion ranges of the US probe. The motions

ranges are summarized with respect to phantom coordinate frames as shown in Figure 3.4b.

Tab. 3.2: Motion range for robot assisted US scanning.

| Direction | Translation (mm) | Rotation (degrees) |
|-----------|---------------------|-----------------------|
| X | ± 10 | ± 15 |
| Y | ± 10 | ± 15 |
| Z | ± 3 | ± 15 |

The coupling medium used throughout the study has a considerable influence on image quality. Sound waves pass through the internal organs at a speed of 1540 m/s. And studies have been conducted to match this speed in other coupling fluids. Water, a water-glycerol mixture, salt solution, and gelatin are all examples of coupling mediums. The speed of sound in water is roughly 1480 m/s, which may be modified to match the speed of sound in tissues by varying the water's temperature. Since the experiments are carried out multiple times and it is difficult to monitor the same conditions every single time, this experiment uses clear water at room temperature and adjusts the parameters in the US system to match the conditions.

3.4 Semantic Segmentation

In this part, segmentation techniques for extracting marker points from images scanned from US probe are presented. This stage is a processing technique used to analyze and group pixels based on their features. Image segmentation is one of the most significant steps of image processing and analysis, to partition a Region of Interest (ROI) based on particular characteristics and attributes.

Due to distortions such as speckles, attenuation, and shadows, segmentation in US images is challenging. The traditional segmentation approaches such as region expanding and threshold-based algorithms are simple and inexpensive, but they are often insufficient for segmenting US images. More robust segmentation methods based on optimization of energy functions and Convolutional Neural Networks (CNN) based segmentation are also available [31]. As a wire phantom of known geometry is scanned and the fishing wire is easily seen in the US images, traditional segmentation algorithms are sufficient. The outcomes of preliminary segmentation using a basic thresholding approach is shown in Figure 3.7.

Filters

Thresholding is the most fundamental segmentation approach, consisting of consistently establishing an intensity value as a threshold, such that objects are differentiated based on their level of intensity. If the image has an 8-bit resolution and the intensity value at any pixel is higher than the threshold, it is assigned a value of 255 (foreground), otherwise, its value is set to 0 (background), resulting in a binary image. This approach does not take pixel position into account, simply intensity. The approach works well for images containing distinct areas. However, because characteristics in ultrasound images are difficult to differentiate, determining a threshold is challenging, resulting in poor

3 Methods and Material

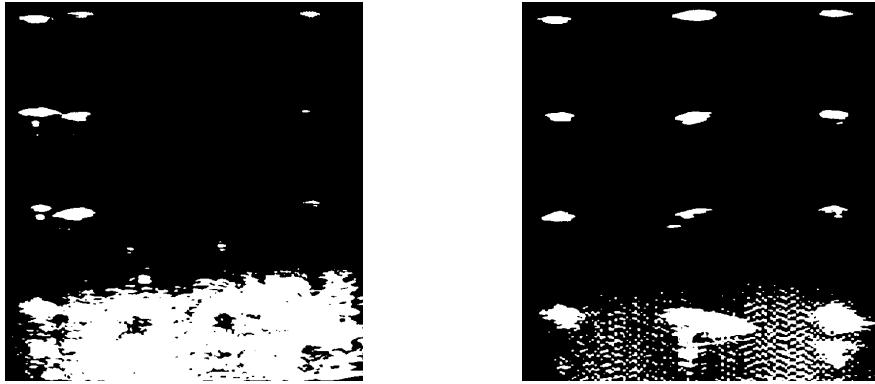


Fig. 3.7: Figures depict the outcomes of the direct thresholding.

segmentation [32]. Median Filter, because of its superior efficiency for certain noise types such as Gaussian, random, and salt and pepper, is one of the most well-known order-statistic filters. The median filter replaces the center pixel of a $M \times M$ neighborhood with the median value of the associated window. It should be highlighted that noise pixels are thought to be quite distinct from the median. Using this concept, a median filter may be used to minimize noise concerns [32].

Dilation is a morphological procedure that widens luminous regions in an image. For that purpose, the new gray value of each pixel is defined as the sum of the old gray values of all pixels in the structuring element centered on this pixel. It is also possible to iterate on dilation filtering. Multiple stages of dilation with a given structuring element are comparable to a single step of dilation with a bigger structuring element. Similarly, erosion expands dark structures in an image by substituting the gray value of each pixel with the minimum of the old gray values of all pixels within the structuring element. The structuring element used to analyze the input image is an important aspect of morphological dilation and erosion operations. A structuring element is a matrix that identifies the pixel in the image being processed and determines the neighborhood in which each pixel is processed. Typically, it is chosen such that it is the same size and form as the items in the input image. For example, a linear structuring element is used to locate the lines in an image.

Automatic Segmentation Algorithm

The first findings reveal that the basic thresholding strategy is unable to segment the marker satisfactorily. Artifacts such as comet tails, speckle, and shadows end up causing thresholding to fail. To mitigate artifacts and segregate the identifiers, a two-step automated segmentation system is introduced and discussed briefly below. The two key steps are as follows:

1. Segmentation operations.
2. Centriod extraction.

The two segmentation phases are repeated in a loop. If the clusters identified at the end of the processes are not satisfactory, the raw image is automatically re-segmented by repeating the two phases while modifying the operation parameters namely the

thresholding factor, kernel size, and structural element size to produce satisfactory results. The two phases are discussed in further detail below. If the clusters are still not detected after various combinations of parameters, the image is flagged and noted. The segmentation algorithm's flow diagram is illustrated in Figure 3.8.

In segmentation operations, the US image is initially processed using a median filter. The median filter brightens the N-fiducial markers, making it simpler to identify them. Different filter sizes are investigated, and a size of 7×7 yields adequate outcomes 3.9b. The filtered image is then thresholded, and the bright spots are readily distinguished from the foreground, as seen in Figure 3.9c. All the spots that are brightened as an outcome of the median filter, including speckles and shadows are persisted after thresholding.

To eliminate distortions, we implement an operation that removes any pixel clusters that are fewer than 10 pixels or larger than 200 pixels in size by resetting the pixel intensity to 0. As a consequence, we obtain a distortion-free image that can be seen in Figure 3.9d. However, there are certain distortions in the immediate neighborhood of the markers caused by the comet tail effect that might have been left out are corrected using morphological techniques. As a disk-shaped structuring element is employed, first expand the marker to create a disk form, and then compress or dilate to restore the initial size of the markers, reducing tiny distortions surrounding the markers as seen in Figures 3.9e and 3.9f.

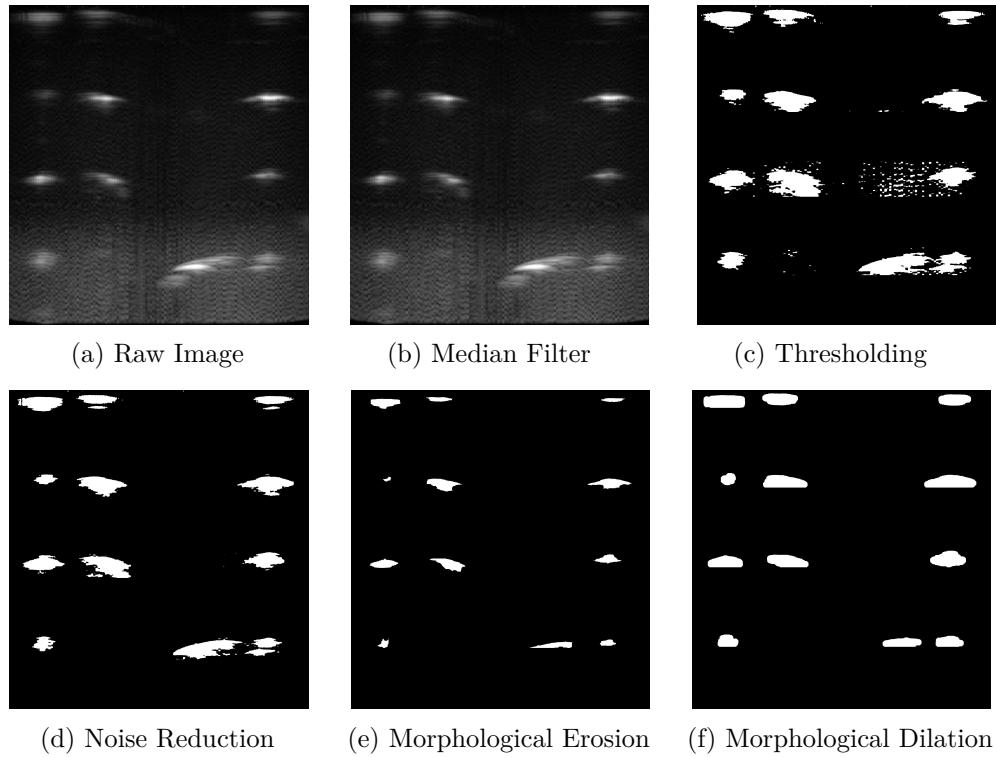


Fig. 3.9: Stages of US image segmentation

The outcome of segmentation operations is a binary image and straightforward clustering produces just two groups of pixel intensities of 0 and 1. The markers must be labeled or identified as different units in order to obtain all 12 marker points. To label them, they must first be located, which is accomplished by detecting the linked components

3 Methods and Material

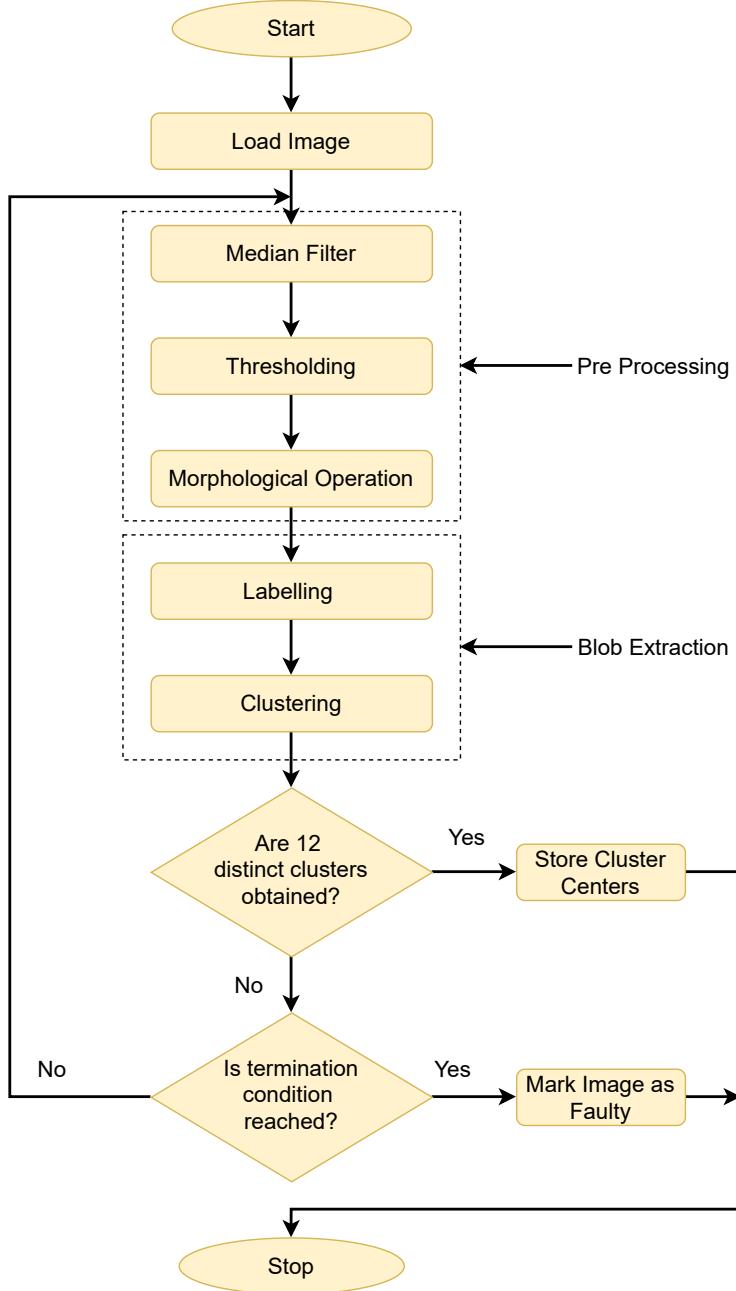


Fig. 3.8: Flow diagram of segmentation procedure.

across the image and marking the connected pixels in the neighborhood with a value distinct from disconnected pixels. The component connection level for 2D pictures might be 4 or 8. Pixels are considered to be linked if their edges meet when the degree of connectivity is set to 4. That is if two adjacent pixels are linked along the horizontal or vertical axis, they are part of the same cluster. If the degree of connectivity is 8, pixels are linked if their edges or corners are in contact and also horizontal and vertical axis. We set the level of interconnection to be 8 in the current investigation. Once all of the marker points have been labeled, the k-means clustering algorithm can readily categorize the clusters. The pixel coordinates of the cluster centers are retrieved and recorded. The outcomes of segmentation are presented in Chapter 4.

The euclidean distance between markers of the left strand and right strand is 30 mm and the distance between adjacent layers of N-fiducials is 10 mm as discussed in Section 3.2. Given the distance between the markers in an image, the segmentation error is computed as follows,

$$X_{image} = \frac{\sum_{m=1}^M D(a(x, z), b(x, z)))}{M} \quad (3.1)$$

$$Z_{image} = \frac{\sum_{m=1}^M D(a(x, z), c(x, z)))}{M} \quad (3.2)$$

where X_{image} and Z_{image} are the mean euclidean distance between marker points $a(x, z)$, $b(x, z)$ and $a(x, z)$, $c(x, z)$ respectively. M is the number of marker points used for computation and D is defined as,

$$D(p(x, z), q(x, z)) = \sqrt{(q_x - p_x)^2 + (q_z - p_z)^2} \quad (3.3)$$

The process is repeated for N ($N = 130$) images acquired in unique poses. The mean deviation from reference is computed using Equation 3.4 and 3.5.

$$X_{mean} = 30 - \frac{\sum_{n=1}^N X_{image_n}}{N} \quad (3.4)$$

$$Z_{mean} = 10 - \frac{\sum_{n=1}^N Z_{image_n}}{N} \quad (3.5)$$

where X_{mean} and Z_{mean} are the mean deviation from ideal values in X and Z directions. The mean segmentation error also incorporates the construction and scaling errors. The mean segmentation error for two independent trials is presented in Chapter 4.

3.5 Spatial Calibration

Tracking 2D US images in three dimensions entails determining the position and orientation of each 2D image in reference to a certain 3D coordinate system. Here, the robot base frame is the primary coordinate system. To compute the hand-eye calibration, the transformation between the probe and phantom frames is to be calculated as discussed in Chapter 2. To recognize the transition between the generated 2D images in the probe frame and the phantom coordinate frame, registration needs to be performed. The goal of registration is to determine the spatial connection between a set of characteristics (homologous points) that can be detected in both image space and phantom space with

3 Methods and Material

established geometric constraints. The construction of the phantom and the methods to extract the markers are discussed in Sections 3.2 and 3.4.

The phantom is composed of the horizontally shifted layers of N-fiducials parallel to the x-y plane of the phantom coordinate system. The intersection of the N-fiducials with the US image plane results in 12 dots in the image as shown in Figure 4.1. The centers of the dots in the images were automatically determined using semantic segmentation, as discussed in Section 3.4. Figure 3.10 depicts the intersection of a US plane with a single N-fiducial. The 3D visualization of the intersection is shown in Figure 3.10b. The registration problem is to map the segmented markers to the N-wire phantom and estimate the transformation ${}^U H_P$ from the phantom coordinate system to the US probe coordinate system.

On the dashed line, coordinates of the point M in the phantom space can be determined based on the similar triangles BLM and MRC as seen in Figure 3.10a.

$$X_M = X_B + \alpha (X_C - X_B) \quad (3.6)$$

$$Y_M = Y_B + \alpha (Y_C - Y_B) \quad (3.7)$$

$$Z_M = Z_B \quad (3.8)$$

Where, X_M, Y_M and Z_M are the x, y and z coordinates of point M in phantom coordinate frame and α can be calculated as shown in Equation 3.9 and LM, LR are obtained as shown in the Figure 3.10a. X_B, X_C and Y_B, Y_C are the x and y coordinates of the start and end of the middle strand in the N-fiducial. The z coordinate of points M and B are identical as they lie in the same x-y plane.

$$\alpha = \frac{LM}{LR} \quad (3.9)$$

The LM/LR ratio was calculated using the positions of the three co-linear dots for every layer of the N-fiducial in the US image (Figure 3.10c). The LM, LR are obtained by calculating the euclidean distance between the coordinates of the cluster centroids obtained from the US images using the Equation of the form 3.10.

$$D(a, b) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.10)$$

where $D(a, b)$ is the distance between the points a, b and x_1, x_2 . The points y_1, y_2 are the x and y coordinates of the points a and b .

The initial estimate of the image scale is previously acquired from the US machine. The scaling is then slightly adjusted to match the fiducials. The known phantom geometry provides the coordinates of the coordinates of vertices in phantom space (X_B, Y_B) and (X_C, Y_C). As each of the three dots is co-linear, there are several solutions for the phantom-to-image transformation ${}^U H_P$, i.e., different solutions for the coordinates of points E and Z in phantom space. As a result, a combination of the geometric technique and the ICP algorithm (Section 3.5) is used to arrive at a global solution.

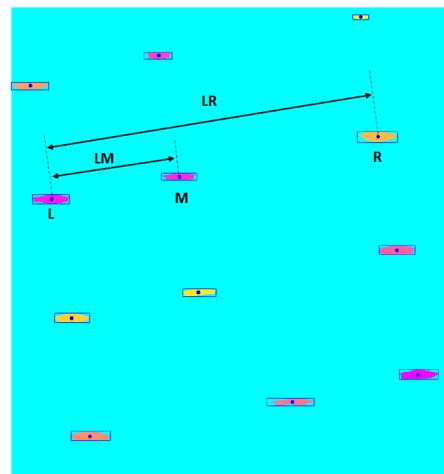
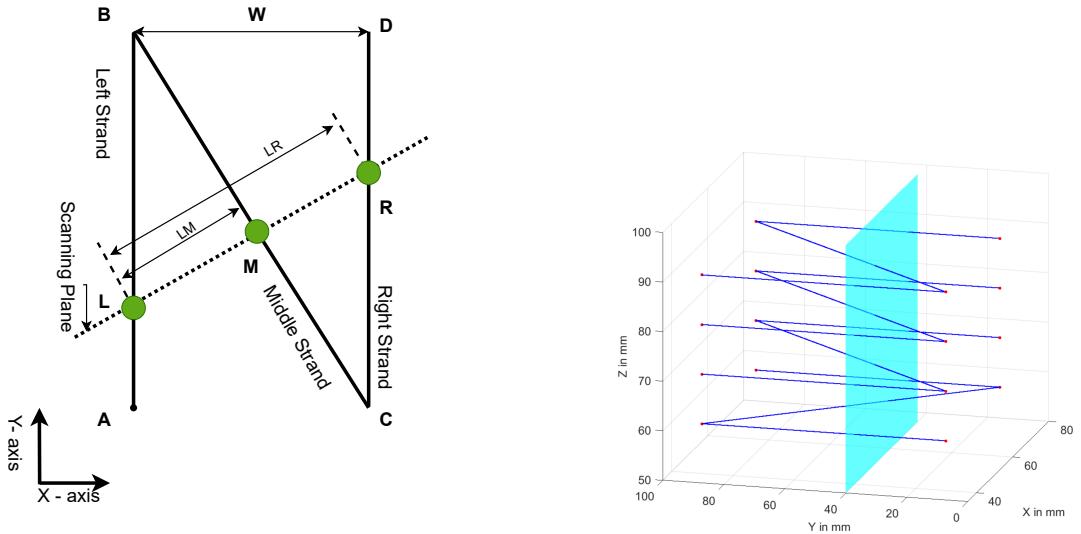


Fig. 3.10: An example for the scanning plane.

3 Methods and Material

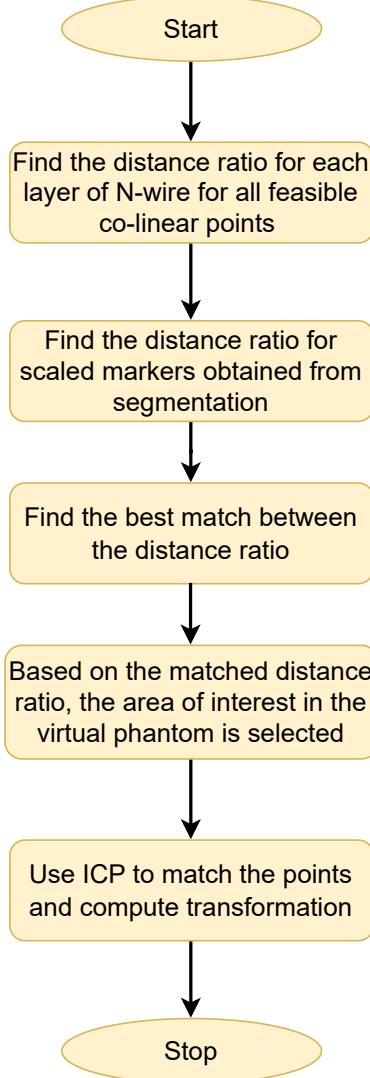


Fig. 3.11: Flow diagram of localization algorithm.

The ICP technique can also be used to solve registration directly. ICP method approaches this problem as a point-to-point registration problem between the image-derived markers and the virtual point cloud phantom shown in Figure 3.5. It matches the corresponding points and accomplishes the required transformation. The algorithm is discussed in greater detail in Section 3.5. Using this method, however, results in a local minimum. To address this issue, an initial estimation must be provided around the unknown global minima. The initial transformation matrix used in this work is shown in Equation 3.11. This aligns the markers obtained from the image to the center of the wire phantom as seen in Figure 4.3a.

$$H_{initial} = \begin{bmatrix} 1 & 0 & 0 & 40 \\ 0 & -1 & 0 & 20 \\ 0 & 0 & -1 & 90 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.11)$$

The problem is solved by exploiting the well-known geometry of the N-wire phantom.

Rather than identifying the middle strand's coordinates, the distance ratio LM/LR is determined for every strand. Distance ratios are calculated for all possible co-linear points in the virtual phantom's left, middle, and right strands of all four layers. The region of interest in the phantom is determined by comparing the ratios acquired from the markers in the US image to the ratios obtained from the virtual phantom and selecting the region with the least error between the ratios. The approximate 3D location of the image in phantom coordinates is the region of interest. Local minima problems can be alleviated by utilizing a tiny proportion of the virtual phantom around the stated region of interest. The ICP algorithm identifies the matching points and computes the transformation. The procedure is summarized in the flow chart as shown in Figure 3.11. The outcomes of localization are discussed in Chapter 4.

The FRE is computed in terms of euclidean distance. For every marker, the euclidean distance to its corresponding match in the virtual phantom is computed using Equation 3.12.

$$FRE = \frac{\sum_{m=1}^M D(p(x, y, z)_m, q(x, y, z)_m)}{M} \quad (3.12)$$

Where D is the euclidean distance which is defined as,

$$D(p(x, y, z), q(x, y, z)) = \sqrt{(q_{x_m} - p_{x_m})^2 + (q_{y_m} - p_{y_m})^2 + (q_{z_m} - p_{z_m})^2} \quad (3.13)$$

Where $p(x, y, z)$ and $q(x, y, z)$ are markers and corresponding matched points in the virtual phantom and M is the number of markers. For N ($N = 130$) images acquired the mean FRE is computed using Equation 3.14

$$Mean\ FRE = \frac{\sum_{n=1}^N FRE_n}{N} \quad (3.14)$$

The mean FRE is obtained for two independent trials is presented in Chapter 4.

Iterative Closest Point

ICP is an iterative technique in which the first stage is to build point correspondences and the second step is to compute a transformation based on these matches. The final registration result is composed of incremental modifications. An ICP framework is made up of the following elements given two non-corresponding point sets [33]:

1. Find corresponding points, e.g. by matching the nearest neighbors.
2. Compute similarity between the matched sets, e.g. root mean square distance.
3. Compute incremental transformation by solving for the least-squares solution using the current correspondences.
4. Apply the obtained transformation to one of the point sets.
5. Repeat the steps until either similarity is less than a certain threshold or the maximum number of iterations is reached.

3.6 Hand Eye Calibration

The topic of hand-eye calibration is covered in the preceding Section 2.3. The major purpose of calibration is to solve the equations from $MX = YN$. As previously described, a simultaneous technique is used to compute the X end-effector to the sensor and Y robot to world matrices using the QR24 calibration algorithm [25]. This section discusses the important steps and equations used to solve the equations.

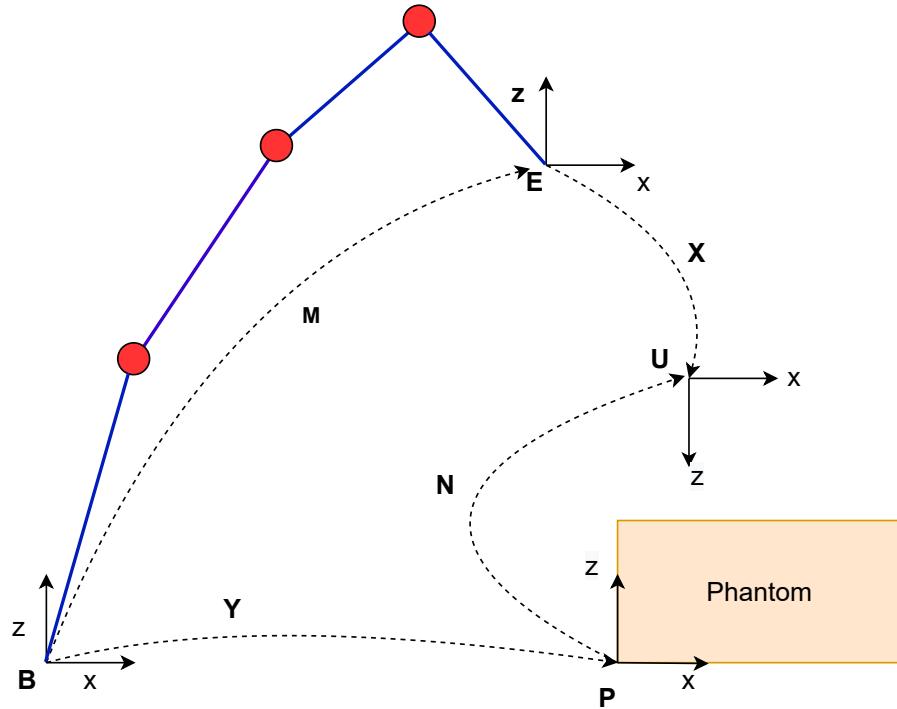


Fig. 3.12: Figure illustrates the various coordinate transformations. Here $M = {}^B H_E$, $X = {}^E H_U$, $Y = {}^B H_P$ and $N = {}^P H_U$.

A simplistic approach to end-effector to sensor and robot to world calibration is to examine the general relationship.

$${}^B H_E {}^E H_U = {}^B H_P {}^P H_U \quad (3.15)$$

Here ${}^B H_E$ denotes the transformation from the robot base to end-effector frames and ${}^P H_U$ denotes the transformation from the phantom coordinate system to the US probe coordinate system. The robot to end-effector transformation is simply accomplished by applying the robot forward kinematics and the phantom to probe transformation is obtained via point registration, as previously explained in Section 3.5. In addition, ${}^E H_U$ is the transformation from the robot end-effector coordinate frame to the ultrasound probe coordinate frame, and ${}^B H_P$ is the transformation from the robot base to the phantom coordinate system. The alterations can be seen in the image 3.12. To calculate these matrices, n measurements for various robot positions are recorded and the corresponding pose from phantom to ultrasound is computed as discussed in Section 3.3, yielding n

equations. Which can be expressed as shown in Equation 3.16.

$$M_i X = Y N_i \quad i = 1, 2, \dots, n \quad (3.16)$$

By considering the non-trivial values of X and Y as vector components.

$$z = [x_{1,1}, x_{2,1}, \dots, x_{3,4}, y_{1,1}, y_{2,1}, y_{3,4}] \in R^{24} \quad (3.17)$$

Equation 3.16 can be represented as a linear equation as shown.

$$Az = b \quad (3.18)$$

where $A \in R^{12n \times 24}$ and $b \in R^{12n}$ and

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \quad b = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \quad (3.19)$$

and A_i and b_i can be expressed as shown in Equation 3.20 and 3.21.

$$A_i = \begin{bmatrix} R[M_i](N_i)_{1,1} & R[M_i](N_i)_{2,1} & R[M_i](N_i)_{3,1} & Z_{3 \times 3} \\ R[M_i](N_i)_{1,1} & R[M_i](N_i)_{2,1} & R[M_i](N_i)_{3,1} & Z_{3 \times 3} \\ R[M_i](N_i)_{1,1} & R[M_i](N_i)_{2,1} & R[M_i](N_i)_{3,1} & Z_{3 \times 3} - E_{12} \\ R[M_i](N_i)_{1,1} & R[M_i](N_i)_{2,1} & R[M_i](N_i)_{3,1} & R[M_i] \end{bmatrix} \quad (3.20)$$

$$b_i = \begin{bmatrix} Z_{9 \times 1} \\ -T[M_i] \end{bmatrix} \quad (3.21)$$

Where $R[M_i] \in R^{3 \times 3}$ is the rotational part of M_i , $T[M_i] \in R^3$ is the translation part of M_i , Z_m denotes a $m \times n$ zero matrix and E_k is a $k \times k$ identity matrix. Using QR-factorisation, this system can then be solved in the least-squares sense. The non-trivial entries of the matrices X and Y are now contained in the resulting vector z as shown in Equation 3.17. It should be noted that this method attempts to identify the best entries for X and Y in terms of quadratic error, that is, it minimizes equation 3.22 where the Frobenius norm is represented by $\| \cdot \|_F$, therefore the resulting matrices are not necessarily orthogonal. Since 24 elements must be estimated, this method is called the QR24 calibration algorithm, and QR-factorisation is utilized to solve the equation system.

$$\sum_{i=1}^n = \| M_i X - Y N_i \|_F \quad (3.22)$$

The poses calibrated with these matrices must be orthonormalized for the final result because X and Y do not have to be orthogonal during computation. Otherwise, rigid body changes based on the discovered calibration matrices would be pointless. Orthonormalization is performed using SVD. Assume N is a tracking system-measured pose, and coordinates in the robot coordinate system are to be computed. Then a non-orthogonal matrix, $Y N$, is computed, and then the closest orthogonal matrix of $Y N$ is estimated.

$$U \Sigma V = SVD(Y N) \quad (3.23)$$

$$YN = UV^T \quad (3.24)$$

3.7 3D US Volume Reconstruction

The results of the calibration are validated by scanning a 3D printed phantom and surface reconstruction using 2D US images. The reconstructed volume is then compared to the Computer Aided Design (CAD) model of the phantom. The US probe mounted at the end-effector of the robot and is driven over the phantom in a straight path and the data is recorded using B-mode imaging as shown in Figure 3.13. 120 images are recorded over the length of 40 mm. The process is repeated from multiple angles. For every recorded data the pose of the manipulator is also noted.



Fig. 3.13: Figure shows the surface of phantom scanned for validation.

The recorded raw data is stored as an image. The reflected sound wave from the phantom appears as high-intensity pixels representing the surface of the phantom as shown in Figure 3.14a. The images are segmented using median filter and thresholding filters to obtain all the pixels representing the phantom surface as shown in Figures 3.14b and 3.14c. The coordinates of the high-intensity pixels obtained are scaled from pixel system to cartesian coordinate system to match the size of the actual phantom as shown in Figure 3.14d. The points are then transformed to the robot base coordinate frame using the corresponding robot pose and calibration matrix as described in Equation 3.25.

$$P_B(x, y, z) = {}^B H_E \times {}^E H_U \times S \times P_I(u, v, w) \quad (3.25)$$

where $P_B(x, y, z)$ is the point on the surface of the phantom in the robot base coordinate frame. ${}^B H_E$ is the transformation from robot base to end-effector frames obtained from

the forward kinematics of the robot and ${}^E H_U$ is the transformation for robot end-effector to US probe frame which is the result of hand-eye calibration. S represents the scaling matrix to convert the obtained points from the pixel system to the cartesian coordinate system. $P_I(u, v, w)$ denotes the segmented point in the pixel system augmented with zero along the v axis to produce a 3D surface point.

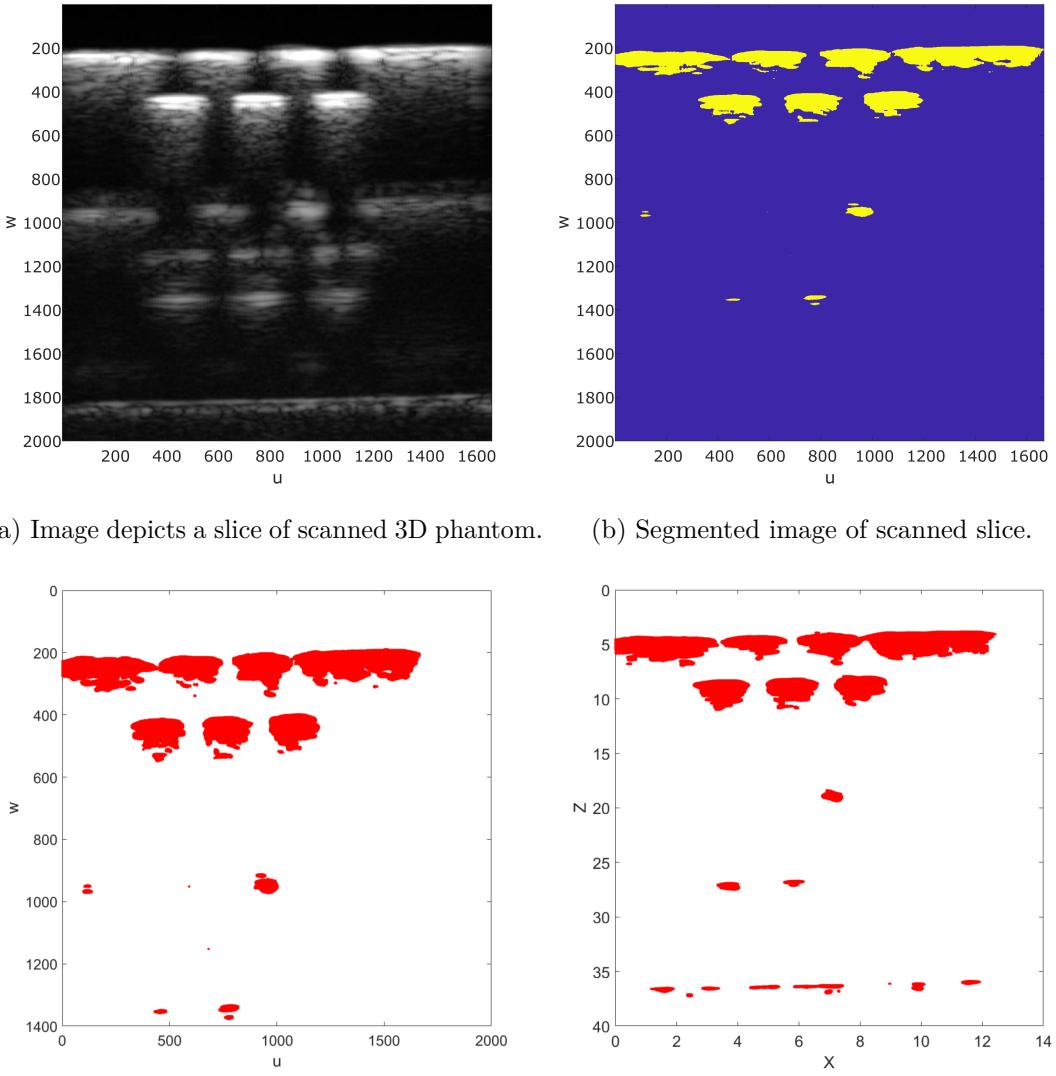


Fig. 3.14: Steps to segment the surface points of the phantom.

Given an image, every surface point obtained is transformed and stored as a point cloud. The process is repeated for every image and the obtained point clouds are merged to eliminate redundancy. Redundancy is seen when the same surface is scanned from a different angle. The merged point cloud is then visualized and compared to the CAD model of the phantom. Visualization of the reconstructed surface is presented in Chapter 4.

4 Results and Discussion

This chapter discusses the details of the results accomplished during the experiment and compares them to the state of art.

4.1 US Scanning and Image Segmentation

The outcomes of US scanning are presented in Figure 4.1. Images obtained when the image plane is perpendicular to the N-fiducial are relatively better as shown in Figure 4.1a, and a specific set of time gain compensation in the US system is employed during recording. When the probe is tilted along the x axis with respect to the probe coordinate frame as shown in Figure 3.4b, Figures 4.1e and 4.1f become distorted, and the time gain compensation of the US system must be adjusted to provide images that are convenient to segment.

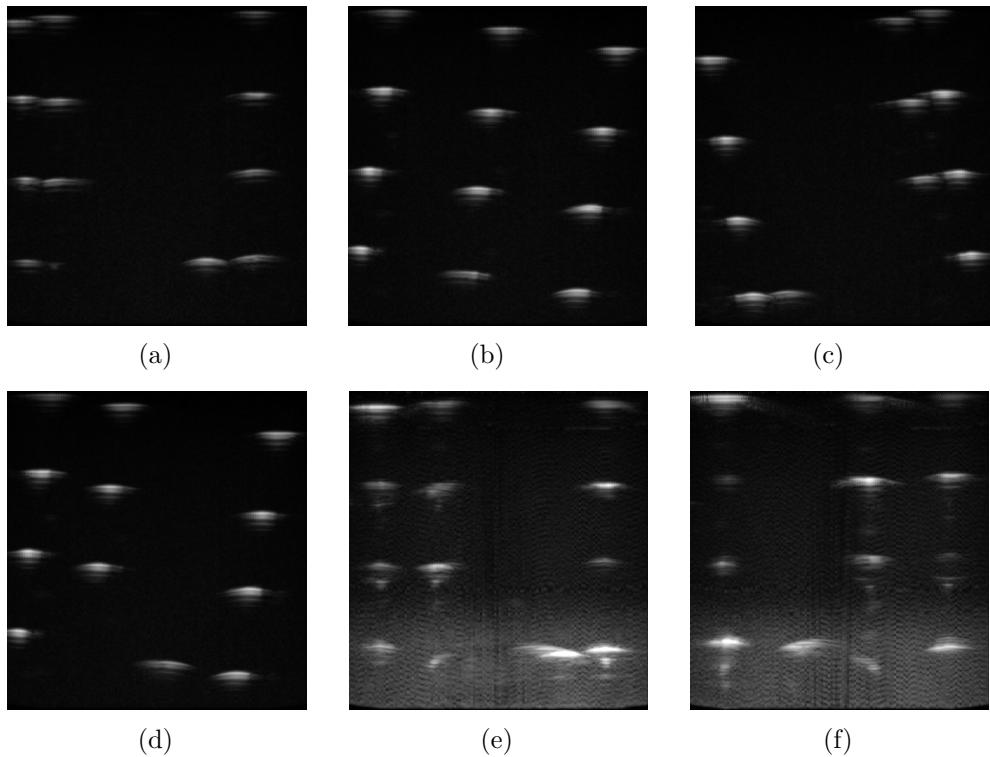


Fig. 4.1: US images acquired by placing the probe at different viewing angles with respect to phantom coordinate frame

The acquired images are segmented using the automatic segmentation algorithm as explained in Section 3.4. Figure 4.2a represents the image plane and its intersection with the wire phantom. Left and right strands of the N-wire fiducial are also indicated. It is

4 Results and Discussion

established that scanning the N-wire phantom generates 12 spots in the image. The 12 clusters are identified and enclosed in the bounding box also, the centers of the clusters are marked as shown.

The euclidean distance between markers of the left strand and right strand is 30 mm and the distance between adjacent layers of N-fiducials is 10 mm as discussed in Section 3.2. Given the distance between the markers in an image, the segmentation error is computed as discussed in Section 3.4.

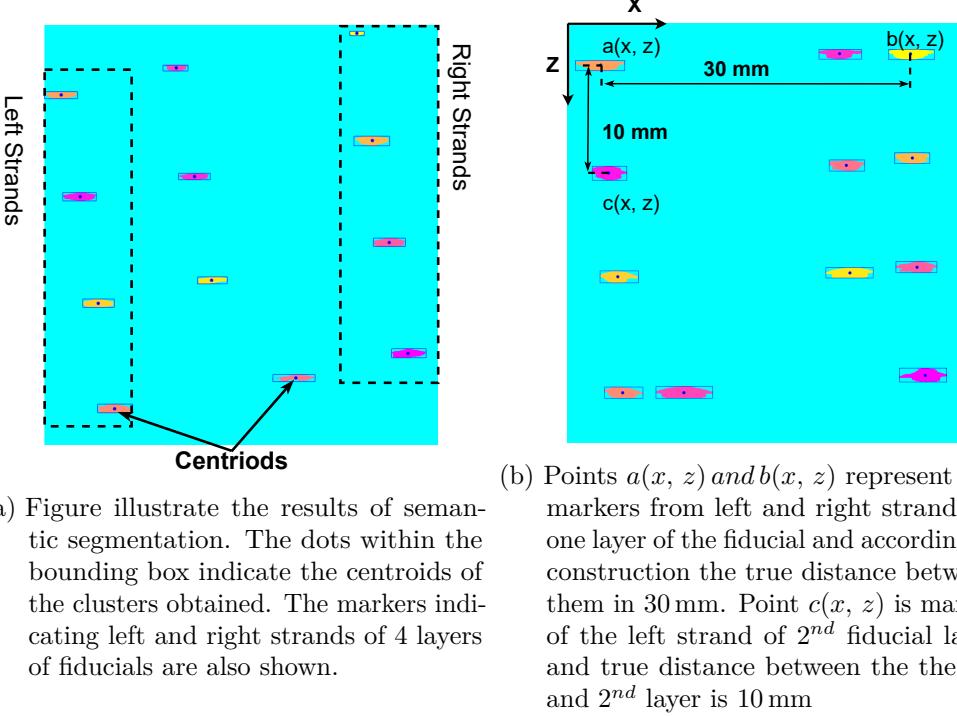


Fig. 4.2: Segmentation results

For multiple trials conducted the mean deviation of euclidean distance computed from the markers to reference size N-fiducial and Standard Deviation (SD) along X and Z directions is presented in Table 4.1. A visual description of reference size is shown in Figure 4.2b. A PC workstation (Intel i7, CPU @2.7 GHz, 32G RAM) is used for computation.

Tab. 4.1: Comparison of mean semantic segmentation error for two independent trials.
And mean time required for computation.

| Trial No. | No. of Images (N) | Mean Error (mm) | | SD | | Time (s/N images) |
|-----------|-------------------|-----------------|--------|--------|--------|-------------------|
| | | X | Z | X | Z | |
| 1 | 130 | 0.3812 | 0.5670 | 0.3957 | 0.1987 | 45.1359 |
| 2 | 230 | 0.6743 | 0.5530 | 0.3955 | 0.2315 | 145.3185 |

Most of the related work in the literature relies on manual segmentation techniques to obtain the markers [16], [6]. Manual segmentation is highly time-consuming and depends

on the expertise of the person selecting the markers from the image. Accuracy also varies with the resolution of the US image. Only a few research groups have applied an automatic segmentation algorithm. Chunxu Shen [13] proposed an effective point-recognition algorithm, which combines point detection and speckle tracking, to address the issue and determine the relation between intersections in US image space and line equations in calibration phantom space. In another research, [17] RANSAC was used to exclude the outliers and estimate the markers entitled by the pattern of N-fiducial alignment. This work shows a simple and computationally fast automatic segmentation algorithm in comparison to the works in the literature. The segmentation process takes about 0.3461 seconds for one image. The computation time for various trials conducted are shown in Table 4.1.

4.2 Registration

Target localization of the markers is achieved using the ICP as discussed in Section 3.5. Figures 4.3 and 4.4 give a visual interpretation of the localization. The blue lines represent the N-fiducials and the red dots represent the markers obtained from segmentation. Marker points after the initial alignment are shown in Figure 4.3a. Using the transformation matrix obtained from ICP the marker points are transformed to match their corresponding points as shown in Figure 4.3b.

For all the N ($N = 130$) images acquired during the acquisition, segmentation provides the 12 markers for every image. Every marker is transformed using the transformation matrix obtained from ICP and compared to virtual phantom as shown in Figure 4.4.

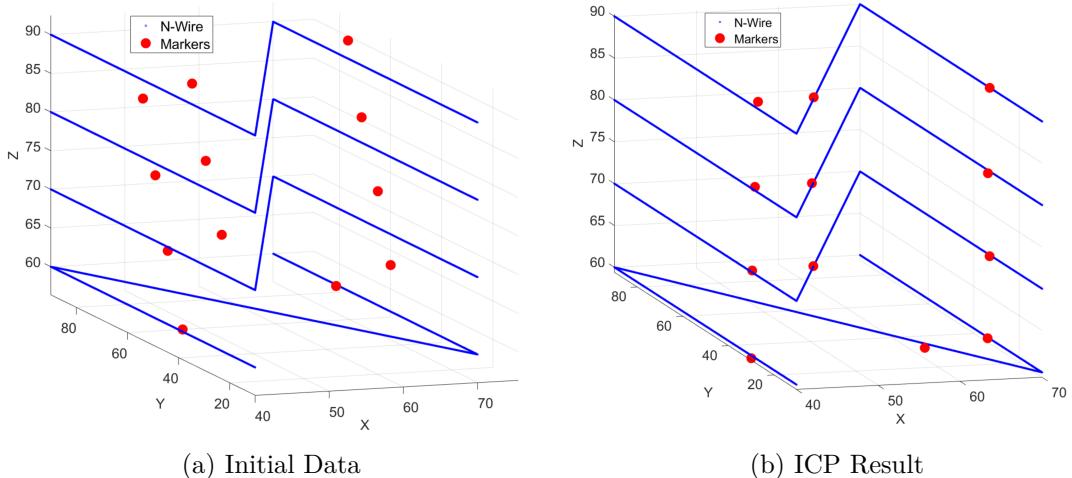


Fig. 4.3: The figures illustrates the results of semantic segmentation. The dots within the bounding box indicate the centroids of the clusters obtained

For every marker in every image the euclidean distance between its corresponding match is computed to obtain FRE as discussed in Section 3.5. The mean FRE and SD for 2 independent trials is summarised in Table 4.2.

4 Results and Discussion

Tab. 4.2: Comparison of mean FRE for two independent trials.

| Trial No. | No. of Images | FRE (mm) | SD |
|-----------|---------------|----------|--------|
| 1 | 130 | 0.7923 | 0.3108 |
| 2 | 230 | 0.7318 | 0.2772 |

Most of the related work [29] [12] in the literature match the markers obtained for the middle strand using the Equations 3.6 and 3.7. Computing the transformation using the middle strand markers leads to good results but using markers from all the strands leads to more accurate results. In the work by Chunxu Shen [29], the mean FRE is 1.140 mm. Another work introduced a complex phantom and performed the registration for three markers using ICP algorithm and achieved a RMS error of 0.505 mm [19]. This work matches all the 12 markers obtain from N-fiducials and computes the transformation using ICP approach giving comparable results to the works in the literature. Table 4.3 summarizes the comparison of FRE between the proposed method in this work and the state of art methods.

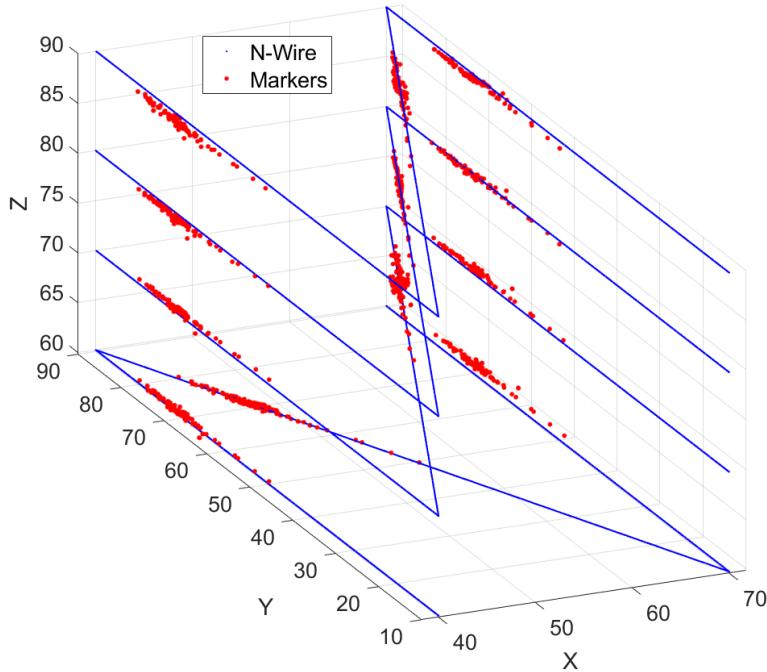


Fig. 4.4: Figure illustrates the results of target localization. The blue lines represent the virtual N-wire phantom and the red dots indicate the centroids obtained from images

Tab. 4.3: Comparison of mean FRE with state of art.

| Method | FRE (mm) |
|--|----------|
| Proposed | 0.7318 |
| Arbitrary wire method [29] | 1.0220 |
| Cross-wire phantom with 8 threads [34] | 1.6700 |
| Complex phantom [19] | 2.3300 |

4.3 Hand-Eye Calibration

Figure 4.5 shows the rotation and translation accuracy of the hand-eye calibration between the robot end-effector and US probe. Initially, the calibration is performed using 25 unique poses. The effect of increasing the number of input images on calibration accuracy is also evaluated. The calibration computation is performed by increasing 20 input images per iteration. The rotation and translation errors are computed for the remaining poses by performing cross-validation. The cross-validation technique used is an exhaustive method, that involves using p-observation as validation data, and remaining data is used to compute the calibration. This method is called as Leave p out cross-validation. Using the calibration matrices ${}^E H_U$ and ${}^B H_P$, the calibration quality is computed by Equation 4.1.

$$I = X^{-1} \times M^{-1} \times Y \times N \quad (4.1)$$

Here, X and Y are the results from hand-eye calibration, and M and N are the matrices from robot kinematics and fiducial registration. This should give an identity matrix if everything would be perfect because they describe $X M = Y N$. As it is impossible in the real world to achieve a perfect calibration the matrix will not be an identity matrix. From this matrix, translation error is computed using norm of x , y and z coordinates using Equation 4.2.

$$e_{trans}[I] = \sqrt{I_{1,4}^2 + I_{2,4}^2 * I_{3,4}^2} \quad (4.2)$$

To determine the rotational accuracy, let $(k, !)$ be the axis-angle representation of $R[I]$. Then the rotation error is defined as

$$e_{rot}[I] = |\theta| \quad (4.3)$$

The average translation and rotation error for 80 poses in validation set is 4.081 mm and 12.4308°. Figure 4.5 also depicts the effect on translation and rotation accuracy by increasing the number of poses. Hence, the deviation of translation and rotation error between computation using 25 images and 105 images, with 80 images in the validation set in each case, is around 0.07 mm. Table 4.4 shows the calibration results for two independent trials. Also, it can be seen that the error increases with orthonormalization as some information is lost with SVD computation as discussed in Section 3.6.

4 Results and Discussion

Tab. 4.4: Comparison of calibration results for two independent trials.

| Trial No. | Mean translation error | Mean translation error after orthonormalization | Mean rotation error | Mean rotation error after orthonormalization |
|-----------|------------------------|---|---------------------|--|
| 1 | 0.8178 mm | 4.081 mm | 9.477° | 12.4308° |
| 2 | 0.93658 mm | 4.1932 mm | 9.67° | 12.3538° |

Resolution in a single ultrasound image is different from one area to another area. This factor can be decomposed into axial and lateral resolutions. The axial does not change with the image depth, but the lateral resolution is depth-dependent [29]. Figure 4.1 shows the four layers of N-fiducials. The markers at the bottom of images are more blurry. This phenomenon would influence the accuracy of the segmentation and registration algorithm, which further harmful to the proposed calibration method. Also registration of a 2D image to a 3D phantom further reduces the quality of the calibration. In such a case, the translation error of 4.081 mm gives a good scope for further research and development. Also in comparison to the actuator-based calibration [16] with a calibration error of 4.91 mm, the proposed method in this work shows an improvement.

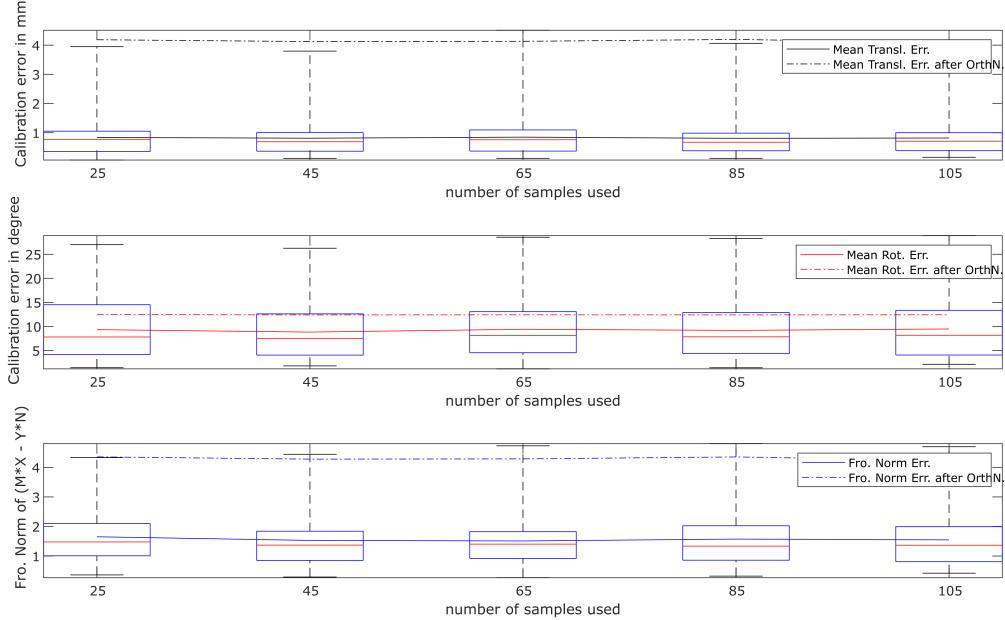


Fig. 4.5: Figure illustrates the evaluation of the hand-eye results. The translation and rotation error for input images from 25 to 105 is plotted.

4.4 Surface Reconstruction using 2D US Images

To validate the results, 3D phantom is scanned as discussed in Section 3.7. Figure 4.6 depicts the surface of the phantom scanned for validation.



Fig. 4.6: Figure illustrates the actual surface of the 3D part phantom

Figure 4.7 and 4.8 shows the visualization of the surface reconstruction of the scanned phantom from the side and top view respectively. With respect to the surface of the printed 3D phantom, the reconstructed surface using 2D images is comparable.

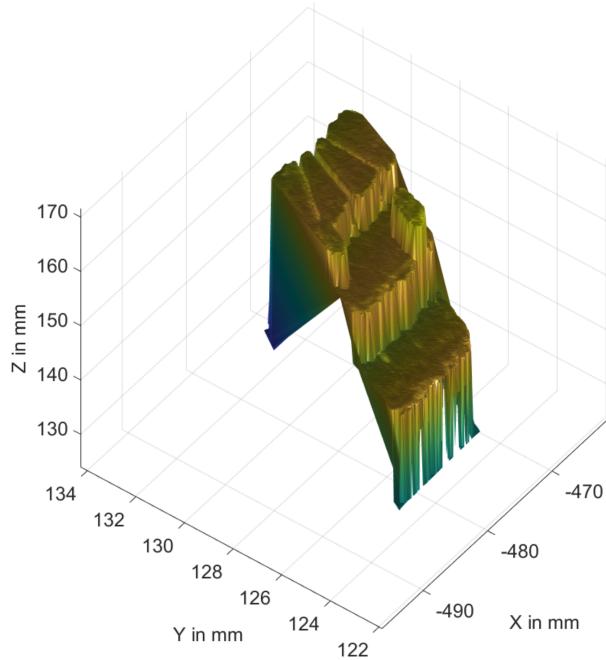


Fig. 4.7: Figure illustrates the isometric view of the reconstructed surface of the 3D phantom.

4 Results and Discussion

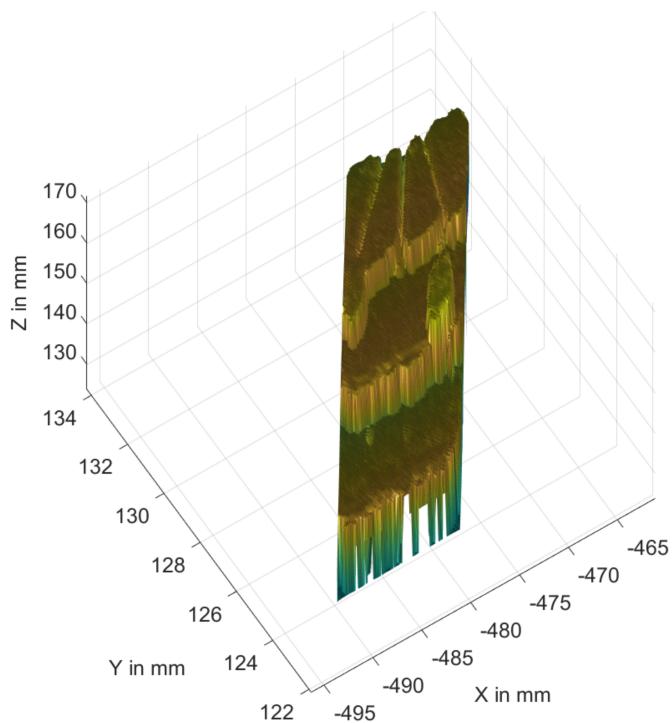


Fig. 4.8: Figure illustrates the reconstructed surface of the 3D phantom.

5 Conclusion

5.1 Conclusion

US probe calibration is an important step to ensure a high quality 3D US reconstruction from 2D US images. In this research, three important steps for robotic ultrasound systems are combined, which include data acquisition, US image segmentation, image markers to N-fiducials registration, probe to robot calibration, and robotic 3D ultrasound reconstruction into a unified pipeline for robotic US system. The proposed framework realizes a fast and precision US probe to robot calibration and eventually achieves accurate robotic ultrasound 3D reconstruction from 2D US images. The proposed framework also has been validated to provide accurate 3D reconstruction using 2D images. The duration of full calibration procedures does not exceed 5 minutes with translation and rotation error of 4.081 mm and 12.4308° respectively, which has great potential to facilitate the advancement of automatic robot ultrasound scanning. In this work, a probe to phantom spatial calibration method was presented and developed using the point-to-point registration algorithm. The experimental results showed that the calibration accuracy increased with the number of data samples used but with a drawback of longer robot operation time.

This work mainly aims to improve the performance of robot-assisted US probe calibration. In comparison to different US calibration phantoms, the N-wire phantom is chosen since it enables a rapid automatic image segmentation for robotic US scanning without manual intervention. Also, position controlled US scanning ensures a sufficiently diverse range of probe poses avoiding the under-determined results when estimating the scanning plane that is the fiducial registration. Moreover, the segmentation time for each image is 0.05 seconds. In contrast, conventional manual image segmentation and post-processing usually last several minutes. Since the US image processing time has been noticeably shortened, more images could be collected from different poses to improve the calibration accuracy. Therefore, the proposed automatic US image calibration framework has potential applications in the robotic-assisted system thanks to its feasibility and stability.

In the experimental part, the proposed framework is analyzed to be feasible. The fiducial registration results show that the transformed US image markers could accurately outline the geometric features in comparison with the respective N-fiducial. To achieve an accurate hand-eye calibration between the US probe and robot, the accuracy of registration plays a vital role. The mean 3D localization error was 0.7049 mm for three repeated experiments. Unlike most ultrasound probe calibration methods found in the literature, the method proposed in this work does not require any additional 3D localizer. Despite the use of less accurate systems compared to tracking devices, the resulting precision is good. The validation results show that the reconstructed point clouds could accurately outline the geometric features compared with the respective geometry model. The hand-eye calibration result is essential to display the reconstructed model to the surgeon in medical scenarios interacting with other surgical instruments.

5 Conclusion

Robotic US imaging is an emerging technology that reduces workload and ensures optimal probe scanning rather than freehand manipulation. A fast, automatic and reliable robot US image calibration method facilitates the deployment and implementation of robotic US systems in practice and therefore enhances the ultimate accuracy and precision of robot-assisted applications. This calibration also enables the AR system to generate and display a hologram of the scanned region of interest achieving the goal of the project. US technology with AR can be used for US-guided interventions simplifying the operator's spatial orientation, with a potential for reducing the inter and intra-operator variability. The operators could be faster and need fewer needle passes for US-guided interventions. It may reduce experience-based differences in performing US-guided interventions.

5.2 Future Work

The algorithm proposed and implemented in this work further offers a multitude of various tasks for improvement, which are briefly listed below.

The advantages of the RANSAC algorithm have been extensively discussed over the years. It is suited for applications in automated image analysis where the image is error-prone such as US images [35] [17]. RANSAC can be used to exclude the outliers and estimate the markers modeled as a 2D plane entitled by the pattern of N-fiducial alignment. This aligns the markers into a straight line which improves the segmentation error.

The virtual phantom used in this work is manually generated by specifying the dimensions of N-fiducials with each strand consisting of thousand points. The accuracy of registration can be improved by increasing the number of points or by tracing the path of N-fiducial using an optical stylus. Also, the extracted markers can be modeled as a plane that can be used to register the 3D pose as a point-to-plane problem. This can be done by the ICP algorithm [33] which will improve the rotation accuracy. Another approach to improve the registration is to use 3D optical trackers at the probe coordinate frame and phantom coordinate frame. Using the poses obtained, the transformation from probe to phantom frame can be computed [19] [29].

In this work, the phantom is scanned by keeping the probe perpendicular to the surface of the phantom for the surface reconstruction using 2D US images. A further improvement of the 3D reconstruction can be done by acquiring 2D images of the phantom from multiple angles and registering them together to form a point cloud. This ensures the side view of the object is also scanned and reconstructed rather than only the top view.

Bibliography

- [1] Osama Shahin. *Minimally Invasive Navigated Liver Intervention: Ultrasound-guided Surgery and Ablation Validation*. PhD thesis, Zentrale Hochschulbibliothek Lübeck, 2014.
- [2] Cristian A Linte, John T Moore, Elvis CS Chen, and Terry M Peters. Image-guided procedures: tools, techniques, and clinical applications. In *Bioengineering for Surgery*, pages 59–90. Elsevier, 2016.
- [3] Michael Unger, Johann Berger, and Andreas Melzer. Robot-assisted image-guided interventions. *Frontiers in Robotics and AI*, 8:221, 2021.
- [4] Mustafa Z Mahmoud, Mohammed Aslam, Mohammed Alsaadi, Maram A Fagiri, and Batil Alonazi. Evolution of robot-assisted ultrasound-guided breast biopsy systems. *Journal of radiation research and applied sciences*, 11(1):89–97, 2018.
- [5] Chunwoo Kim, Felix Schäfer, Doyoung Chang, Doru Petrisor, Misop Han, and Dan Stoianovici. Robot for ultrasound-guided prostate imaging and intervention. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 943–948. IEEE, 2011.
- [6] Johan Sarrazin, Emmanuel Promayon, Michael Baumann, and Jocelyne Troccaz. Hand-eye calibration of a robot-ultrasound probe system without any 3d localizers. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 21–24. IEEE, 2015.
- [7] Renáta Elek, Tamás D Nagy, Dénes A Nagy, Bence Takács, Péter Galambos, Imre Rudas, and Tamás Haidegger. Robotic platforms for ultrasound diagnostics and treatment. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1752–1757. IEEE, 2017.
- [8] Elvis CS Chen, Terry M Peters, and Burton Ma. Guided ultrasound calibration: where, how, and how many calibration fiducials. *International journal of computer assisted radiology and surgery*, 11(6):889–898, 2016.
- [9] Nadja A Farshad-Amacker, Till Bay, Andrea B Rosskopf, José M Spirig, Florian Wanivenhaus, Christian WA Pfirrmann, and Mazda Farshad. Ultrasound-guided interventions with augmented reality in situ visualisation: a proof-of-mechanism phantom study. *European radiology experimental*, 4(1):1–7, 2020.
- [10] Heather Venables. How does ultrasound work? *Ultrasound*, 19(1):44–49, 2011.
- [11] Thomas Kuiran Chen, Adrian D Thurston, Mehdi H Moghari, Randy E Ellis, and Purang Abolmaesumi. A real-time ultrasound calibration system with automatic accuracy control and incorporation of ultrasound section thickness. In *Medical*

Bibliography

- Imaging 2008: Visualization, Image-Guided Procedures, and Modeling*, volume 6918, page 69182A. International Society for Optics and Photonics, 2008.
- [12] Guillermo Carbajal, Andras Lasso, Álvaro Gómez, and Gabor Fichtinger. Improving n-wire phantom-based freehand ultrasound calibration. *International journal of computer assisted radiology and surgery*, 8(6):1063–1072, 2013.
 - [13] Chunxu Shen, Liushuai Lyu, Guangzhi Wang, and Jian Wu. A method for ultrasound probe calibration based on arbitrary wire phantom. *Cogent Engineering*, 6(1):1592739, 2019.
 - [14] Tiexiang Wen, Cheng Wang, Yi Zhang, and Shoujun Zhou. A novel ultrasound probe spatial calibration method using a combined phantom and stylus. *Ultrasound in Medicine & Biology*, 46(8):2079–2089, 2020.
 - [15] Rika Rokhana, Ketut E Purnama, Eko M Yuniarso, and Mauridhi H Purnomo. Estimation of wire phantom’s position in ultrasound probe calibration based on polynomial equation. In *Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence*, pages 137–141, 2017.
 - [16] Terry K Koo and Nathaniel Silvia. Actuator-assisted calibration of freehand 3d ultrasound system. *Journal of healthcare engineering*, 2018, 2018.
 - [17] Maria Chatrasingh and Jackrit Suthakorn. A novel design of n-fiducial phantom for automatic ultrasound calibration. *Journal of medical physics*, 44(3):191, 2019.
 - [18] Hui Zhang, Filip Banovac, Amy White, and Kevin Cleary. Freehand 3d ultrasound calibration using an electromagnetically tracked needle. In *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*, volume 6141, page 61412M. International Society for Optics and Photonics, 2006.
 - [19] Ralf Bruder, Florian Griese, Floris Ernst, and Achim Schweikard. High-accuracy ultrasound target localization for hand-eye calibration between optical tracking systems and three-dimensional ultrasound. In *Bildverarbeitung für die Medizin 2011*, pages 179–183. Springer, 2011.
 - [20] Mili Shah, Roger D Eastman, and Tsai Hong. An overview of robot-sensor calibration methods for evaluation of perception systems. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*, pages 15–20, 2012.
 - [21] Y. Shiu and S. Ahmad. Finding the mounting position of a sensor by solving a homogeneous transform equation of the form $ax = xb$. In *Proceedings. 1987 IEEE International Conference on Robotics and Automation*, volume 4, pages 1666–1671, 1987.
 - [22] R.Y. Tsai and R.K. Lenz. A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.
 - [23] Rong-hua Liang and Jian-fei Mao. Hand-eye calibration with a new linear decomposition algorithm. *Journal of Zhejiang University-SCIENCE A*, 9(10):1363–1368, 2008.

- [24] Zhiqiang Zhang, Lin Zhang, and Guang-Zhong Yang. A computationally efficient method for hand–eye calibration. *International journal of computer assisted radiology and surgery*, 12(10):1775–1787, 2017.
- [25] Floris Ernst, Lars Richter, Lars Matthäus, Volker Martens, Ralf Bruder, Alexander Schlaefer, and Achim Schweikard. Non-orthogonal tool/flange and robot/world calibration. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 8(4):407–420, 2012.
- [26] Jackrit Suthakorn, Narucha Tanaiutchawoot, and Cholatip Wiratkapan. Ultrasound calibration with ladder phantom at multiple depths for breast biopsy navigation system. *Theoretical and Applied Mechanics Letters*, 10(5):343–353, 2020.
- [27] Po-Wei Hsu, Graham M Treece, Richard W Prager, Neil E Houghton, and Andrew H Gee. Comparison of freehand 3-d ultrasound calibration techniques using a stylus. *Ultrasound in medicine & biology*, 34(10):1610–1621, 2008.
- [28] Emad M Boctor, Ameet Jain, Michael A Choti, Russell H Taylor, and Gabor Fichtinger. Rapid calibration method for registration and 3d tracking of ultrasound images using spatial localizer. In *Medical Imaging 2003: Ultrasonic Imaging and Signal Processing*, volume 5035, pages 521–532. International Society for Optics and Photonics, 2003.
- [29] Guangzhi Wang Jian Wu Chunxu Shen, Liushuai Lyu. A method for ultrasound probe calibration based on arbitrary wire phantom. *A method for ultrasound probe calibration based on arbitrary wire phantom*, 24(1):3–11, 2019.
- [30] Wei Chen, Sheng Hu, Hao Cao, Tao Huang, Xiaowen Wu, Ling Lu, and Jiwen Peng. Review on research process of sound reduction materials. In *IOP Conference Series: Materials Science and Engineering*, volume 612, page 052062. IOP Publishing, 2019.
- [31] Ziyang Wang. Deep learning in medical ultrasound image segmentation: A review. *arXiv preprint arXiv:2002.07703*, 2020.
- [32] Vivian Bass, Julieta Mateos, Ivan M Rosado-Mendez, and Jorge Márquez. Ultrasound image segmentation methods: A review. In *AIP Conference Proceedings*, volume 2348, page 050018. AIP Publishing LLC, 2021.
- [33] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, volume 1611, pages 586–606. International Society for Optics and Photonics, 1992.
- [34] Awais Ahmad, M Cenk Cavusoglu, and Ozkan Bebek. Calibration of 2d ultrasound in 3d space for robotic biopsies. In *2015 International Conference on Advanced Robotics (ICAR)*, pages 40–46. IEEE, 2015.
- [35] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

List of Abbreviations

| | |
|---------------|-------------------------------|
| AR | Augmented Reality |
| CAD | Computer Aided Design |
| CT | Computed Tomography |
| CNN | Convolutional Neural Networks |
| FRE | Fiducial Registration Error |
| ICP | Iterative Closest Point |
| MRI | Magnetic Resonance Imaging |
| RANSAC | Random Sample Consensus |
| RMS | Root Mean Square |
| ROI | Region of Interest |
| SD | Standard Deviation |
| SVD | Singular Value Decomposition |
| US | Ultrasound |

Declaration

Hiermit versichere ich, dass ich meine Abschlussarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Datum:
(Unterschrift)