Path Planning for Trajectory Guided Freehand Ultrasound Scan

by

Qian Lin

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

Author .....
Department of Mechanical Engineering
May 12, 2023

Certified by .....
Brian W. Anthony
Principal Research Scientist
Thesis Supervisor

Accepted by .....
Nicolas G. Hadjiconstantinou
Chairman, Mechanical Engineering Committee on Graduate Theses

Path Planning for Trajectory Guided Freehand Ultrasound Scan

by

Qian Lin

Submitted to the Department of Mechanical Engineering
on May 12, 2023, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

Medical imaging plays a crucial role in medical diagnosis and analysis. 3D medical imaging provides more comprehensive and greater anatomical detail of internal body structures when compared to traditional 2D images, allowing for more accurate measurement of organ and tumor volume, and prediction and monitoring of some disease progression. 3D medical images can be obtained through various imaging modalities, including magnetic resonance (MR), computed tomography (CT), and ultrasound (US). Among those modalities, freehand ultrasound is preferred for its cost-effectiveness, non-invasiveness, portability, safety, versatility, and real-time information.

However, the lack of information on the position and orientation of the ultrasound probe makes it challenging to obtain 3D images from 2D ultrasound slices. Without the expert knowledge, the user may not acquire precise images on the region of interest (RoI). To address this issue, we proposed a novel path planning framework that provides real-time guidance for freehand ultrasound and reconstructs 3D images in real-time. A low-cost RGB-D camera with IMU module is mounted on a regular ultrasound probe to estimate the spatial placement of the probe with respect to the RoI, and the acquired ultrasound images are analyzed and registered into 3D voxel grid. After the user performs initial scan, the system guides the user to find missing areas shaded by obstacles such as bones, resulting in more accurate, detailed, and efficient 3D ultrasound imaging. We validated our system on an ultrasound phantom and demonstrated its ability to investigate the area beneath the obstacle. Additionally, we developed a visualization system for real-time probe movement guidance and

image display.

This study demonstrates the feasibility of implementing an online path planning approach with real-time guidance and high-attenuation area avoidance for freehand ultrasound scanning, even in scenarios where prior knowledge of the scanning area is not available. The proposed path planning system not only enhances the efficiency and precision of ultrasound imaging in clinical settings, but also facilitates the acquisition of high-quality 3D ultrasound images by non-expert users in a more convenient manner, potentially allowing for long-term health monitoring.

Thesis Supervisor: Brian W. Anthony
Title: Principal Research Scientist

Acknowledgments

As I am finishing the first step of my graduate study, I would like to express my sincere gratitude to my supervisor, Dr. Brian Anthony, for his invaluable guidance, support, and encouragement throughout my research project. His expertise and constructive feedback have been instrumental in shaping my ideas and improving the quality of this thesis.

I am grateful to my labmates, especially April and Xuan, who provided me with their experience in equipment selection and use, and collaborated with me on creating the validation phantom.

I also want to thank my family for their unwavering support, love, and encouragement throughout this journey. Their constant belief in me has been a great source of inspiration and motivation. I am also deeply grateful to my friends, especially Yiwei, Jianjia, Biyang, and Zhibo, who encouraged me and supported me during the tough times.

Finally, I would like to express my appreciation to Buddhism, which has given me constant guidance in all aspects of my life. The wisdom I derived from Buddhism has also been invaluable in my research. I would like to thank Thousand Buddha Temple and all the monks and friends I met there, especially Master Xianfeng and Lina, who have provided me with excellent examples of how to apply the teachings of Buddhism from beyond the world to the worldly life and work.

Once again, thank you to everyone who has contributed to my academic journey. Your support and encouragement have been invaluable.

Contents

List of Figures

Chapter 1

Background and Motivation

1.1 Volumetric Medical Image

Medical imaging technologies play an important role in clinical diagnosis and medical treatment. Popular image modalities include X-ray, magnetic resonance (MR), and computed tomography (CT), ultrasound (US).

These technologies have been applied in the measurement of in vivo organ volumes, which has drawn much attention. The organ volumes can provide the whole organ view for a better understanding of the body status of patients. Optimizing the imaging parameters is essential to image quality; organ volumes can also provide situational awareness and context for 2D slice acquisition. Moreover, organ or tumor volume measurement is widely used in monitoring and longitudinal analysis of disease progression.

In particular, the measurement and monitoring of volumetric information in some diseases is very important and even essential [12].

1.1.1 Existing Image Modalities for 3D Volume Measurement

Computed Tomography (CT) Image

Computed tomography (CT) is a medical diagnostic imaging technique that utilizes X-rays to measure the projection of an object from multiple surrounding directions, and then construct a 3D image of the linear attenuation coefficient throughout the object. The images are typically obtained as a sequence of parallel axial slices, which can be "stacked" into a 3D image, showing the 3D volume and properties of the object [75]. Since the location and orientation of the slice images are known, and the slices are parallel to each other, the reconstruction of 2D X-ray images into 3D image only involves some basic translation and would also be more accurate when compared with ultrasound imaging.

Segmentation of volume reconstructions can be used for quantitive measurement. Many researches have been conducted in the area of CT image automatic segmentation. Classical methods for segmentation involve applying threshold on the image intensity and considering the shape information of the interested domain [63]. Lim et al. [54] proposed a method for CT liver image automatic segmentation based on intensity distribution and priori knowledge of adjacent slices. The volume measurement utilizing this method presented an average error rate of 3% when compared to manual segmentation. Automatic segmentation methods are developed based on machine learning algorithms utilizing the prior knowledge related to the interested domain [63], which also happens in other imaging modalities [2, 56].

Magnetic Resonance (MR) Image

Magnetic resonance (MR) imaging is a non-invasive medical imaging technique that uses a magnetic field and radio waves to create detailed maps of internal body structures [8]. MR images are formed by detecting the energy emitted from the body tissues in response to the magnetic field and radio waves. MR images are commonly used to diagnose and evaluate a variety of conditions, including brain and spinal cord injuries, tumors, heart, liver, and other organs. MR is non-invasive, does not use ionizing radiation, and can provide highly detailed and accurate images that can assist in guiding medical interventions and treatments.

In 1995, Järvinen et al. [45] validated the accuracy of volume measurement with MR imaging on the right atrium, revealing that the measured volumes had a good correlation with the true volumes obtained from water displacement. Mahieu-Caputo et al. [58] measured the lung volume of fetal from MR images, and with a relationship between the lung volume size and gestational age, concluded a criterion to diagnose pulmonary hypoplasia in congenital diaphragmatic hernia. Mayr et al. [60] compared the tumor volume estimated by tracing the tumor region size in MR slices and the volume measured by a "ellipsoid volume" fomula based on diameter measurement in 60 patients with advanced cervical cancer, and concluded that the surrogate "ellipsoid volume" fomula was not able to predict the tumor size for radiation therapy (RT) treatment.

Ultrasound (US) Image

Ultrasound (US) imaging is a medical imaging technique that employs high-frequency sound waves to create images of internal body structures. To acquire ultrasound images, an ultrasound transducer will be placed on the skin over the target area.

The transducer will first emit sound waves and then change into receiving mode to collect the returning echoes relected from the organs and tissues inside the body, which will be analyzed to produce real-time images of that area. Ultrasound was first introduced in medical imaging in the 1950s, initially focusing on cardiac and obstetrical applications [20, 103]. Over time, ultrasound imaging is commonly used to diagnose and treat various diseases, and evaluate various organs, such as the heart, liver, and kidneys, as well as to monitor fetal development during pregnancy. It is non-invasive, does not use ionizing radiation, and can provide detailed images that can assist in guiding medical interventions and treatments. In the beginning, ultrasound was mainly used in hospitals, but with the development of cost-effective and portable equipment, as well as its non-ionizing and safe nature, its use has expanded to physician offices, trauma settings, and even space [15, 52, 57].

Previously, some organ volume was usually calculated with an ellipsoid formula based on the dimension estimated from ultrasound images, which simply modeled the organ as an ellipsoid with the height, width, and length of H, W, and L, respectively, and the volume can be estimated as $\frac{\pi}{6}HWL$. Bakker et al. [5] compared the renal volume determined from ellipsoid formula and that from MR imaging, concluded that the ellipsoid formula can not provide precise and consistent evaluation of the volume of kidney. In a study conducted by Tong et al. [96], 8 observers measured 15 *in vivo* prostate images through ellipsoid formula from 2D transrectal ultrasound (TRUS) images, and 3D TRUS images, proving that the 3D TRUS images have higher reliability than the ellipsoid formula estimation. Rogers et al. [81] compared the volume estimation accuracy of B-mode ultrasound, 3D tomographic ultrasound (tUS), CT, and MR imaging using an *ex vivo* porcine arteries phantom, and suggested that the tUS was the most accurate. Benjamin et al. [6] developed a system for estimating renal volumes with freehand ultrasound scan by augmenting a depth

camera and utilizing simultaneous localization and mapping (SLAM), resulting an estimation error within 3%, which was better than the ellipsoid fomula and CT images.

When compared to ultrasound, other medical imaging technologies such as MR, CT, contrast angiography (CA), and single-photon emission computed tomography (SPECT) are typically considered more expensive, involving exposure to ionizing radiation, less portable, or have a higher risk of complications from contrast media [70]. Particularly, in obstetrics and gynecology examination, according to the National Institute of Child Health and Human Development, the lower cost, real-time capability, safety, and operator comfort and experience have made ultrasound the preferred choice for imaging during pregnancy [79]. Meanwhile, ultrasound evaluation of blood flow in the fetal brain has proven to be more precise (85%) than amniotic fluid testing (76%), indicating noninvasive ultrasound as a recommended replacement for invasive testing [68]. Ultrasound shows cost-effective for diagnosing various medical conditions, including gallstones, biliary pancreatitis, and adrenal masses, and has comparable accuracy when compared with other imaging modalities such as MR and CT [26, 73]. Additionally, ultrasound can be used as a nonionizing alternative to colonoscopy and can be used for frequent reimaging in patients with inflammatory bowel disease to avoid excessive radiation exposure [39]. In many other medical conditions, such as vascular flow evaluation, cardiac exmination, and cancer diagnosis, ultrasound offers fast and accurate diagnosis, with a lower cost [7].

1.1.2 Organ Volume Measurement in Disease Diagnosis and Monitoring

Chronic Kidney Disease (CKD)

Chronic kidney disease (CKD) is a prevalent medical condition involving sustained kidney function deterioration over time, the extent of which can be determined by the glomerular filtration rate (GFR) [22,25]. Chapman et al. [12] followed 241 adults with CKD for 8 years, revealing that height-adjusted total kidney volume (htKTV) has negative correlations with GFR, and the increace of htKTV occurs earlier than the decrease of GFR. Chapman et al. [12] concluded a baseline of htTKV $\geq$ 600 cc/m as a better predictor for CKD and renal dyfunction than the previous ones.

The total kidney volume and the its growth rate are highly correlated to the advanced stages of CKD, and can also be used as biomarker in autosomal dominant polycystic kidney disease (ADPKD) predication and monitoring [109]. ADPKD is one of the major late-onset genetic multisystem disorders characterized by the enlargement of numerous kidneys cysts, leading to CKD and end-stage renal disease (ESRD) [27, 37, 109]. There is currently no cure for this disease, but treatment can slow down the progression of the disease, which can be benifit a lot from identifying the disease at early stage and tracking the progression with the biomarker of renal volume. According to a study on 241 patients with ADPKD conducted by Grantham et al. [34], the total kidney volume and cyst volume showed an exponential growth in most studied patients, with an average annual rate of 5.27$\pm$3.92.

As stated in [74], ultrasound imaging is a commonly utilized method for identifying ADPKD due to its cost-effectiveness and safety compared with MR and CT scans. However, it should be noted that in the early stages of ADPKD, small cysts may potentially go undetected in ultrasound imaging examination [74]. In earlier

years, ellipsoid formula was applied to the ultrasound images to estimate the kidney volume. However, this method lacks accuracy and reproducibility [5]. The renal volume estimation method proposed by Benjamin et al. [6] made accurate kidney volume acquisition with freehand ultrasound possible.

Chronic Liver Disease

In a study conducted in 1998, Lin et al. [55] measured the liver volume of 39 patients without liver diseases and 44 patients with chronic liver disease with CT images, revealing that the actual liver volume of alcoholic hepatitis patients was significantly higher than the predicted volume based on their height and weight, while the liver volume of chronic viral hepatitis patients tended to be lower than predicted. Hence, the actual liver volume with respect to the predicted volume can be an indicator for the cause and severity of liver diseases [55]. Polycystic liver disease (PLD) is identified by the growth of multiple cysts, gradually and continuously causing the liver enlargement [29]. Neijenhuis1 et al. [66] studied the height-corrected liver volume (htLV) of 82 patients with PLD and concluded that the patients with larger htLV would have severer symphotoms and lower life quality, suggesting that htLV should be considered as a criteria for exploring new treatments for PLD symptoms.

Currently, due to the significant increase in obesity rates, nonalcoholic fatty liver disease (NAFLD) has become more prevalent than alcoholic liver disease, and is the leading cause of both steatosis and raised liver enzymes, with a worldwide estimated prevalence of 20% in the general popular and up to 70% among type 2 diabetes mellitus patients [11,48,59]. The first identifiable stage of NAFLD, hepatic steatosis, can be recognized by the fat volume in the liver exceeding 5% [85]. Although steatosis is typically considered non-threatening regarding to the risk of developing advanced

liver disease, it is still a significant contributor to cirrhosis due to the high prevalence [99]. Because the speed of sound in fat is lower than in healthy tissue, the 3D distribution map of speed of sound in the liver can serve as a biomarker to this disease [30, 71, 72].

1.2 3D Ultrasound

Dispite the advantages of ultrasound imaging mentioned above, the conventional 2D ultrasound uses 1D ultrasound probes to provide 2D images in an arbitrary plane, but due to the lack of information on the position and orientation of the probe, it would be harder to acquire 3D information from ultrasound images. Without expert knowledge, the user might fail to acquire precise images of the region of interest (RoI). To deal with the limitation mentioned above, 3D US was proposed to help with acquiring spatial anatomic information. This provides the clinics with the 3D volume view and panoramic view of the RoI, which also provides real-time visual assistance during surgeries [43].

1.2.1 Data Acquisition

Generally, 3D ultrasound image can be collected in two different ways: (1) directly collect 3D images with 2D array transducers; (2) reconstruct 3D image with 2D images collected by 1D array transducers and the spatial information of the probe, which requires localization of the space position of the probe [43].

The 2D array transducers send out acoustic beam steering in both the horizontal and vertical dimensions to obtain a volumetric scan [108]. The elements of the 2D array transducers can generate a pyramid-shaped beam, whose echoes can be pro-

cessed to form 3D ultrasound images in real time. By modifying the phased array delays to direct and focus the beams on the ROI, the transducers can stay motionless during the scanning process [43]. This type of transducer does not need localization or reconstruction, resulting in a more precise result. However, the 2D array transducer is still facing challenges in fabrication, signal processing, time consuming and fabrication cost, and the view field is very small [43].

Besides the 2D array transducers, some mechanical 3D probes are also proposed for acquisition of 3D ultrasound images [43]. The mechanical 3D probe moves a regular 1D linear array transducer within the probe, and reconstructs 3D ultrasound images from the 2D images acquired by the 1D transducer [76]. The mechanical 3D probe does not need external localizer and are easy to operate. However, they are larger than conventional linear probes, and the static holding is required during the image acquisition, which can result in errors. Moreover, they require a specific motor to fit within the probe, which limits their versatility [43].

The convential probes can also be applied to 3D ultrasound image acquisition, but some external localization methods are required, which will be discussed in Section 1.2.2.

1.2.2 Localization Method

There are two main types of external localization methods for the regular 2D imaging transducers: (1) place the transducer in desired positions utilizing mechanical localizers (as is classified by Huang et al. [43]) or robotic arms; (2) estimate the real-time spacial placement of the transducer with some sensors [6, 92].

The mechanical localizers can provide accurate position and orientation of the transducers, and can be controlled to follow optimal paths and reduce scanning

time. The mechanical localizers are typically enormous and heavy, losing the advantage of portability of ultrasound imaging. Another popular method for transducer localization is to mount the transducer on an articulated robotic arm. The type of transducers held by the robotic arm can also be changed based on the scanning condition. However, to reduce the error introduced by the positioning errors on the joints, the robotic arm should be as short as possible, which limits the view range. Therefore, in clinical practice, freehand scans are much preferred. Additionally, a force-controlled freehand ultrasound probe proposed by Gilbertson et al. [31] can control the contact force between the transducer and body surface for desired tissue deformation, which overcomes the shortage of human operator and enhances the ability of freehand ultrasound.

According to Huang et al. [43], acoustic positioners, optical positioners, and magnetic field sensors have been applied to freehand ultrasound scanning to estimate the pose (position and orientation) of the probes in real time. Typically, these localization methods need to attach some particular labels (optical markers, electromagnetic tags, etc.) to the ultrasound transducers and track the spacial placement of these tags with an external device from some distance. These additional setups would introduce inconvenience and expensive cost to ultrasound imaging, violating the flexibility and low-cost nature of ultrasound imaging. Moreover, this type of setup requires the labels to be identifiable to the tracking device, which restricts the orientation of the probe. Meanwhile, obstacles are not allowed between the labels and the tracking device, which further limits the motion of the probe. Thus, this type of solution is not ideal in clinical use.

Alternatively, the spacial displacement of ultrasound probe can be estimated through mounting sensors on the probe. Horvath et al. [40] projected structured light onto the patient skin surface, and attached a video camera to the probe for

localization. Similarly, a designated adhesive strip with markers is affixed to the skin, and detected by a camera attached mounted on the transducer in a study conducted by Rafii-Tari et al. [78]. Although these methods were validated to be accurate and cheap, fixing markers on the skin would reduce the flexibility, restricting the probe within a small range around the designed pattern. Researchers are also exploring the use of ultrasound images for transducer localization with fully developed speckle patches and deep learning [62,77,80]. This method does not require additional device for position tracking, but the error accumulates over time, and the requirement of prior knowledge might not support the scanning in an arbitrary in real time. A total variation regularization method was proposed, which employs a variational formulation in the manifold of Euclidean transformations to improve the accuracy of pose data generated by general tracking systems [24].

Another markerless localization method of exploiting the feature of the skin surface and structure from motion (SfM) with a web camera was proposed, and the location estimation error for probe moving 200mm was tested to be 2mm [44]. This study proved the feasibility of empolying light-weight camera in 3D freehand ultrasound, but the orientation estimation still remain to be verified. Benjamin et al. [6] suggested a method of mounting a low-cost RGB-D camera and and inertial measurement unit (IMU) to a convential ultrasound transducer, allowing the localization during unrestricted freehand scanning.

1.2.3 3D Volume Reconstruction

With the localization information of the transducer when the US image is acquired, a series of 2D US images can be augmented and reconstructed into 3D images. The speed and accuracy of 3D volume reconstruction are important for achieving real-time

3D ultrasound image acquisition. According to the implementation, there are two popular types of real-time volume reconstruction algorithms: Voxel-Based Methods, and Pixel-Based Methods [43].

For voxel-based methods, a discritized voxel grid in a predefined area is defined. For each voxel in the voxel grid, the corresponding pixels in the B-mode images can be located, and a resulted value (e.g., mean value or maximum value) will be assigned to the voxel. A simple method can be find the nearest pixel to the voxel from the 2D imamges, which will be located on the voxel's normal to the nearest 2D image [82,90]. Furthermore, the voxel-based methods can be improved with interpolation between the involved pixels, mostly in a distance-weighted fashion [43, 97].

Pixel-based methods are more popular in 3D ultrasound image resconstruction, where the nearest voxels of each pixel in the B-mode images will be found, and the value of the pixel will be assigned to the corresponding voxels. In cases where multiple pixels are assigned to a same voxel, the value of the voxel can be determined with different methods, that is, the resulting value of the voxel can be the average value, the maximum value, the most recent value or the first value [16, 32, 67, 69, 97]. When some vacant voxels occur in the reconstructed volume, some gap filling methods would then be applied, such as bilinear interpolation between the two nearset nonempty voxels around the empty voxel [41].

Thus, to fast and completely acquire the 3D information of a given domain and avoid estimation for the gap filling process, especially for freehand ultrasound without expertise knowledge, path planning is needed to assist image acquisition.

1.3 Related Works

Although path planning has not been discussed in terms of guided freehand 3D ultrasound imaging, many works in different area have been done to achieve intelligent ultrasound image acquisition. Furthermore, in a broader vision area, the techniques applied in automatic view planning can be utilized in this research. Meanwhile, the problem of path planning has been well discussed in the fields of computational graphics and robtics, so some ideas in those areas can also be adpated into this problem.

1.3.1 Robotic Ultrasound Acquisition Systems

Many researchers have applied technology in robotics and computer vision to build robotic ultrasound systems for image acquisition, which makes remote medical imaging and diagnosis possible. Li et al. [53] classified the levels of autonomy in robotic ultrasound acquisition systems based on the autonomous capabilities [107]. At its most basic level, a robotic ultrasound system involves "human-robot cooperation" in which the robot and the human operator work together to move the ultrasound transducer [53, 93]. Visual servoing techniques based on ultrasound images are usually integrated in this type of system for tracking specific features and unintended patient motion [1, 88, 89]. An enhanced type of this "human-robot cooperation" system allows the robot completely control the motion of the probe, but needs the user to specify the initial position [13, 49, 61]. The second level is called task autonomy, which enables the robot automatically acquire ultrasound images along a pre-defined path. The third level, also known as conditional autonomy, refers to robotic ultrasonography systems that are capable of autonomously planning and carrying out ultrasound acquisitions without any guidance from a human operator, but still un-

der their supervision. The three most common topics discussed in this area are contact force control, scanning path planning, and image quality optimization [53], the latter two of which overlap with our interest.

The automatic acquisition of ultrasound imaging requires automatic scanning path generation. Mustafa et al. [64] developed a robotic system to automatically scan liver area. This platform utilized the natural visual marker of human body for target area localization using optical camera, and the scanning trajectory was then generated with interpolation inside the area. Huang et al. [42] introduced a depth camera to the robotic ultrasound system for 3D object surface reconstruction and determining the target area. The system would then execute linear sweeps for multiple times according to the area size. These methods, however, do not consider the body surface and tissue information, which might lead to low image quality and limited scanning applicability. Some studies utilized MRI data [38, 47, 100] for deformable registration with real-time RGB-D images of the patient, which enhances the scanning versatility. Graumann et al. [35] proposed an algorithm to generate the scanning path for a certain RoI for robotic ultrasound scan with the prior knowledge of 3D point cloud of patient surface and co-registered segmentation of the volume. In this work, the scanning orientation is fixed to the normal direction of the surface at the scanning point, and the generated paths follow the pre-defined cosine function. While these systems can easily and fast generate ultrasound scanning trajectories for a domain, its ability will be limited in clinical use due to the complexity of human body, such as the obstruction of bones and the position and shape of interested organs. Nakadate et al. [65] proposed an incremental scanning path planning method utilizing real-time US images to search for the desired view of the carotid artery. This method involves detecting carotid landmarks using real-time image processing algorithms, but may not be easily applicable to other organs without obvious

landmarks.

Another popular research topic in this area is to find optimal probe orientation and position to acquire the information of a target plane or point for better image quality. Many studies have have investigated using tomographic data from CT or MR images to predict ultrasound image quality [50, 102, 111]. A view-planning platform with automatic adaption to patient data and image-based probe tracking based on preproduced CT or MR images was proposed [101], which can find optimized probe locations for acquiring the standard view planes defined by the user. Acoustic window planning has been proposed to adjust scanning orientation and surface points for optimized acquisition at target points which were previously generated with the information from MR or X-ray images [33]. Additionally, some automatic image quality assessment techniques have also been developed for ultrasound imaging, which can be further applied to real-time ultrasound acquisition [4, 86, 106].

1.3.2 Freehand Ultrasound Trajectory Guidance

Freehand ultrasound scanning is typically performed based on the experience of the operator, and guidance or path planning for this method has not been well studied. Droste et al. [19] developed a first real-time freehand ultrasound probe movement guidance system for obstetric standard plane acquisition using a neural network called US-GuideNet. This network receives ultrasound image signals and corresponding probe motion information to guide the probe movement for acquiring a desired standard plane. Zhao et al. [110] further developed a landmark retrieval-based method utilizing a Transformer-VLAD network for real-time obstetric freehand ultrasound scanning guidance. This network can transform obtained images into global descriptors and search for the nearest landmark to localize the probe, and the

probe position is visualized accordingly. The performance of this network was tested on a clinical 3D US simulator.

1.3.3 Robotic Path Planning

The topics of path planning and object reconstruction have been widely discussed in the field of robotics. Our problem can be identified as Coverage Path Planning (CPP), which aims to find a path to pass all the points in an given area or volume [28]. This type of task occurs in many applications, such as automatic cleaning robot [17], demining robot [3] and lawn mower robot [36]. Although this problem can be further modeled as many classical path planning problems, most of them are NP-hard [28].

Based on whether the algorithm is proven to guarantee complete coverage of the given space, CPP algorithms can be classified as heuristic or complete. With the assumption of whether the environment is fully known, CPP algorithms can also be classified as offline or online [28]. Typical CPP algorithms discretize the target region into small cells [14] and represent with adjacent graph using line sweeping [28]. Wong et al. [104, 105] proposed a coverage algorithm for mobile robots that uses natural landmarks detection for topological mapping.

Considering the real-time ultrasound acquisition and the need of volume reconstruction, we are mostly interested in grid-based online CPP algorithms. Grid-based methods are the most common solution to CPP problems, which use cells with values indicating obstacle presence or free space, typically represented by squares but also triangles, so the completeness of these methods depends on resolution [23, 28]. However, grid maps have a disadvantage of increasing memory usage exponentially because the resolution remains the same irrespective of the complexity of the environment [95]. Additionally, they necessitate precise localization to ensure the consistency

of the map [51].

1.3.4 3D Reconstruction and View Planning

3D object reconstruction is a technique to build the model a real 3D object from images scanned at different perspectives. A typical object reconstruction environment setup includes a range camera, positioning system, fixtures, and the target object. The *range camera* can detect the 3D shape of the object. Since different sides of the object needs different perspective, *positioning systems* are needed to measure the relative position of camera and object. To determine the relative position, some kinds of *fixtures* are also needed [87]. View planning algorithms work in *viewpoint space*, where each viewpoint consists of the sensor pose and measurement parameters [94].

As is depicted in Figure 1-1, the workflow of object reconstruction shapes in a cycle of four steps – plan, scan, register, and integrate. *Plan* means computing a

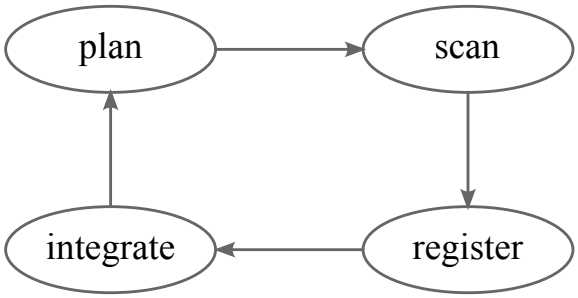


Figure 1-1: Workflow of object reconstruction [87]

sequence of next-best-view (NBV) lists. Then the sensor should move as the the NBV lists indicate, and take *scan*. In consideration of the existing of position error in positioning systems, the images acquired in *scan* should be *registered*, and then *integrated* into a single and non-redundant model. This process works in a loop until meets some criteria.

The topic of view planning developed from the demand of 3D models for complex objects, and can be applied in automated high-quality 3D object reconstruction [87]. View planning works in the step of *plan* in the reconstruction process mentioned in Figure 1-1.

Algorithms to find the NBV sequences to construct an arbitrary object with the existing of positioning error have been developed [98]. These algorithms first generate a set of candidate views in the viewpoint space, and evaluate the views with multi-resolution strategy.

1.4 Discussion

Medical imaging is an important part in clinical diagnosis and treatment, while 3D volume information plays a crucial role in diagnosis and stage identification of some diseases, such as CKD and NAFLD. Among various imaging modalities, ultrasound imaging is preferred due to its cost-effectiveness, convenience, real-time feedback and safety. However, the accurate reconstruction of 3D ultrasound images requires additional real-time location and orientation information of the transducer. A suitable solution to the ultrasound probe localization, in consistent with the low-cost and portability nature of ultrasound imaging, is considered to be mounting a low-cost RGB-D camera and IMU module on the probe.

Nontheless, expert knowledge is required to achieve the 3D information quickly and accurately. To address this issue, a path planning system can be proposed to guide the freehand ultrasound scanning to construct a 3D ultrasound image. The path planning problem in freehand ultrasound has not been adequately explored, so several related areas have been reviewd in this context. Different from the research topics discussed in Section 1.3, this project focuses on guiding non-expert users to

fast and completely acquire an entire interested volume with real freehand scan in real time.

Chapter 2

Problem Formulation for Freehand Ultrasound Scan

In general, the scanning area is discretized into voxels, which are pixels in the 2D case. As mentioned in Section 1.3.3, the aim of Coverage Path Planning (CPP) is to find a scanning trajectory that can cover almost all voxels and use the fewest scanning poses, if possible.

In an early work [10], the criteria for CPP problem, or region filling operations, were defined as follows:

1. The whole target region (or all target points) must be covered.
2. The region should be filled without overlapping paths.
3. The paths should be continuous, sequential, and non-repetitive.
4. The paths should be obstacle-free.
5. For simplicity in operation, simple motion trajectories should be used (straight lines and circles in most cases).

6. An "optimal" path is desired if possible.

However, in some cases, not all criteria can be satisfied due to the complexity of the real environment, such as the occlusions due to ribs or other anatomical structures in ultrasound imaging.

To adapt the CPP problem to freehand ultrasound, we need to consider the properties of ultrasound scanning. The complex structure with bones and entrapped air of the human body makes some areas unobservable by ultrasound. Generally, each scan motion consists of a sequence of images and probe poses. As shown in Figure 2-1, there are two common scanning types used in clinical practice: linear scan and fan scan.

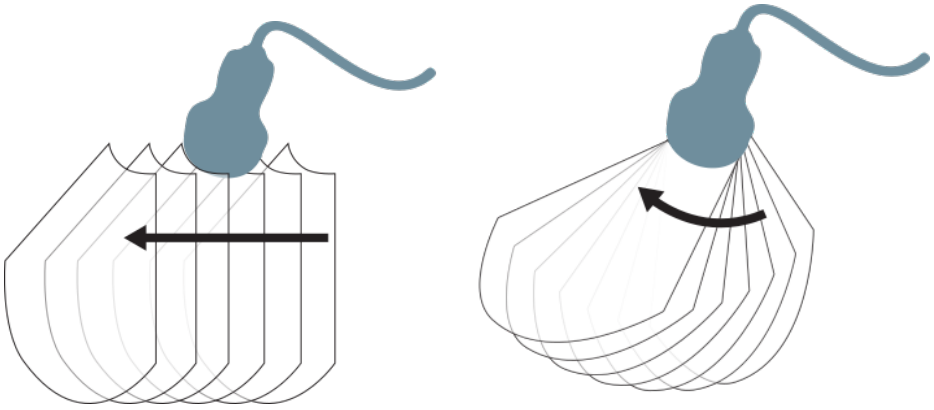
Furthermore, path planning for freehand ultrasound scan should also consider the image quality. In ultrasound scanning, the transducer emits ultrasound waves and detects the reflected waves. Given the properties of the scanning area and the transducer type, the angle between the ultrasound beam and target boundaries (e.g., tissue, organ, tumor), pressure, and wavelength would influence the reflection and, consequently, the image quality. Also, due to attenuation when ultrasound travels through the medium, the brightness and contrast of the acquired images would be low after passing through high-attenuation areas. A good path planning algorithm should also try to avoid such areas and acquire the information underneath.

Therefore, the criteria for path planning for freehand ultrasound can be summarized as follows:

1. Most of the target region is covered.

2. The paths should be continuous sequences.

3. Obstacles (e.g., bones) should be avoided or compensated for during scanning.



(a) (b)

Figure 2-1: Two commonly used scanning trajectories for commercial ultrasound probe: linear scan and fan scan. Figure 2-1(a) shows linear scan, where the probe moves translationally, while maintaining the angle between the probe and the scanning surface. Figure 2-1(b) shows fan scan, where the probe rotates about the contact point of the probe and the surface.

4. Simple motions, such as linear scan and fan scan, should be used.

5. The scanning parameters, such as the pose, should be adjusted to acquire better images if possible.

6. An "optimal" path with the smallest set of scanning actions is desirable.

We define a scanning sequence as the sequence of scanning poses, which can be either linear scan or fan scan. Two adjacent motions should be either different in type or discontinuous in space.

Moreover, for freehand scanning, the human operator cannot strictly follow the suggested trajectories, requiring additional principles to examine whether the objective is reached.

To outline the main points for this project, a basic workflow of the proposed path planning approach is described in 2-2. First, to quickly collect the information

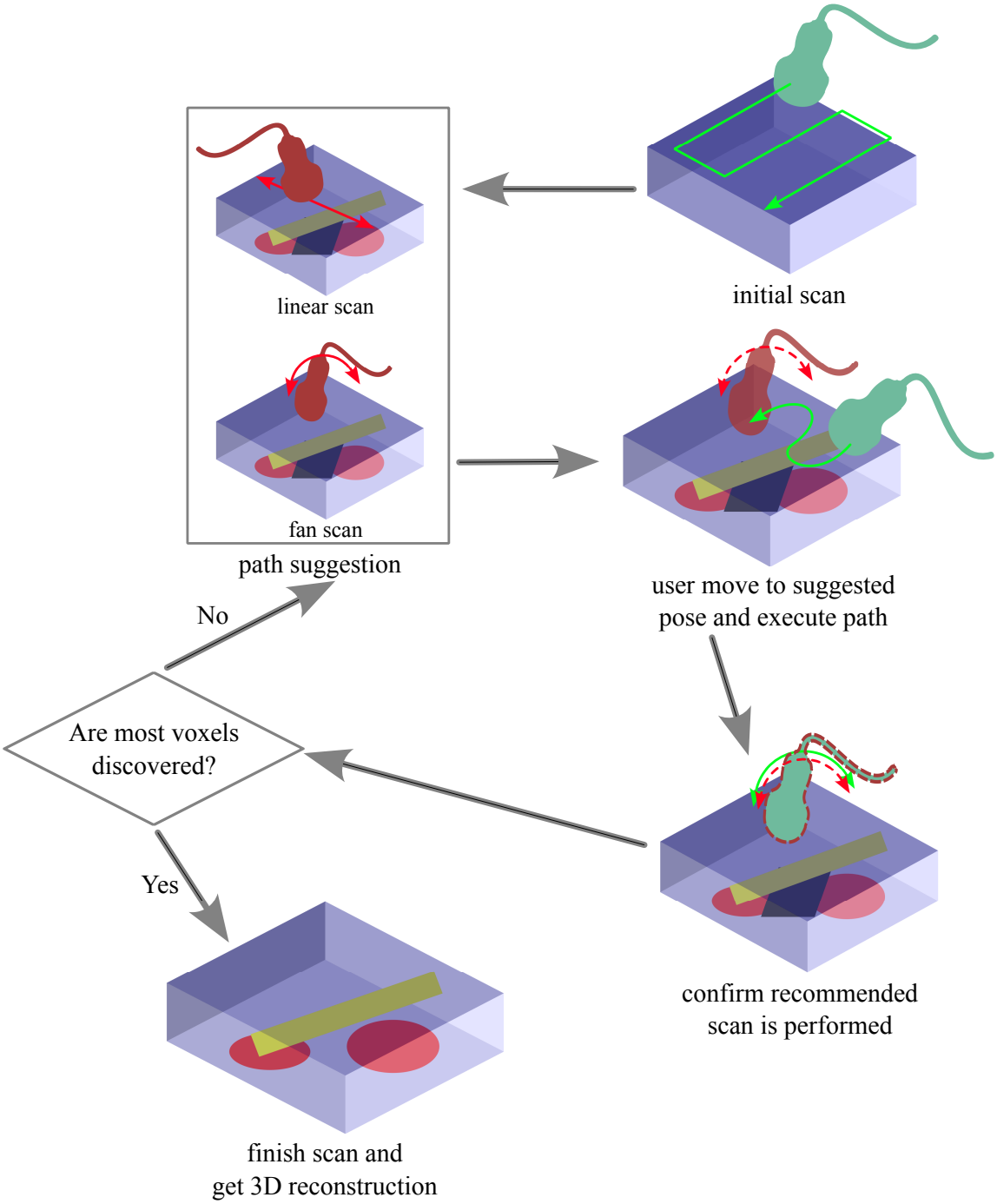


Figure 2-2: The basic outline and workflow of the proposed method

within the region of interest, a initial scan will be suggested, that is, do several linear scans within the region. Then the algorithm will start to analyze the missing area, and iteratively predict a most informative pose for the next scan to approach. The type of scan to perform around the new pose will also be suggested. The human operator can move the transducer to the suggested pose either directly by lifting the probe and moving to the position through the air, or keeping the probe in contact with the region surface. Keeping in contact is not required but suggested so that the proposed system can continue to gather additional information during the two scanning path. Ideally, the algorithm can confirm the operator has performed the desired scan at the selected pose, and then the algorithm will repeats finding the next best pose with a suggested scanning type to perform at the new pose. Once the algorithm finds most voxels have been discovered, or no new informative pose can be found, the loop will end and a 3D model of the region can be reconstructed.

Chapter 3

Next Best Scanning Pose Searching in 2D Simulation

To further study the path planning problem, we discuss the problem of finding the best scanning pose in a 2D simulation in this chapter.

3.1 Path Planning Algorithm with Greedy Policy

We employ the greedy method [21] to make every step the most informative since the inner structure of the scanning area is unknown. The strategy can be written as Algorithm 1. where the score consists of pose score and scan score. The pose score indicates the continuity of the current pose p with previous pose, while the scan score is the sum of the scores of the voxels in the predicted scanning area at the current pose p. The involved voxels of a certain scanning pose is determined with ray tracing method according to Bresenham algorithm [9]. The scan score at a single voxel is

Algorithm 1 With greedy search

initial scan: voxels $\{v_0, v_1, \cdots, v_{n_0}\} \rightarrow \mathcal{V}_{known}$
Update priority map
while known voxels percent < threshold or cannot find new pose **do**
    Find all possible candidate poses as $\mathbb{P}$
    Evaluate the scores of all poses in $\mathbb{P}$
    $p \leftarrow$ a pose with the highest score in $\mathbb{P}$
    Update $\mathcal{V}_{known}$with scan at p
    Update priority map
end while

defined based on whether it is unknown or in a high-priority area as follows:

$$s_v = \begin{cases} -100, & v \text{ nontransmittable or other prohibited zone} \\ 1, & v \text{ is unknown or in high-priority area} \\ 0, & \text{other} \end{cases} \tag{3.1}$$

Voxels in the interested area can be assigned with higher priority.

In some configurations, there might exist some unreachable areas. Hence, the pose searching process would also end if no new pixels can be discovered.

3.2 Path Planning in 2D Simple World

The algorithm was first applied to a 40mm × 40mm simple 2D world with a direct straight line, as shown in Figure 3-1, where the scanning beam will emit from the top. The target object is the ellipse in the middle, shaded by a long and thin obstacle. The "image" acquired in this case is 1D, i.e., a straight line, and the "probe" in this scenario is a point sensor that sends out a 1D single line for acquisition. Each pixel in the acquired image cooresponds to 0.1 mm in the world.

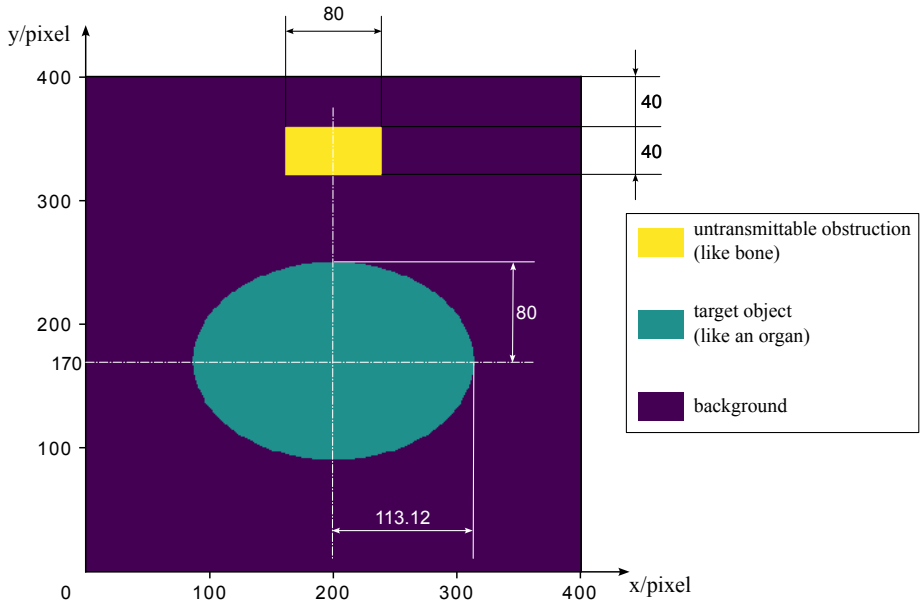


Figure 3-1: Schematic of 2D world

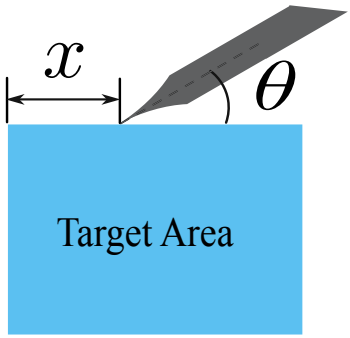


Figure 3-2: Degrees of freedom of the sensor

In this 2D world, the candidate pose sets were generated uniformly in the workspace. Note that the end of the point sensor must be located on the top surface of the target area, so the sensor has two degrees of freedom (DoF), orientation and lateral position. Therefore, with the lateral position of the tip point of the sensor (x) and the intersection angle between the sensor beam and the upper boundary (θ), the pose of the sensor can be defined, as shown in Figure 3-2.

Considering that line sensors in the real world would still have thickness, which limits the intersection angle, we limit the angle θ between 30° and 150°.

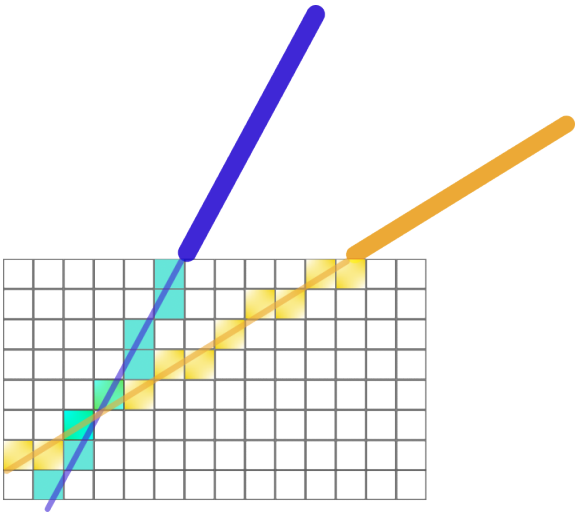


Figure 3-3: Ray tracing method in 2D case. The boxes colored in blue are the pixels reached by sensor represented by the blue line, and the yellow boxes are reached by the yellow line sensor.

For each pair of x and θ, we can retrieve the intersected voxels at each layer along the longitudinal direction, which is described in Figure 3-3.

The system was implemented on the Robotic Operating System (ROS Noetic) [91]. It consists of two main nodes, one for calculating the data received to determine the best next pose, and for visualizing the acquired information and pose. This node

sends the best pose to the other node, which runs the simulation, generates line images, and sends the image back to the calculation node. The calculation node only has knowledge of the size of the world and can only obtain line images associated with the suggested poses. In the calculation node, the size of the world is set as 40mm width $\times$ 40mm length, with the center located in $[0, 0]$. The grid size for image registration, reconstruction and path calculation is 0.1mm.

The algorithm operated as shown in Figure 3-4. As is shown in Figure 3-4(a) – 3-4(b), the direct beam performed an initial scan, and the missing area in the middle was shaded by the obstacle. Knowing the existing of obstacle, the algorithm adjusted the orientation of the beam to acquire the missing information (Figure 3-4(c) – 3-4(d)). The scanning ended in Figure 3-4(e) when most (>96%) pixels are known, where the object area was identified, and the upper and side boundaries of the obstacle area was also discovered. The bottom boundaries of the obstacle, and some area beneath the obstacle remained undiscovered.

However, the undiscovered area below the non-transmittable obstacle cannot be reached with the current configuration.

3.3 Path Planning in 2D Simulation with Ultrasound 1D propagation

This algorithm was also applied in 2D simulation with an ultrasound beam. In this case, the probe is a 1D transducer emitting an ultrasound beam, and the 1D ultrasound propagation of 1D probe was simulated with finite element method.

The wave equation for the displacement y of any particle at location x in the

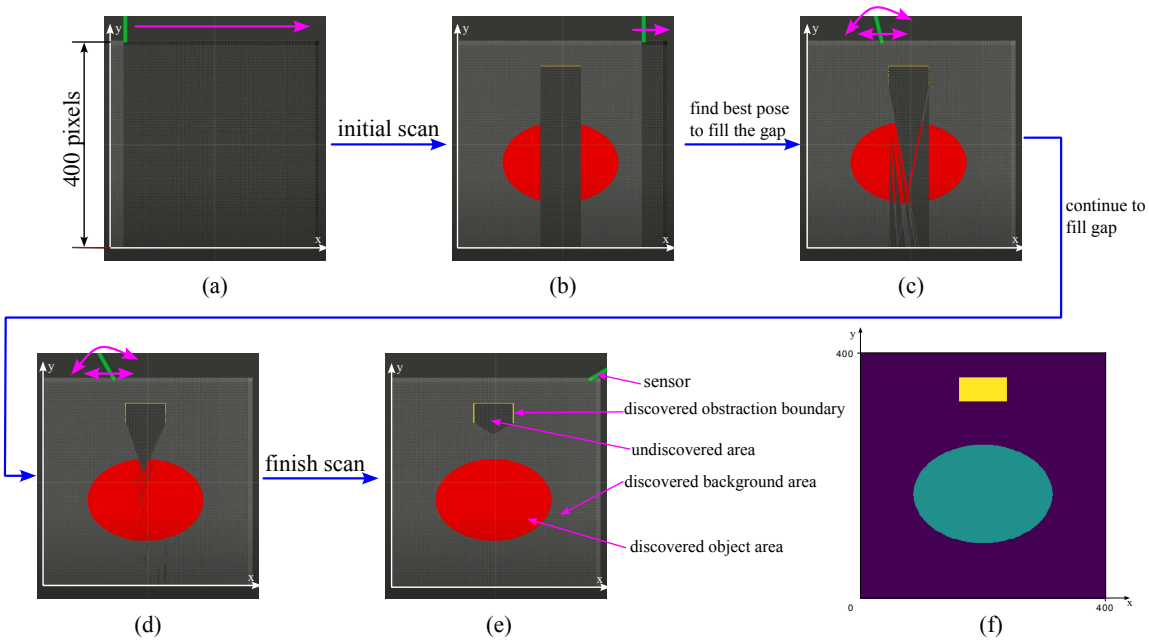


Figure 3-4: Resulting scanning path with straight line. As is legended in Figure (e), the green rod represents the sensor location; the red areas are detected to be the obejct; the white areas are recognized as the background; the black areas remain undiscovered; and the yellow lines are the boundaries of the obstacle. Figure (a) shows the initial state. Figure (b) describes the information acquired during initial scan. Figure (c) – (d) indicate the algorithm was trying to fill the gaps caused by the obstacle during initial scan, and Figure (e) depicts the final state. The 2D world schemetic is displayed in Figure (f) again for comparison, the detailed legend and dimensions can be found in Figure 3-1.

ultrasound field, at time t, can be expressed as:

$$\rho \frac{\partial^2 y}{\partial t^2} = B \frac{\partial y^2}{\partial x^2} \tag{3.2}$$

where ρ is the density of the medium and B is the bulk modulus. The speed of sound can be derived as $c = \sqrt{\frac{B}{\rho}}$. When accounting for damping with a damping ratio β, the wave equation becomes:

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial y^2}{\partial x^2} - \beta \frac{\partial y}{\partial t} \tag{3.3}$$

Polynomial interpolation is used on both sides of the wave equation, yielding the following equation for a selected simulation time step Δt and grid distance Δx:

$$\begin{aligned} \frac{y(x,t+\Delta t) - 2y(x,t) + y(x,t-\Delta t)}{\Delta t^2} =& c^2 \frac{y(x+\Delta x,t) - 2y(x,t) + y(x-\Delta x,t)}{\Delta x^2} \\ & - \beta \frac{y(x,t+\Delta t) - y(x,t-\Delta t)}{\Delta t} \end{aligned} \tag{3.4}$$

This equation can be used to calculate the displacement y at the next time step, which is given by:

$$\begin{aligned} y(x,t+\Delta t) =& \frac{1}{1+\frac{1}{2}\beta\Delta t} \Big[c^2 \frac{\Delta t^2}{\Delta x^2} y(x+\Delta x,t) + c^2 \frac{\Delta t^2}{\Delta x^2} y(x-\Delta x,t) \\ & + (2 - 2c^2 \frac{\Delta t^2}{\Delta x^2}) y(x,t) - y(x,t-\Delta t) + \frac{1}{2}\beta\Delta t y(x,t-\Delta t) \Big] \end{aligned} \tag{3.5}$$

After iterating Equation 3.5 for a finite number of steps, the particle displacement

field for 1D ultrasound propagation can be estimated. Using the equation of motion,

$$-\frac{\partial P}{\partial x} = \rho_0 \frac{\partial y^2}{\partial t^2} \tag{3.6}$$

the pressure signal P at the receiver can be derived. Then we can apply descritized Fourier Transform to the signal, filter the information out with a low pass filter, and detect the envelope of the signal, as is shown in Figure 3-5. This method allows for the generation of 1D B-mode images.

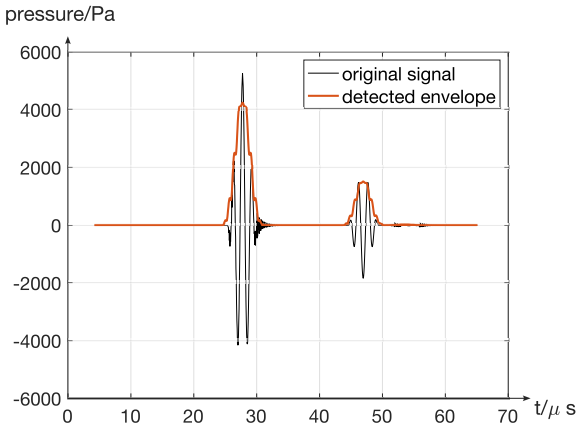


Figure 3-5: Envelope detection of the simulated signal

Based on those geometrical information, ultrasound 1D propagation would be simulated and the obtained signal would then be sent out for further path planning. To simplify the calculation related to geometrical features, the target object was designed as a circle with a diameter equal to half of the whole square, as shown in Figure 3-6. Given the pose of the sensor, the length of the ultrasound beam traveling within the Region of Interest (RoI) can be determined, along with the location and length of the intersection between the beam and the target. It is also possible to determine whether the ultrasound beam reaches the high-attenuation area and the

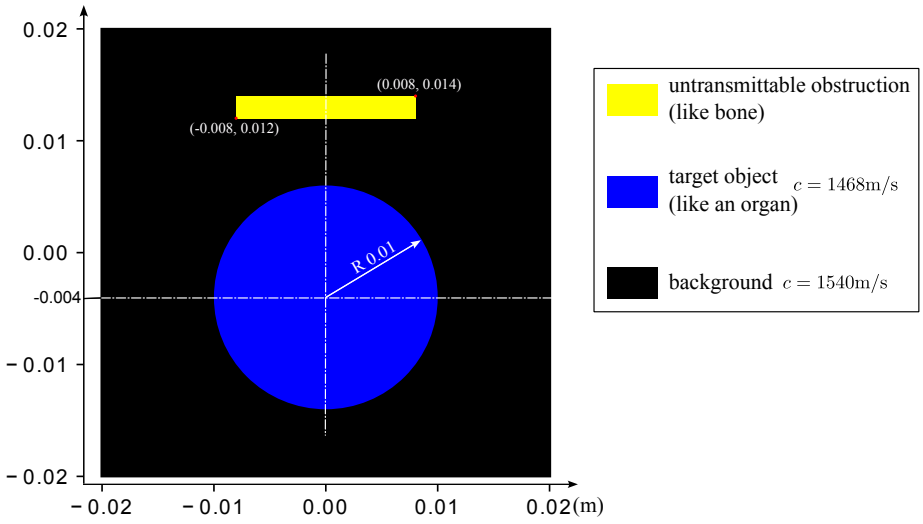


Figure 3-6: 2D world configuration for ultrasound simulation. The white strip on the top represents a high ultrasound attenuation area. The blue circle represents the target object, and the black square serves as the RoI to be explored.

length the beam travels within the area.

The simulation was conducted using an excitation signal frequency of 1 MHz. The density of the surrounding area was set to 1000 kg/m^3, while the bulk modulus was chosen as 2.3716×10^9 N/m^2, resulting in a speed of sound of $c_0 = 1540$ m/s in this area. The energy transmitted through the white area was assumed to be low and was ignored. The density of the blue round area in the middle was set to 1.1 times that of the background density, resulting in a speed of sound of $c_t = 1468.33$m/s in the object area. The damping ratio for attenuation was set to $\beta = 0.1c^2$.

The sampling rate was set to 20 MHz, which determined the time step Δt, while the space grid distance was $\Delta x = c_0 \Delta t$. The setup for pose calculation and image simulation was similar to the simple world simulation described in Section 3.2, with the simulation node in this case simulating 1D ultrasound propagation and generating 1D B-mode images.

Under the simulation conditions described above, the algorithm performed as depicted in Figure 3-7. As with the straight beam case, the simulated 1D ultrasound probe first executed an initial scan, and the missing area in the middle was shaded by the obstacle, as shown in Figure 3-7(a) – 3-7(b). Knowing the existense of obstacle, the algorithm started to adjust the orientation of the ultrasound beam to acquire the missing information, as shown in Figure 3-7(c) – 3-7(d). In Figure 3-7(e), when most pixels were discovered, the algorithm could not find a new scanning pose where the ultrasound beam can reach more than 10 new pixels, so the searching process ended. Similar to the previous example, the object area was identified, and the upper and side boundaries of the obstacle area was also discovered. The bottom boundaries of the obstacle, and some area beneath the obstacle remained undiscovered.

In this best scanning pose searching problem, which was developed from coverage path planning (CPP) problem, the image quality was not considered. Thus, at the left and right corners of the circle in the middle, the pixels were discovered during initial scan when the ultrasound beam direction was almost tangential to the object surface. Under this scanning condition, the reflection at the object boundaries was low, resulting in unclear images around the two corners. However, the algorithm considers those pixels as "discovered", so they would not need to be examined again with appropriate view angels. Similarly, the pixels at the bottom boundary of the target object were also explored during the initial scan. However, due to attenuation, the image intensity around that area was also low.

3.4 Results Discussion

The previous section presented two simple 2D cases provide an intuitive understanding for path planning in trajectory-guided ultrasound scanning. In both cases, an

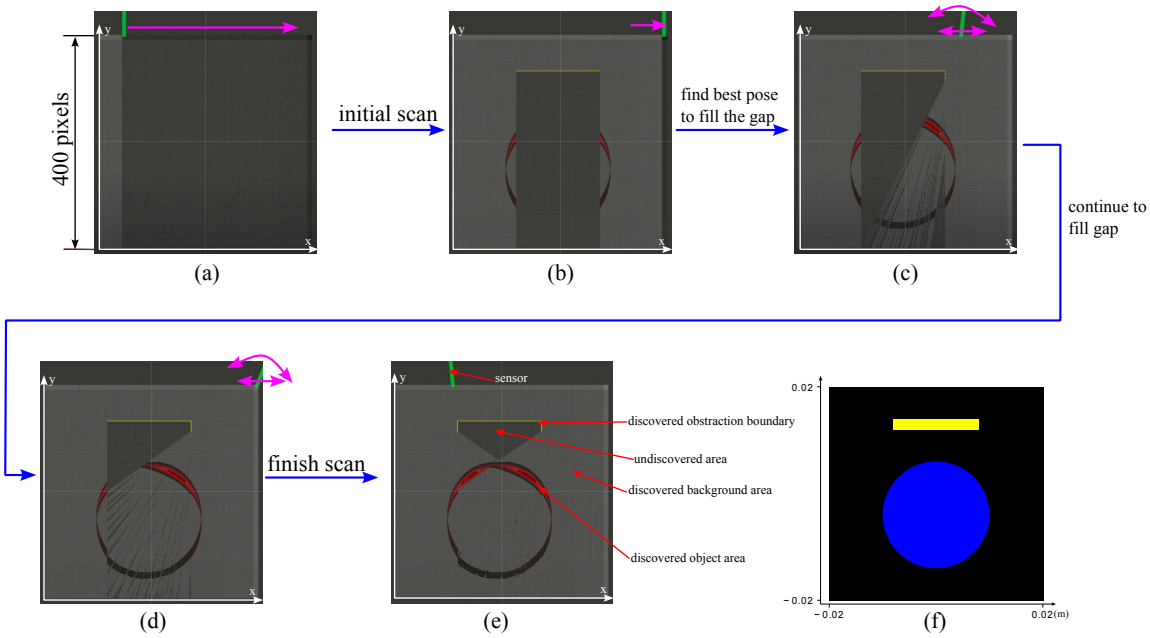


Figure 3-7: Resulting scanning path with ultrasound 1D propagation. As is legended in Figure (e), the green rod represents the sensor location; the red (and dark red) areas are detected to be the obejct; the white areas are recognized as the background; the gray areas remain undiscovered; and the yellow lines are the boundaries of the obstacle. Figure (a) shows the initial state. Figure (b) describes the information acquired during initial scan. Figure (c) – (d) show the algorithm was trying to fill the gaps caused by the obstruction area during initial scan, and Figure (e) is the final state. The 2D world schemetic for this simulation is again included in Figure (f) for comparison, the detailed legend and dimensions can be found in Figure 3-6.

area beneath the obstacle remained undiscovered after the guided scan. This area is the blind spot shaped by the occlusion area, and can only be accessed when the intersection angle θ is less than 30° or greater than 150°, which violates the workspace definition, as is described in Figure 3-8. In real ultrasound scans, the angle θ can

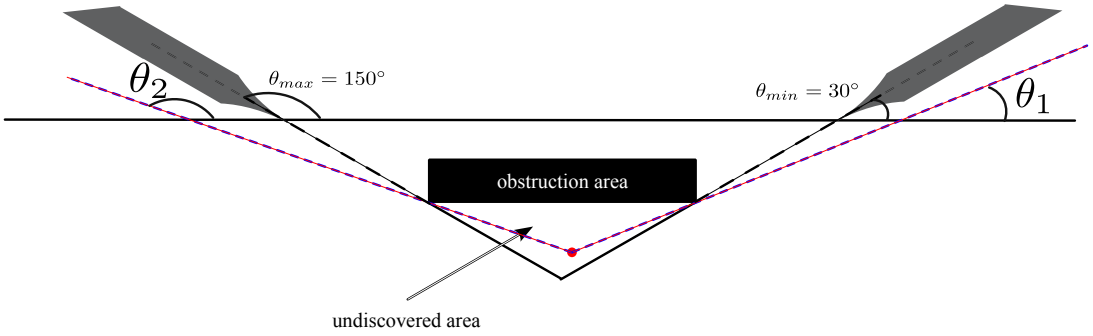


Figure 3-8: Blind spot under current configuration. Obviously, the red point can only be accessed at $\theta_1 < 30°$, or $\theta_2 > 150°$, which exceed the angle limit.

also influence image quality, restricting the angle to a narrow range. Therefore, some undiscoverable area may still exist even with appropriate trajectory guidance.

In these two experiments, the best next scan pose was calculated only as a single pose, resulting in a pose sequence that may be discontinuous at some points, i.e., the probe may jump frequently from one side to the other. To avoid frequent and severe changes in pose, an additional score indicating the distance of the new pose to the previous one was included.

It should be acknowledged that the greedy policy employed in this work may lead to local minima. However, since we can only optimize the next pose based on past information, the greedy policy is acceptable. Additionally, the inherently safe nature of ultrasound imaging makes the total scanning time a non-sensitive factor, so a time-optimal solution is not necessary, and redundant information can actually improve image quality.

Chapter 4

Path Planning Algorithm Description

In Chapter 3, we presented two simple 2D examples to illustrate the fundamentals of the path planning algorithm for ultrasound scanning. In this chapter, we describe how we adapt and apply this algorithm to real-time ultrasound scanning with 2D images in order to construct a 3D volume, and the hardware system required for its implementation.

4.1 Hardware System

4.1.1 Probe Localization

In order to estimate position and orientation of the probe in real time, we use the Intel Realsense depth camera D435i (Intel, Santa Clara, California, USA) for the implementation of the path planning algorithm. This camera is a RGBD camera equipped with an inertial measurement unit (IMU). Figure 4-1 shows the appearance

of the camera and its associated coordinate system.

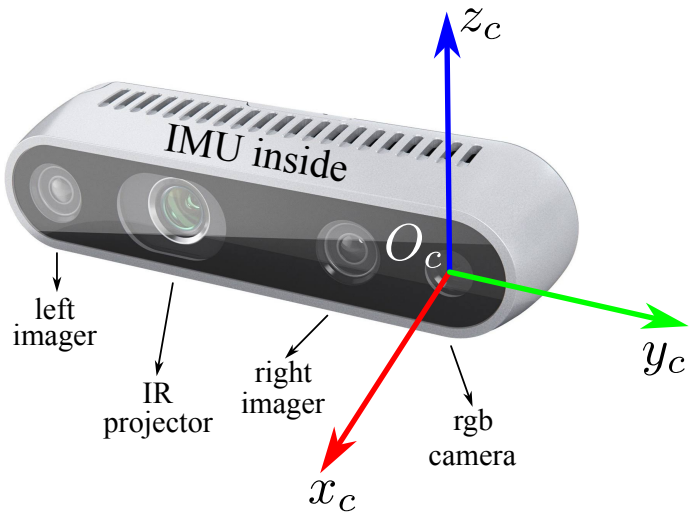


Figure 4-1: Intel Realsense depth camera D435i. This camera has two camera sensors ("left imager" and "right imager"), an Infrared Projector ("IR projector"), and an RGB camera (labeled as "rgb camera"). It provides depth data with a resolution of 1280×720 at a frame rate of 90 fps, and RGB data with a resolution of 1920×1080 at 30 fps.

The camera has two stereo camera sensors ("left imager" and "right imager"), an Infrared Projector ("IR projector"), and an RGB camera (labeled as rgb camera"). The stereo depth module (two imagers and the IR projector) has a field of view (FOV) of $87° \times 58° \times 95°$ (horizontal, vertical, and diagonal, respectively) and is available in all lighting conditions, with a resolution of 1280 $\times$ 800. The IR projector sends out static infrared patterns at the wavelength of 850nm $\pm$ 10 nm to the environment to help the depth measurement. The projector has a FOV of $90° \pm 3° \times 63° \pm 3° \times 99° \pm 3°$. The camera can provide steoro depth data with a maximum resolution of 1280×720 at a frame rate of 90 frames per second (fps) and RGB data with a resolution of up to 1920 $\times$ 1080 pixels at a frame rate of 30 fps. It is capable of measuring depth between the range of 0.2m to over 3m. This camera is equipped with an additional

inertial measurement unit (IMU) module, enabling the estimation of its 6 degrees of freedom (DoF) space pose, and the integration error of the IMU data would be corrected with the optical information.

Figure 4-1 shows the appearence and associated coordinate system of this camera. By default, the origin O_c of this camera coordinate system is located at the center of the RGB camera, and the x_c, y_c, and z_c axes point towards the forward, left, and upright sides of the camera, respectively. As is displayed in Figure 4-2, the camera is attached to the ultrasound probe with some 3D-printed connectors.

In order to simplify the computation, we have defined a specific coordinate system for the probe as shown in Figure 4-2. The x_p axis is aligned with the transducer traversal element arrangement direction, the y_p axis is aligned with the longitudinal direction, and the z_p axis is aligned with the upright direction of the probe. Moreover, we have relocated the origin to the bottom center point of the probe. Therefore, the rotation matrix R_p^c from the camera frame to the probe frame can be written as given in Equation 4.1.

$$R_p^c = R_z(-90^\circ) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.1}$$

The translation between the probe and the camera, denoted as d_p^c, with respect to the camera coordinates depends on the shape and size of the connectors.

4.1.2 Coordinate system

This section provides an overview of the coordinate systems used in the proposed system. The coordinate systems are illustrated in Figure 4-3, which shows four main frames: the world frame, the camera frame, the ultrasound probe frame, and the

Figure 4-2: The ultrasound probe with D435i camera attached by 3D-printed connectors. The coordinate system $O_p - x_p y_p z_p$ defines the probe frame, where x_p axis points to the traversal element arrangement direction of the transducer, which will be the width direction u of the acquired B-mode image; y_p axis is parrallel to the longitudinal direction; z_p axis is the upright direction, whose negative direction is the height direction v of the resulting image. O_p is the bottom center point of the probe, which coinsides with the top center of the image.

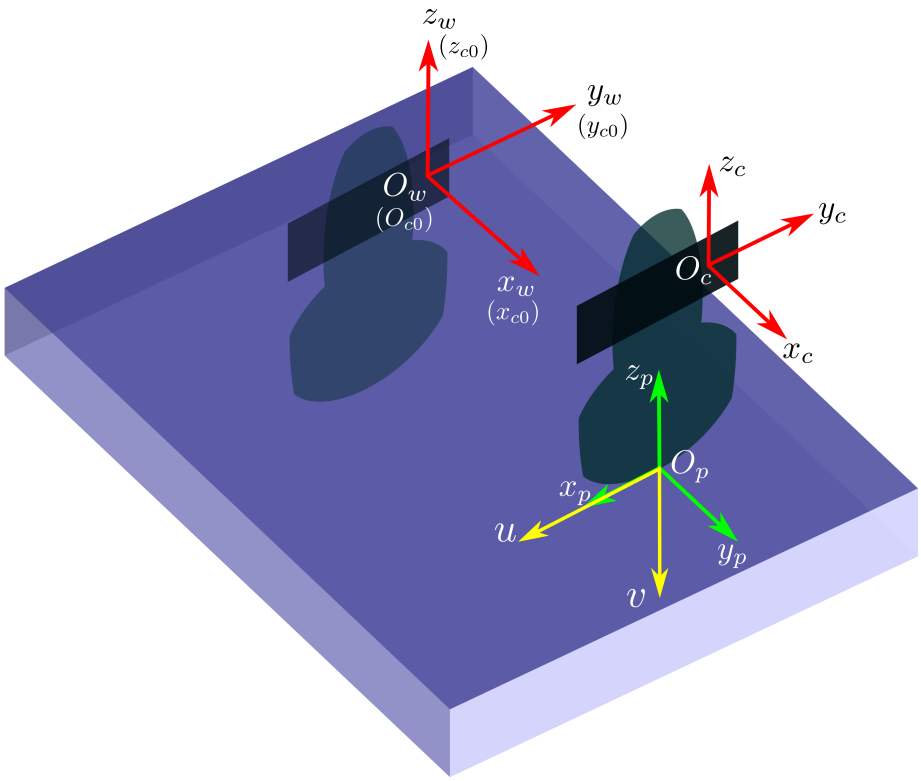


Figure 4-3: Coordinate systems associated with the D435i camera, world, and ultrasound probe. The red frame $O_w - x_w y_w z_w$ represents the world frame. Typically in the experiments, the origin of world frame is located around the center of the top surface of the region of interest (the blue box). The green frame $O_p - x_p y_p z_p$ represents the ultrasound probe frame, as is defined in Figure 4-2. The red frame $O_c - x_c y_c z_c$ indicates the camera frame described in Figure 4-1. The yellow frame $(\mathbf{u}, \mathbf{v})$ indicates the ultrasound B-mode image frame, whose width direction is $\mathbf{u}$, and height direction is $\mathbf{v}$.

ultrasound B-mode image frame. The world frame ($O_w - x_w y_w z_w$) is established by the camera when the measurement program starts running, and its origin and coordinates are defined by the starting pose of the camera. The camera frame ($O_c - x_c y_c z_c$) information can be obtained directly from the depth camera with respect to the world coordinate system. The probe frame ($O_p - x_p y_p z_p$) has a static rotation from the camera frame given in Equation 4.1, along with a rigid translation d_p^c. The ultrasound B-mode image frame is denoted by the yellow frame ($\mathbf{u}, \mathbf{v}$), where $\mathbf{u}$ and $\mathbf{v}$ are the width and height directions of the image, respectively.

Let us denote the time at camera initialization as t_0. If we place the probe normally onto the region of interest (RoI) surface at t_0, the world coordinate system will be established at frame $O_w - x_w y_w z_w$ (also known as the camera frame at t_0, i.e., $O_{c_0} - x_{c_0} y_{c_0} z_{c_0}$) in Figure 4-3. Then the rotation and translation from the world frame to the probe frame at t_0 ($O_{p_0} - x_{p_0} y_{p_0} z_{p_0}$) will be equal to R_p^c, and d_p^c. At an arbitrary time t after t_0, let T_c^w be the pose information of the camera frame read from the camera, then the probe pose (rotation matrix R_p^w and displacement d_p^w) in the world frame can be written as

$$\begin{bmatrix} R_p^w & d_p^w \\ 0 & 1 \end{bmatrix} = T_p^w = T_c^w T_p^c = \begin{bmatrix} R_c^w & d_c^w \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_p^c & d_p^c \\ 0 & 1 \end{bmatrix} \tag{4.2}$$

For a pixel (u_n, v_n) in the B-mode image acquired at t, its corresponding location in the probe frame is $[u_n s_u, 0, -v_n s_v]^T$, where s_u and s_v are the resolution of the ultrasound image in $\mathbf{u}$ and $\mathbf{v}$ direction, respectively. Therefore, its position $[X, Y, Z]^T$

in the world frame can be written as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R_p^w \begin{bmatrix} u_0 s_u \\ 0 \\ -v_0 s_v \end{bmatrix} + d_p^w \tag{4.3}$$

To aviod repetitive computation of matrix multiplication, we can derive the positions of other pixels from the top center point in the image, which coincides with the origin of prboe coordinates, i.e., d_p^w. Let the column vectors of R_p^w be $R_p^w = [\mathbf{n_p^w}, \mathbf{t_p^w}, \mathbf{b_p^w}]$, then we can easily map the unit vector of the B-mode image, $\Delta\mathbf{u}$ and $\Delta\mathbf{v}$, to the world frame:

$$\begin{aligned} \Delta\mathbf{u^w} &= s_u \mathbf{n_p^w} \\ \Delta\mathbf{v^w} &= -\, s_v \mathbf{b_p^w} \end{aligned} \tag{4.4}$$

Similar to the ray tracing method based on Bresenham algorithm [9], but in two directions, the corresponding points of other pixels can then be computed with $\Delta\mathbf{u}$, $\Delta\mathbf{v}$, d_p^w, and their relative displacement to the top center point in the image frame.

4.1.3 Verasonics Ultrasound Machine

The ultrasound machine used in this project is Verasonics Vantage 256 (Verasonics, Kirkland, Washington, USA). This machine is a research-grade ultrasound machine that provides high flexibility and programmability for a wide range of ultrasound applications. It has a maximum sampling rate of 80 MHz, which enables high-resolution imaging, and can be programmed to operate in both B-mode (brightness mode) and RF (radiofrequency) mode in real time, and supports a wide range of transducer types and frequencies. and velocity, making it a valuable tool in car-

diovascular research. During the data acquisition of the proposed system, the data would be streamed out in real-time through Ethernet to process and record.

4.1.4 Data Streaming and Communication

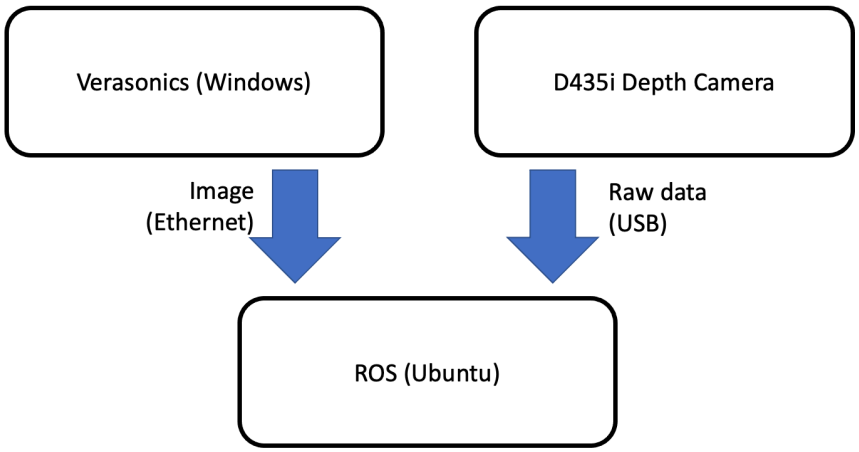


Figure 4-4: Data transfer and communication in the system

As is depicted in Figure 4-4, in the proposed system, the Verasonics machine is connected to a Windows desktop and the acquired B-mode images are transferred to a laptop running ROS through Ethernet. The pose of the probe is estimated by the RealSense D435i camera, which is also connected to the laptop. The coordination and synchronization of different processes and calculations is handled by ROS, which ensures seamless communication between different components of the system.

However, since the acquired 2D images are reshaped into 1D for streaming out, and the data is received and processed by line in ROS, there is a time delay between receiving the first byte of data from the Verasonics machine and successfully processing all the lines to form an image. Therefore, the pose information related to the image should be collected at the time when first byte of image information is

received to ensure accurate and timely synchronization of data.

4.2 Target Area Identification and Workflow Description

As is described earlier in Figure 1-1, although this proiect focuses on path planning for the scanning pose, as part of this process we need to decide a fast solution to the identification of important area to move the process forward.

In this project, the main focus is on path planning for the scanning pose, but it is still necessary to identify important areas for online and real-time path planning, as is described earlier in Figure 1-1. Specially, when working with a typical B-mode image acquired in an ultrasound phantom and the identification of important areas, we use segmentation of the area with inclusions and shaded areas from the background. It is possible to complete this segmentation task roughly based on the image intensity at each pixel.

However, the high-reflection area and high-attenuation area are not absolutely correlated, so the high-attenuation area can not be directly identified through intensity differences alone. Given that the shadow in the image must be caused by the high-attenuation area above it, the high-attenuation area, or obstacle to avoid, can be located based on the shaded area. For example, we can traverse an image by column, and for a single column, as is framed out in Figure 4-5, if a high-intensity area occurs in a shallow pixel (area A in the Figure), and is followed by a low-intensity area in lower pixels (area B), then we can label area A as the high-attenuation area.

The basic methodology to fill the shaded area would be similar to the 2D cases

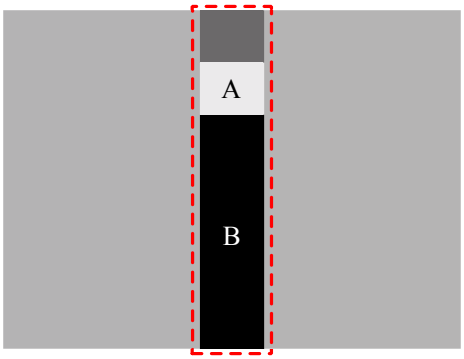


Figure 4-5: Simple intersity criteria for extracting the obstraction area. When traversing an column in the image from top to the bottom, if we first find a high-intensity area (labeled as A) followed by some low-intensity area (labeled as B), then we can consider area A as the obstruction area, or the obstacle area.

discussed in Chapter 3, while for the blurred area, the number of voxels involved would be relatively small compared with the total voxel number. Therefore, a more efficient solution is to record the blurred areas once identified and find a better scanning pose for the areas that remain blurred after most voxels are discovered. Then the path planning with greedy search adapted to real-world freehand ultrasound is described in Algorithm 2.

The evaluation of candidate poses and estimation of better scanning pose for blurred areas will be discussed in the following sections.

Algorithm 2 Path planning for freehand ultrasound

initial scan: voxels $\{v_0, v_1, \cdots, v_{n_0}\} \rightarrow \mathcal{V}_{known}, \{b_0, b_1, \cdots, b_{m_0}\} \rightarrow \mathcal{V}_{blurred}$
Update priority map
while known voxels percent < threshold or cannot find new pose **do**
    Find all possible candidate poses as $\mathbb{P}$
    Evaluate the scores of all poses in $\mathbb{P}$
    $p_o \leftarrow$ a pose with the highest score in $\mathbb{P}$
    Decide a best path $p_1, p_2, \cdots, p_s$ around p_o, and wait for the user to execute
    while Path $p_1, p_2, \cdots, p_s$ is not finished **do**
        Acquire a new image A and pose p
        Update $\mathcal{V}_{known}, \mathcal{V}_{blurred}$with image A at pose p
        Update priority map
    end while
end while
for each b in $\mathcal{V}_{blurred}$ **do**
    if b is still blurred **then**
        Find a better scanning pose for b
    end if
end for

4.3 Multi-resolution Strategy For Path Planning in Large Area

4.3.1 Candidate Pose Generation and Evaluation

Similar to the 2D cases, we first need to find all possible poses for next scan for further evaluation. The spacial placement of transducer has 6 DoF, the location, $[x, y, z]$, and the orientation, roll (ϕ), pitch (θ), yaw (ψ). However, not all parameters are free to change in order to acquire informative ultrasound images.

First of all, the transducer surface emitting the ultrasound should be in contact with the obejct surface (phantom surface, or skin surface when acquiring in vivo images). Therefore, the displacement of z would be determined by x and y with

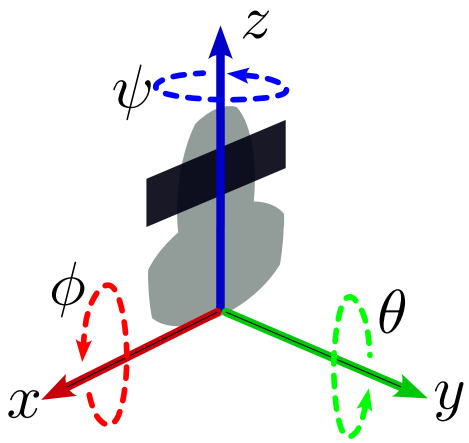


Figure 4-6: The six degrees of freedom (DoF) of the probe.

the surface function $f(x, y)$. Specially, for a typical ultrasound phantom whose top surface is flat, z will be a constant value. The variation of angular position θ does not change the plane the transducer is examining, which might provide more information, but the new information would be limited. Hence, we would not consider changing this angular position in path planning. The choice of ϕ might introduce interspace of air between the transducer and the target surface, adversely affecting the image quality, so this angle should be limited within a reasonable range. Since the variation of ϕ coinsides with fan scan, and doing linear scan with a particular value of ϕ would be hard to execute, we would first decide the displacement of x, y, and ψ, and then choose from linear scan or fan scan.

Thus, the candidate pose set can be generated by evenly sample of $[x, y, \psi]$ in the workspace while keeping $z = f(x, y)$, $\theta = \phi = 0$.

The evaluation of a candidate pose is similar to the 2D case, that is, predicting the voxels can be accessed by imaging at the candidate pose, but the Bresenham algorithm [9] is applied to both $\mathbf{u}$ and $\mathbf{v}$ directions utilizing Equation 4.4. The related voxels are located and evaluted by column in the predicted image. For a

single column, the 2D version ray tracing method performs the same with the 1D beam, counting the new voxels and the voxels with high proirity the column beam reaches, and once the column beam reaches the obstacle area, a negative score will be assgined to the column, and the 2D version ray tracing method will move to the next column for evalution. The final score for this candidate pose is the sum of the scores of all the columns.

4.3.2 Multi-resolution Strategy

Moving from 2D domain to 3D domain, the number of voxels and candidate poses increases sharply, which is hugely time-consuming. To expendiate the path searching process, a multi-resolution strategy described in [98] is utilized and adapted. That is, first evaluate the candidate poses in a lower resolution R_1, and find the top n_1 candidate poses for next resolution layer. Repeat the evaluation with new resolution and reduced candidate pose set until achieving the final resolution. Suppose there are k resolution layers, and resolution R_k is the final resolution we reconstruct the 3D model with, then with the resolution R_k, we need to find the top one pose, and this pose would be the output of this multi-resolution best pose searching process. Namely, the next best pose searching process in the scenario of freehand ultrasound scan can be described in Algorithm 3.

4.4 Normal Vector Estimation for Pose Adjustment of Target Scanning Point

As is shown earlier in the 2D ultrasound simulation in Section 3.3, even when most of the voxels in the RoI are discovered, there might still exist some blurred area where

Algorithm 3 Next best pose searching for freehand ultrasound with multi-resolution strategy

Generate candidate pose set $\mathbb{P}_0$
for $i = 1 : k - 1$ **do**
  for each candidate pose p in $\mathbb{P}_{i-1}$ **do**
    Evaluate p with resolution R_i
  end for
  Find top n_i poses $\{p_{i,1}, p_{i,2}, \cdots, p_{i,n_i}\}$
  $\mathbb{P}_i \leftarrow \{p_{i,1}, p_{i,2}, \cdots, p_{i,n_i}\}$
end for
Evaluate each p in $\mathbb{P}_k$ with resolution R_k
Find top pose p_{best}
return p_{best}

the image is not clear and could be improved with an additional pose. In ultrasound imaging, the angle of incidence represents the angle at which a beam deviates from a line that is perpendicular to the tissue surface. Typically, a larger angle of incidence means in fewer reflected sound waves to the transducer, causing a darker and more unclear image. Therefore, in order to obtain a clear ultrasound image of some inner body structures (e.g., organ shape), the optimal orientation of the ultrasound beam should be perpendicular to the surface of the target inner body area, i.e., follow the direction of the normal vector at the inner surface [18]. To satisfy this criterion, the path planning algorithm should suggest a probe pose that is close to the normal direction of some unclear points. This requires normal vector estimation.

Considering the fact that human operators cannot follow the instructions precisely, here we can use a simple normal vector estimation method – using the normal vector of a plane tangent to the surface to approximate the normal vector to a point at the surface. Therefore, this problem can be turned into a least-square plane fitting estimation problem. Thus, the solution to this problem can be reduced to an analysis of the eigenvectors and eigenvalues (or PCA Principal Component Analysis) of a

covariance matrix created from the nearest neighbors of the query point introduced in [83].

For each inquired point P_q in the point cloud, we can find its k neighbor points, $P_1, P_2, \cdots, P_k$, and the centroid point $\overline{P}$ of the k neighbors can then be estimated by taking

$$\overline{P} = \sum_{i=1}^{k} P_i \tag{4.5}$$

Then the covariance matrix at point P_i can be computed as

$$C = \frac{1}{k} \sum_{i=1}^{k} \xi_i (P_i - \overline{P})(P_i - \overline{P})^T \tag{4.6}$$

where ξ_i is the possible weight for P_i, if any, or would be set to 1. For a 3D point cloud, $C \in \mathbb{R}^{3\times3}$, so we can find the eigenvalue and eigenvectors of C, λ_j and $\mathbf{v}_j$, $j = 1, 2, 3$, through

$$C\mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad j = 1, 2, 3 \tag{4.7}$$

Let λ_1 be the smallest eigenvalue, then the normal vector $\mathbf{n}_q$ at point P_q can be approximated by $\mathbf{v}_1$ or $-\mathbf{v}_1$.

Since the sign of the normal vector $\mathbf{n}_q$ given by PCA method is ambiguous, Rusu [83] also stated that the direction can be determined with the viewpoint v_p when it is known. Then normal vector should satisfy that

$$\mathbf{n}_q \cdot (v_p - P_q) > 0 \tag{4.8}$$

Applying this normal estimation method into this path planning problem, we utilized an open-source library, Point Cloud Library (PCL) [84], for point cloud processing. During the image registration and reconstruction process, the information

of the surface points would be recorded in a point cloud, and the locations of all the blurred areas would also be tracked. After most region of the RoI is discovered and filled, check each voxel in the blurred areas. If no better image has been acquired at this voxel (denote the center point of this voxel as P_v), estimate the normal vector $\mathbf{n}_i$ at this voxel, then in most cases we can find the intersetion point P_s of this vector $\mathbf{n}_i$ and the outer surface, or find the point P_s satisfying

$$P_s \in P_{\text{outer surface}}, s.t. \min_{P_s} \{\arg(\overline{P_s P_v}, \mathbf{n}_i)\} \tag{4.9}$$

Then we can find the normal vector $\mathbf{n}_s$ at P_s of the outer surface with the information gathered through initial scan, and choose the plane decided by $\mathbf{n}_i - P_s - \mathbf{n}_s$ for a new scan.

Chapter 5

Algorithm Validation and Experiments

In the last chapter, we described the basic workflow of the proposed algorithm. A system for real-time freehand ultrasound scanning guidance was developed and tested on an ultrasound phantom.

5.1 Experimental Setup

5.1.1 Ultrasound phantom for experiments

To evaluate the functionality of the proposed framework, especially the obstacle (high-attenuation area) avoidance feature, an ultrasound phantom was fabricated, as is shown in Figure 5-1. The phantom consisted of a potato and an agar cube placed in the lower layer. The agar cube serves as the target "organ" for examination, as it exhibits strong reflection and good penetrability in ultrasound. The potato has a high reflection at the surface, but high-attenuation inside. To introduce a shaded

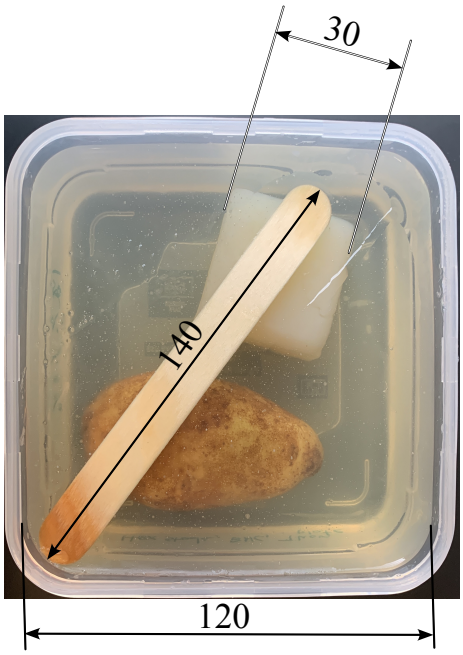


Figure 5-1: The ultrasound phantom fabricated for testing and validation

area, a wooden stick was place on the top of the two objects.

Since the phantom was handmade, it contains some defects, such as air holes, which would introduce noise to the acquired image.

5.1.2 Visualization System for Probe Guidance and Real-time Reconstruction Preview

A crucial element of real-time probe guidance for freehand scanning is a user-friendly visualization system. To this end, we developed a visualization system to display suggested scanning poses, and the reconstructed 3D image in real time based on RViz [46], a 3D visualization tool of ROS, which is shown in Figure 5-2. The 3D view in the middle of the interface provides information about the phantom, probe, and camera, as well as the data reconstructed from the acquired images. During

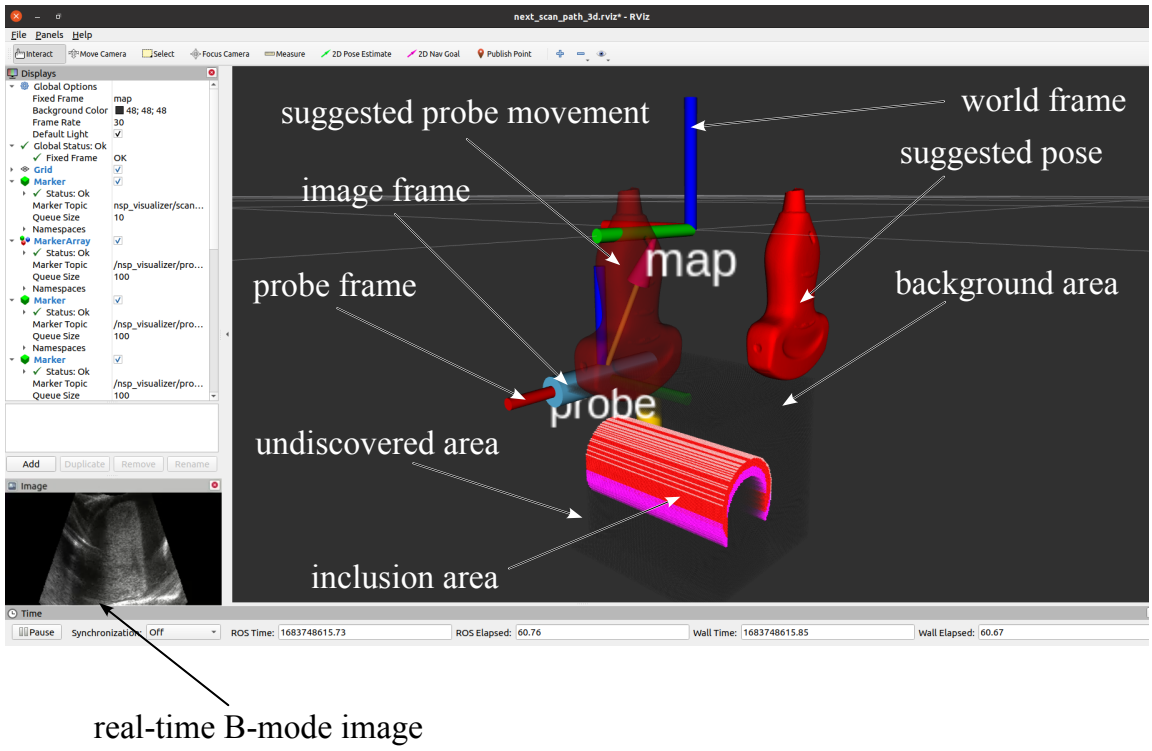


Figure 5-2: Visualization system for real-time probe guidance and reconstruction. As is labeled, the voxels in the undiscovered area will be invisible, while voxels in the inclusion area will appear in red colors, the obstacle area in yellow, and the background area in black. The ultrasound probe mesh in red placed on the top of the phantom indicating the suggested pose; the transparent probe mesh moves along the suggested path. The real-time ultrasound image streamed from the Verasonics machine is displayed in the lower left corner.

the real-time path planning and image acquisition process, once an B-mode image is received by the computational system, the visualization system will change related voxel colors in the screen based on the information from the image. The voxels in the undiscovered area will remain invisible, while voxels in the inclusion area will appear in red colors, the obstacle area will be yellow, and the background area will be in black. The method for area classification and corresponding voxel determination will be discussed in Section 5.1.5. The camera pose and probe pose are also displayed in the 3D view. The ultrasound probe mesh in red placed on the top of the phantom indicating the most informative pose at this stage, guiding the user to move the probe, while the transparent probe mesh moves continuously, showing the suggested movement of the probe and directing the user to perform either a linear scan or fan scan. A screen in the lower left corner displays the ultrasound image streamed from the Verasonics machine.

The system has a feature of automatically confirming whether the suggested path was executed by evaluating the Euclidean distance between the current pose and the suggested pose. However, during the experiments, the position and orientation estimation of the probe exhibited lage error at times, so the feature was disabled to aviod false judgement. The user will be required to move the probe until the frame in the visualization platfrom indicating the probe aligns with the suggested probe pose, and once the user confirms that the suggested path has been achieved, they can report to the computation system for next path computation and suggestion (by pressing the enter botton in the terminal in current setup).

5.1.3 Parameter Selection

Most parameters in the Verasonics system are tunable. The high voltage parameter controls the amplitude of the measuring wave pulse, which directly influences the energy reflected. Higher value of this parameter will result in higher contrast and intensity of the B-mode image. Noticing that the defects in the phantom can absorb energy and hence decrease the image intensity, we set the high voltage to the highest possible value, 50 V, or the B-mode images acquired might not be bright enough for the following assessment and processing. The defects in the phantom also cause some detectable reflection in the background area, which interferes the image segmentation. Since most of the reflection caused by the defects is not strong, a parameter of the "sensor cutoff" was set to 0.9 to filter out such reflection and form a clear boundary of the inclusion area. Despite the fact that the images acquired in real time may not be as clear as those from commercial ultrasound machines equipped with enhancement algorithms, the B-mode images streamed out from the Verasonics machine were normalized to a range of 0 to 255 in intensity and had a size of 459 pixels in height and 394 pixels in width. The GE C1-6D Ultrasound Probe (GE Healthcare, Chicago, IL, USA) was utilized in the experiments, with each pixel in the B-mode image representing a square with a side length of 0.0238mm. To facilitate faster path calculation, we need to prevent the total voxel number being too high, so the resolution of the voxel grid was set to be much lower than that of the ultrasound image, that is, 1mm, 2mm, and 5mm in the three trials, respectively.

The initial pose of the camera defines the world frame, so at the initialization of the camera, the projection point of camera frame origin to the top surface of the phantom, i.e., the center of the RGB camera, would be positioned near the center of that surface. At the same time, the probe would be held normally to the

phantom top surface. Certainly, these operations can not be executed accurately by human hands, but the system can tolerate small deviations. To prevent the probe from hitting the edge of the phantom container, the target area in the calculation node was defined as a box with demension of 120mm in length, 120mm in width, and 60mm in height. The displacement of the probe frame in camera frame was measured as $[-0.0398, -0.0175, -0.1065]$ (m), so the center of the phantom was defined as $[0, 0, -0.13]$ (m) in the world frame to ensure the origin of probe frame to be close to the top surface of the phantom.

The best next scanning pose searching was conducted in a limited workspace, as described in Section 4.3.1. To avoid unnecessary computation, the position coordinates x and y were confined within a 100mm $\times$ 100mm square centered at the center of the phantom top surface. The rotation angle θ was discretized steps of 6° within the range of $[0, 180^\circ]$. To determine whether to perform a linear or fan scan, the algorithm calculated the total score of both types of scans by evaluating their respective forward and backward 15-step trajectories. For linear scans, the step size was set to the length of a single voxel (1mm, 2mm, and 5mm in the three trials, respectively), whereas for fan scans, the step size was 5°.

5.1.4 Real-time Image Registration and Post-processing Volume Reconstruction

It is important to consider computation time when developing a real-time guidance system. The algorithm was tested on the system three times with different voxel length of 1mm, 2mm, and 5mm. The multi-resolution strategy introduced in Section 4.3.2 was implemented to optimize the speed of pose searching process, which involves adding additional layers with voxel lengths that are 2 and 4 times to the original

layer. To speed up the image registration process, a pixel-based method in the latest value fasion was adopted, and with the method described in Equation 4.4, reduplicative transformation can be prevented and time-efficient image registration can be achieved.

In practice, the real-time 3D image is primarily used as a tool for guiding probe movement rather than as diagnostic evidence, so it is not necessary to have precise segmentation and reconstruction during the scanning process. Therefore, to obtain a better 3D image, a precise post-processing volume reconstruction is required. Since the main objective of this project is not 3D image reconstruction, a simple reconstruction method was employed after the path planning process. This involved converting each pixel with information in the 2D B-mode image into a space point in a point cloud. By doing so, a dense point cloud can be generated for areas with inclusion.

5.1.5 Information Extration in Acquired B-mode Images

Due to the defacts in the handmade ultrasound phantom, extracting information from the B-mode images can be challenging. To simplify the image processing during the path planning, only the obstacles are identified through a sudden intensity change along the $\mathbf{v}$ direction in the ultrasound B-mode images, as is described in Section 4.2. The inclusion area related voxels are also colored during the real-time reconstruction based on the intensity value in the corresponding pixels of the B-mode images, but the algorithm does not distinguish the inclusion from the background area for further path planning. In post-processing volume reconstruction, the cubic inclusion and the obstacle were segmented through the intensity difference after filtering.

Figure 5-3 displays two original B-mode images acquired during the experiments and their segmentation with the simple intensity-based method. Both images were

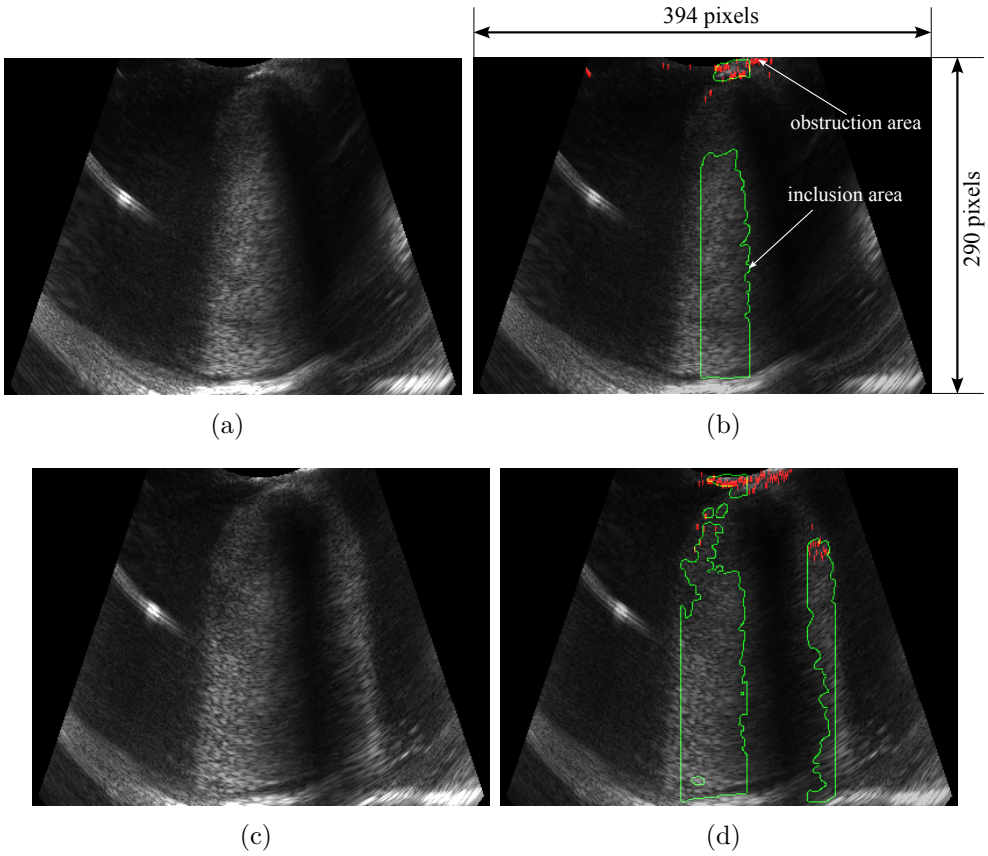


Figure 5-3: Simple segmentation for two typical B-mode images in the experiment (top and bottom edges were cropped). The red points refer to the segemented obstruction area, while the green lines represent the boundaries of the segmented inclusion area. Figure 5-3(a) and Figure 5-3(c) are original B-mode images, while Figure 5-3(b) and Figure 5-3(d) are resulting segmented images with inclusion and obstacle areas for Figure 5-3(a) and Figure 5-3(c), respectively.

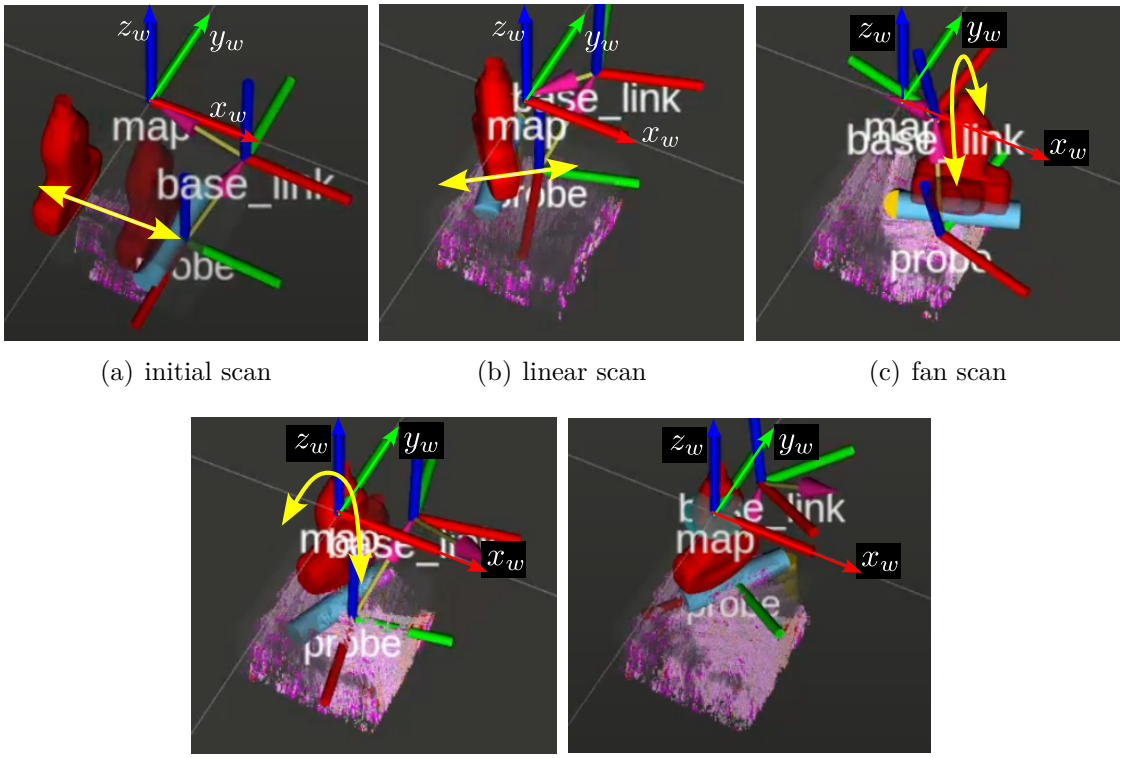
captured over the cubic inclusion and obstacle. It can be observed that there are irregular dark patterns within the area of the cubic inclusion, and also some false bright areas on the top and the left. The bottom of the container also showed high reflection in the B-mode images. A Gaussian filter with the window size of 5×5 was applied to the images, but only part of the pattern was removed, and the images started losing details.

Hence, a pixel would be considered to be inside the inclusion when the maximum intensity value within a 5×5 window around it passed a threshold of 60. Some column and row criteria were also applied to remove the erroneous bright area and the container bottom. While for the obstacles, since the number of pixels involved was much less than the inclusion, the original image was utilized.

Obviously, some areas are still erroneously segmented, such as some inclusion areas near the bottom were not identified in both Figure 5-3(b) and 5-3(d), and there are outliers in the detected obstacle points, but the resulting images are adequate to reconstruct a rough 3D point cloud.

5.2 Experimental Results

When using a voxel length of 1mm, the entire process took approximately 5 minutes. The suggested scanning trajectories of the proposed system together with the path planning and reconstruction process are depicted in Figure 5-4. The placement of the voxel grid in the screen was consistent with the phantom orientation in the real world, where the obstacle extended from the right inner corner to the left outer corner. In Figure 5-4(a) shows the initial scan suggested by the system, which involved covering the area roughly with a simple linear scan. Then we reported the completion of initial scan through hitting the enter botton in the terminal when we thought the initial



(a) initial scan (b) linear scan (c) fan scan

(d) fan scan (e) final state

Figure 5-4: The suggested scanning path sequence in the experiment when voxel length was 1mm. The voxels in the undiscovered area is invisible, while voxels in the inclusion area appear in red and pink colors, the background area in gray, and the obstacle area is yellow. The "map" frame represents the world frame, which is still and was defined when at the camera initialization; the "base_link" frame indicates the pose of the camera; the "probe" frame is also included; the two axes in yellow and blue denote the image frame that is being processed. The ultrasound probe mesh in red placed on the top of the phantom indicating the most informative pose at this stage, guiding the user to move the probe, while the transparent probe mesh shows the suggested movement of the probe, directing the user to perform either a linear scan or fan scan. Figure 5-4(a) shows the initial scan. Figure 5-4(b) suggests a linear scan. Figure 5-4(c) and Figure 5-4(d) are both fan scans. Figure 5-4(e) is the finishing state.

scan was accomplished. Next, the system identified a missing area in the initial scan, so a linear scan was suggested and drawn in the screen, as is shown in Figure 5-4(b). After this path was reported as finished, the system determined that the most informative scanning path for the next scanning step was a fan scan near the obstacle, as is indicated in 5-4(c). Then another fan scan near the obstacle was also suggested, as is shown in Figure 5-4(d). Finially, most of the voxels were filled, and the final voxel grid in the real-time path planning and reconstruction system is shown in Figure 5-4(e).

Utilizing the information extraction and image reconstruction method for post-processing described in the previous section, the final point cloud resulted from this experiment is displayed in Figure 5-5 with three views. Although outliers caused by position and segmentation errors exis, the shape of the cubic inclusion and the obstacle stick can still be recognized.

The second validation experiment was conducted on the same phantom with a voxel length of side of 2mm for real-time path planning and image exhibition. The scanning process was completed in approximately 4 minutes, and the suggested trajectory sequence and resulting real-time mesh are presented in Figure 5-6.

Similar to the first experiment, the state in Figure 5-6(a) indicated the initial scan, and after the achievement of initial scan, the system determined the next scanning path should be a fan scan near the obstacle, as is instructed in Figure 5-6(b), which was followed by another fan scan (Figure 5-6(c)), two successive linear scan (Figure 5-6(d) and 5-6(e)), one more fan scan (Figure 5-6(f)) and another linear scan (Figure 5-6(g)). Eventually, the final state was achived, as is displayed in Figure 5-6(h), and the final point cloud after post-processing is displayed in Figure 5-7.

As previously discussed in the 1mm case, the outliers are mainly caused by the localization and segmentation error, but the shapes of the cube and stick are still

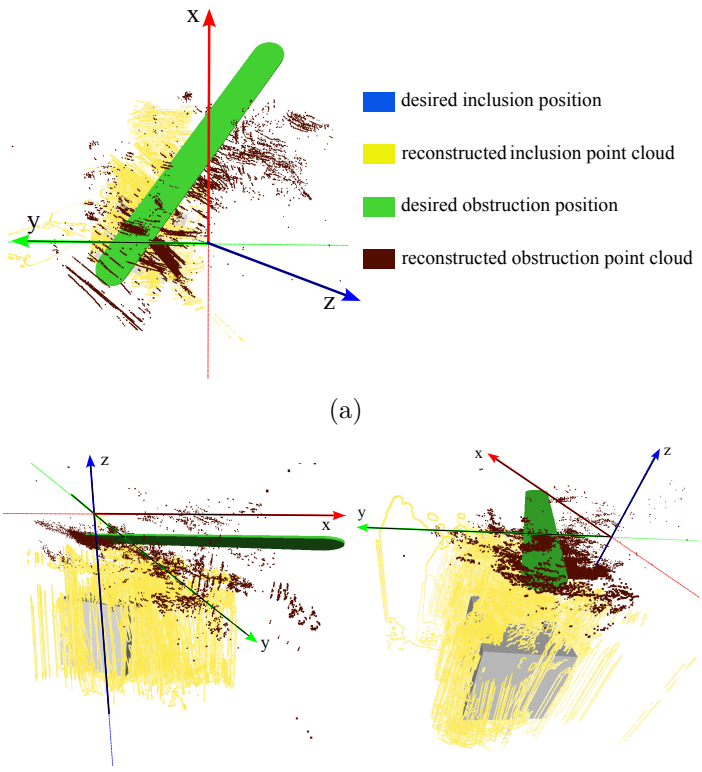
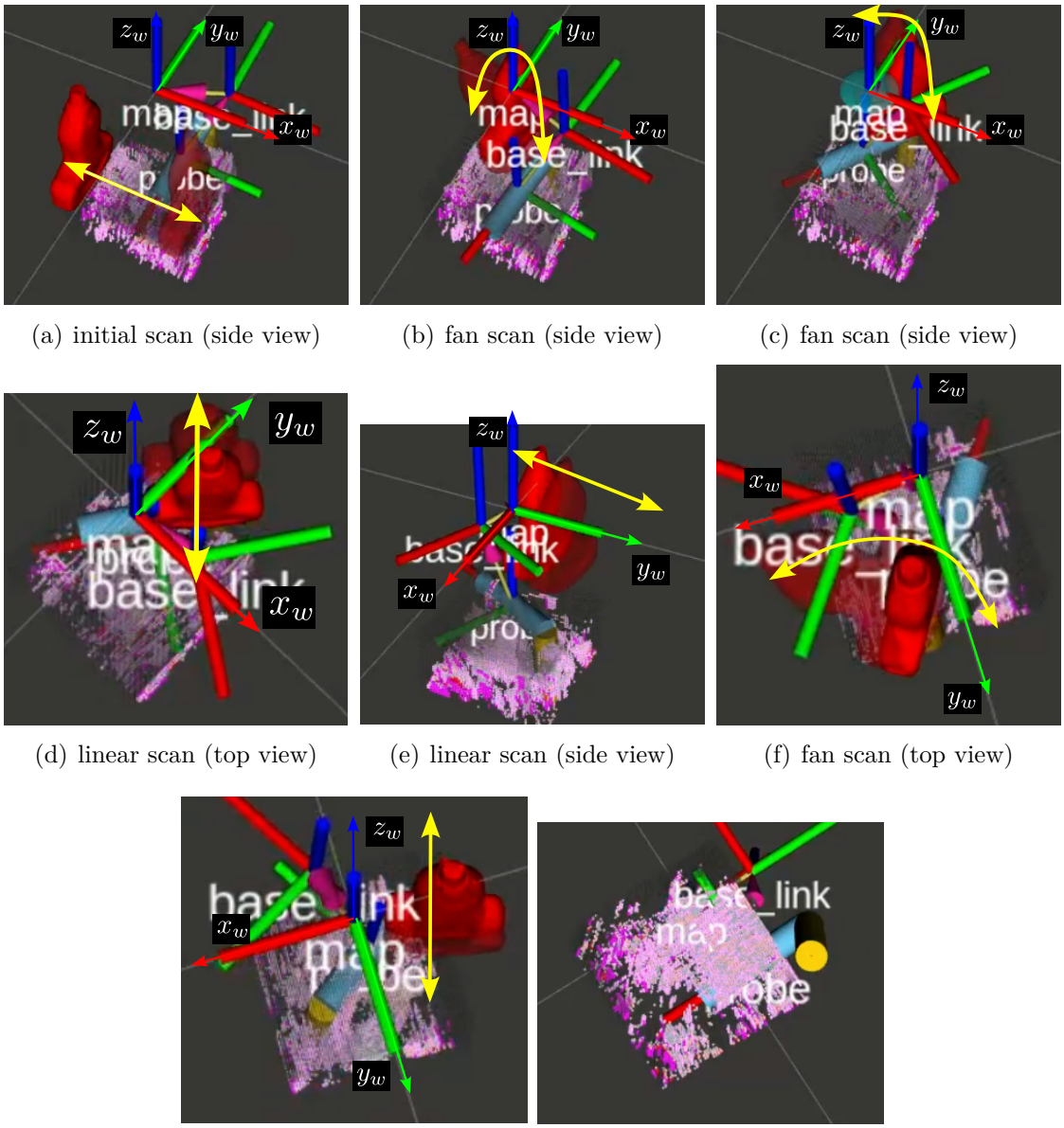


Figure 5-5: Three views of the resulting point cloud. Figure 5-5(a) is the top view, and Figure 5-5(b) and 5-5(c) are side and front view, respectively. In the figures, the green rod and blue cube are the estimated true position of the obstacle and the inclusion; while the red point cloud and the yellow point cloud are the reconstructed point of the obstruction and the inclusion, respectively.



(a) initial scan (side view) (b) fan scan (side view) (c) fan scan (side view)

(d) linear scan (top view) (e) linear scan (side view) (f) fan scan (top view)

(g) linear scan (top view) (h) final state (bottom view)

Figure 5-6: The suggested scanning path sequence in the experiment when voxel length was 2mm. Figure 5-6(a) shows the initial scan, and Figure 5-6(b) – 5-6(g) suggests fan scan, fan scan, linear scan, linear scan, fan scan, and linear scan, sequentially. Figure 5-6(h) is the final state of the real-time path planning experiment viewed from the bottom.

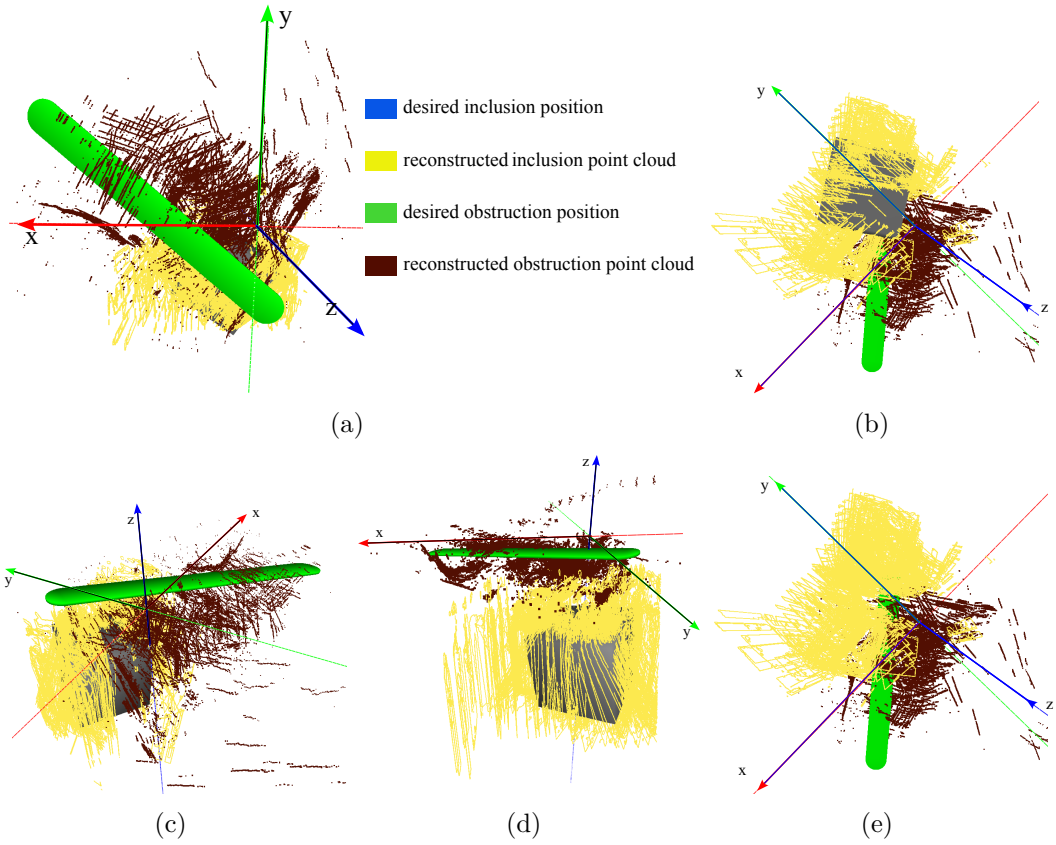
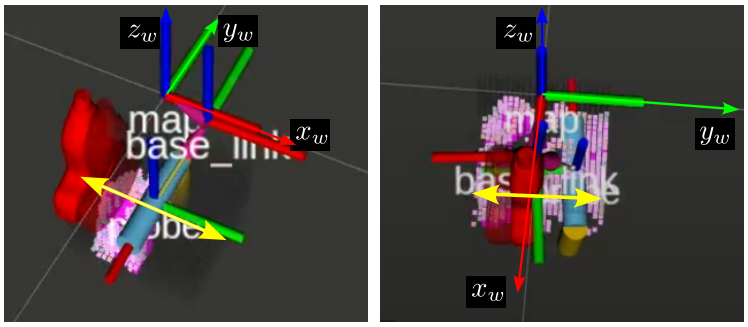
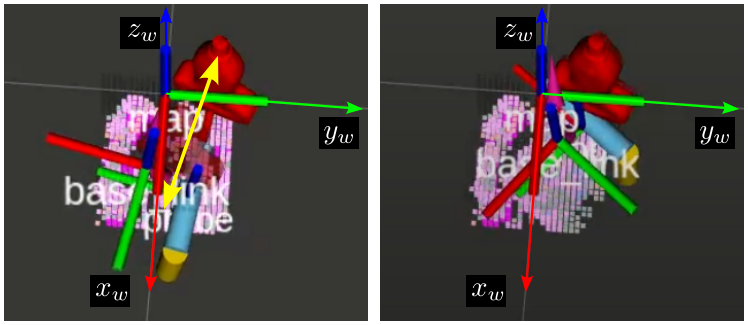


Figure 5-7: The resulting point cloud when voxel length was 2mm. Figure 5-7(a) and Figure 5-7(b) are the top view and bottom view, respectively, while Figure 5-7(c) and 5-7(d) are side and front view, respectively. Figure 5-7(e) is the same view with 5-7(b), but the true inclusion mesh was removed for better observation. Some outliers were removed and the true object meshes were enlarged for better observation.

recognizable in the post-time reconstructed point cloud. It is worth noting that a small interspace can be observed under the obstacle, which is clearly visible in the bottom view (Figure 5-7(e)).



(a) initial scan (side view) (b) linear scan (top view)



(c) linear scan (top view) (d) final state (top view)

Figure 5-8: The suggested scanning path sequence in the experiment when voxel length was 5mm. Figure 5-8(a) shows the initial scan. Figure 5-8(b) and 5-8(c) are both linear scans. Figure 5-8(d) is the finishing state. Figure 5-8(b) – 5-8(d) were viewed from the top.

In the third trial, a larger voxel side length of 5mm was used, and the process was completed within 3 minutes. The planned path sequence and resulting real-time mesh are displayed in Figure 5-8. Similarly to the previous trials, the initial scan was performed first, as described in Figure 5-8(a). Then two linear scans were determined and executed, as indicated in Figure 5-8(b) and 5-8(c). The freehand ultrasound

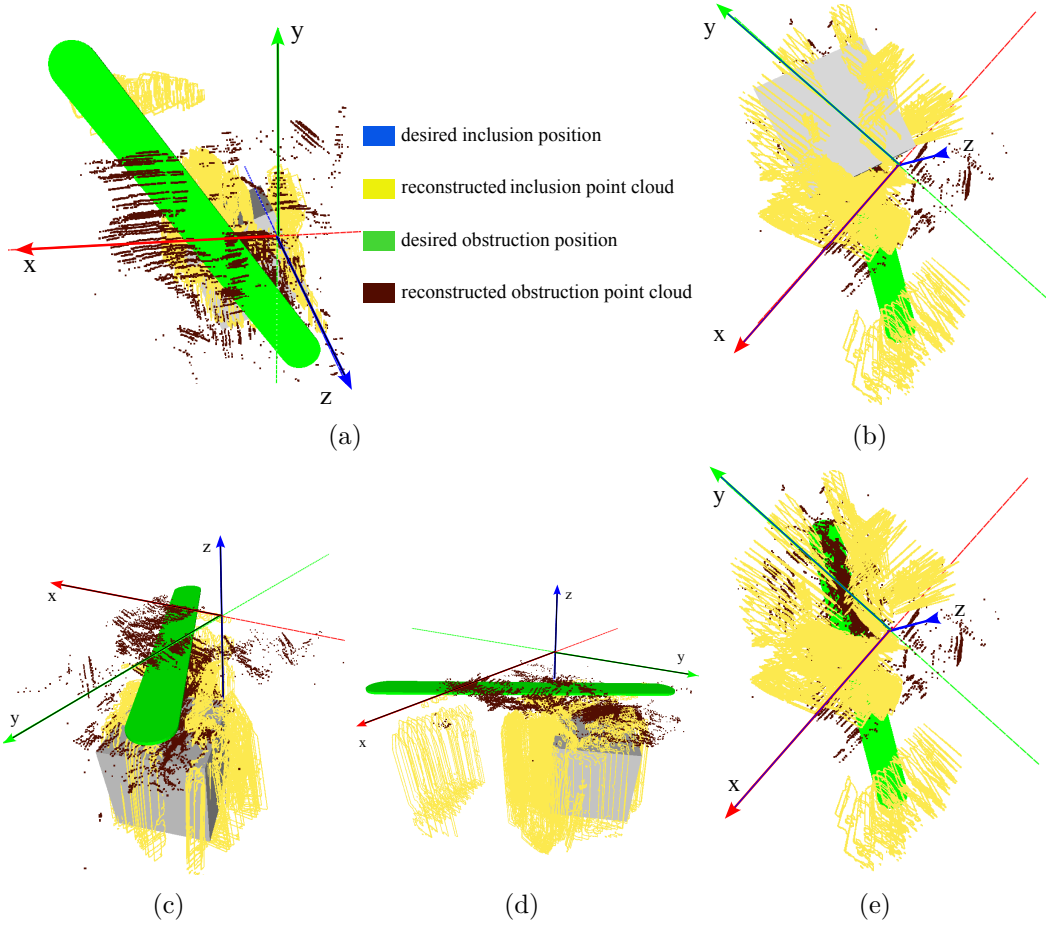


Figure 5-9: The resulting point cloud when voxel length was 5mm. Figure 5-9(a) and Figure 5-9(b) are the top view and bottom view, respectively, while Figure 5-9(c) and 5-9(d) are side and front view, respectively. Figure 5-9(e) is the same view with 5-9(b), but the true inclusion mesh was removed for better observation. Some outliers were removed and the true object meshes were enlarged for better observation.

acquisition was comsidered complete after the two linear scans, and Figure 5-8(d) shows the final state.

The resulting post-time reconstructed point cloud is depicted in Figure 5-9. Larger and more vacancies in the reconstructed cubic inclusion can be observed compared to previous trials.

5.3 Results Discussion

The voxel size of 1mm is the maximum achievable resolution of the visualization system at current setup. Both the path sequence and the reconstructed results confirmed the success of the real-time path planning and scanning guidance system. Despite the presence of errors and outliers in the reconstructed volumes, the cubic inclusion was still recognizable in all three experiments.

Based on the post-processed reconstructed meshes, the 1mm case worked the best, resulting in a dense and complete mesh, but a dense voxel grid also increased the computational time generating the next best path, which caused a time delay of approximately 2 seconds. However, since path recalculating is only required when the previous scanning path is finished, this time-consuming step does not occur frequently during the experiment.

In the 2mm case, the time delay caused by path recalculating was deciseconds, which is acceptable because the human operator also needs time to react and move to the next path. Fan scans were suggested several times in this case, yet the gap caused by the obstacle was not fully filled.

In the 5mm case, the time consumed for path planning could be entirely neglected, but a obvious gap occurs in the reconstructed mesh. This is predictable because the algorithm did not suggest fan scans during the experiment. Although some vacant

areas might be reached by some unconscious motions of the human operator, the large vacant in the final mesh indicates that this resolution is too sparse.

Therefore, we can conclude that the 5mm voxel length is too sparse for this scenario, while the 2mm voxel length is acceptable despite some small gaps, and the 1mm voxel length is ideal.

The time cost for the three experiments was initially overestimated because the probe needed to be be moved slowly between two different scanning paths, or the depth camera and the IMU module would lose tracking ability. Nevertheless, since the tests were performed by a same person familiar with the system, the time consumed in each trial can still imply that the time expended would increase as the increasing of resolution. Therefore, a optimal resolution can be selected between 1mm and 2mm for voxel length, which depends on the preference between the scanning speed and quality. A more intuitive trajectory guidance tool can also be developed for better real-time guidance.

While some outliers may result from incorrect segmentation, such as the mixture of obstacle and inclusion, or classifying the container boundaries or other defacts as the inclusion, there is another type of error originating from localization. Specifically, the expansion of the cube in one direction might arise from the accumulating error in the depth estimation of the camera and the integrating error of the IMU. The points outside the normal range were also introduced by localization failures. Therefore, the stability and consistency of the localization strategy should be improved for further studies.

Furthermore, due to the similarity of the shape of potato and the defects in the phantom, the potato contour was not identifiable in any of the experiments. Additionally, the phantom has a simple structure, and the unclear area caused by unsuitable scanning pose does not exist in this phantom, so the better pose suggestion

for unclear area has not been tested. Hence, a better segmentation method will be desired for more complex scenarios.

Overall, the validation system and the experiments demonstrate the potential of using online path planning with real-time guidance and 3D mesh reconstruction for freehand ultrasound. The system can guide the user to explore the information hidden under the high-attenuation areas, enabling users unfamiliar with ultrasound imaging to acquire 3D freehand ultrasound image accurately and efficiently.

Chapter 6

Conclusions and Future Work

In summary, this thesis presented a novel approach to improve the accuracy, completeness and efficiency of freehand ultrasound imaging through real-time path planning and guidance. The proposed system utilized a low-cost RGB-D camera and IMU module to estimate the spatial position and orientation of the ultrasound probe and provide real-time guidance for the user. The system was validated on an ultrasound phantom with high-attenuation obstacle on the top, proving the ability of the system in real-time path planning and scanning guidance, and bypassing the high-attenuation area to explore the hidden information.

This study demonstrated the possibility of online path planning with real-time guidance and high-attenuation area avoidance for freehand ultrasound scanning, even in scenarios where prior knowledge of the scanning area is not available. By avoiding high-attenuation areas and improving the view angle of the region of interest, the proposed path planning framework not only improved the accuracy and efficiency of ultrasound imaging in clinical settings but also enabled non-expert users to acquire high-quality 3D ultrasound images in a more convenient way, indicating the potential

of this system for long-term health monitoring.

Based on this initial demonstration of path planning system for freehand ultrasound trajectory guidance, more potential reasearches and applications can be explored in the future:

1. Integrating some model-based image segmentation methods such as deep learning, and prori knowledge of body surface data or other medical images for better reconstruction in some particular areas.

2. Utilizing visualization techniques such as Augmented Reality (AR) to provide more intuitive and comprehensible real-time guidance.

3. Analyzing the speed of sound together with the structural information for better acquisition of the region of interst.

4. Applying the improved system to human studies to realize path planning under complicated body structures.

5. Alternatively, intergrating with articulated robotic arm systems to enable the automatic ultrasound scanning in complex body area.

Bibliography

[1] Purang Abolmaesumi, Septimiu E Salcudean, Wen-Hong Zhu, Mohammad Reza Sirouspour, and Simon Peter DiMaio. Image-guided control of a robot for medical ultrasound. *IEEE transactions on robotics and automation*, 18(1):11–23, 2002.

[2] Zeynettin Akkus, Alfiia Galimzianova, Assaf Hoogi, Daniel L Rubin, and Bradley J Erickson. Deep learning for brain mri segmentation: state of the art and future directions. *Journal of digital imaging*, 30:449–459, 2017.

[3] Marija Đakulovic and Ivan Petrovic. Complete coverage path planning of mobile robots for humanitarian demining. *Industrial Robot: An International Journal*, 39(5):484–493, 2012.

[4] Maria Antico, Damjan Vukovic, Saskia M Camps, Fumio Sasazawa, Yu Takeda, Anh TH Le, Anjali T Jaiprakash, Jonathan Roberts, Ross Crawford, Davide Fontanarosa, et al. Deep learning for us image quality assessment based on femoral cartilage boundary detection in autonomous knee arthroscopy. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 67(12):2543–2552, 2020.

[5] Jeannette Bakker, Marco Olree, Robert Kaatee, Eduard E. de Lange, Karel G. M. Moons, Jaap J. Beutler, and Frederik J. A. Beek. Renal volume measurements: Accuracy and repeatability of us compared with that of mr imaging. *Radiology*, 211(3):623–628, 1999. PMID: 10352583.

[6] Alex Benjamin, Melinda Chen, Qian Li, Lei Chen, Yi Dong, Carolina A. Carrascal, Hua Xie, Anthony E. Samir, and Brian W. Anthony. Renal volume estimation using freehand ultrasound scans: An ex vivo demonstration. *Ultrasound in Medicine & Biology*, 46(7):1769–1782, 2020.

[7] S Michelle Bierig and Anne Jones. Accuracy and cost comparison of ultrasound versus alternative imaging modalities, including ct, mr, pet, and angiography. *Journal of Diagnostic Medical Sonography*, 25(3):138–144, 2009.

[8] AM Blamire. The technology of mrithe next 10 years? *The British journal of radiology*, 81(968):601–617, 2008.

[9] Jack E Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems journal*, 4(1):25–30, 1965.

[10] Zuo Llang Cao, Yuyu Huang, and Ernest L. Hall. Region filling operations with random obstacle avoidance for mobile robots. *Journal of Robotic Systems*, 5(2):87–102, 1988.

[11] Naga Chalasani, Zobair Younossi, Joel E Lavine, Anna Mae Diehl, Elizabeth M Brunt, Kenneth Cusi, Michael Charlton, and Arun J Sanyal. The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the american association for the study of liver diseases, american college of gastroenterology, and the american gastroenterological association. *Hepatology*, 55(6):2005–2023, 2012.

[12] Arlene B Chapman, James E Bost, Vicente E Torres, Lisa Guay-Woodford, Kyongtae Ty Bae, Douglas Landsittel, Jie Li, Bernard F King, Diego Martin, Louis H Wetzel, et al. Kidney volume and functional outcomes in autosomal dominant polycystic kidney disease. *Clinical Journal of the American Society of Nephrology*, 7(3):479–486, 2012.

[13] Pierre Chatelain, Alexandre Krupa, and Nassir Navab. Confidence-driven control of an ultrasound probe. *IEEE Transactions on Robotics*, 33(6):1410–1424, 2017.

[14] Howie Choset. Coverage for robotics–a survey of recent results. *Annals of mathematics and artificial intelligence*, 31(1):113–126, 2001.

[15] Kendell Cole. Sonography's expansion into space. *Journal of Diagnostic Medical Sonography*, 24(6):380–387, 2008.

[16] Pierrick Coupé, Pierre Hellier, Noura Azzabou, and Christian Barillot. 3d freehand ultrasound reconstruction based on probe trajectory. In *8th International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 3749, pages 597–604. Springer, 2005.

[17] R Neumann De Carvalho, HA Vidal, P Vieira, and MI Ribeiro. Complete coverage path planning and guidance for cleaning robots. In *ISIE'97 Proceeding of the IEEE International Symposium on Industrial Electronics*, volume 2, pages 677–682. IEEE, 1997.

[18] Carlo Di Mario, Stanley Madretsma, David Linker, Salem HK The, Nicolaas Bom, Patrick W Serruys, Elma J Gussenhoven, and Jos RTC Roelandt. The angle of incidence of the ultrasonic beam: a critical factor for the image quality in intravascular ultrasonography. *American Heart Journal*, 125(2):442–448, 1993.

[19] Richard Droste, Lior Drukker, Aris T Papageorghiou, and J Alison Noble. Automatic probe movement guidance for freehand obstetric ultrasound. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 583–592. Springer, 2020.

[20] Inge Edler and Carl Hellmuth Hertz. The early work on ultrasound in medicine at the university of lund. *Journal of Clinical Ultrasound*, 5(5):352–356, 1977.

[21] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical programming*, 1:127–136, 1971.

[22] Garabed Eknoyan, Norbert Lameire, K Eckardt, B Kasiske, D Wheeler, A Levin, PE Stevens, RW Bilous, EJ Lamb, JJKI Coresh, et al. Kdigo 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney int*, 3(1):5–14, 2013.

[23] Alberto Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal on Robotics and Automation*, 3(3):249–265, 1987.

[24] Marco Esposito, Christoph Hennersperger, Rüdiger Göbl, Laurent Demaret, Martin Storath, Nassir Navab, Maximilian Baust, and Andreas Weinmann. Total variation regularization of pose signals with an application to 3d freehand ultrasound. *IEEE transactions on medical imaging*, 38(10):2245–2258, 2019.

[25] Simon DS Fraser and Tom Blakeman. Chronic kidney disease: identification and management in primary care. *Pragmatic and Observational Research*, 7:21–32, 2016. PMID: 27822135.

[26] Mireen Friedrich-Rust, Guenther Schneider, Rainer M Bohle, Eva Herrmann, Christoph Sarrazin, Stefan Zeuzem, and Joerg Bojunga. Contrast-enhanced

sonography of adrenal masses: differentiation of adenomas and nonadenomatous lesions. *American Journal of Roentgenology*, 191(6):1852–1860, 2008.

[27] Patricia A. Gabow. Autosomal dominant polycystic kidney disease. *New England Journal of Medicine*, 329(5):332–342, 1993. PMID: 8321262.

[28] Enric Galceran and Marc Carreras. A survey on coverage path planning for robotics. *Robotics and Autonomous Systems*, 61(12):1258–1276, 2013.

[29] Tom JG Gevers and Joost PH Drenth. Diagnosis and management of polycystic liver disease. *Nature reviews Gastroenterology & hepatology*, 10(2):101–108, 2013.

[30] Goutam Ghoshal, Roberto J Lavarello, Jeremy P Kemmerer, Rita J Miller, and Michael L Oelze. Ex vivo study of quantitative ultrasound parameters in fatty rabbit livers. *Ultrasound in medicine & biology*, 38(12):2238–2248, 2012.

[31] Matthew W Gilbertson and Brian W Anthony. Force and position control system for freehand ultrasound. *IEEE Transactions on Robotics*, 31(4):835–849, 2015.

[32] David G Gobbi and Terry M Peters. Interactive intra-operative 3d ultrasound reconstruction and visualization. In *Medical Image Computing and Computer-Assisted InterventionMICCAI 2002: 5th International Conference Tokyo, Japan, September 25–28, 2002 Proceedings, Part II 5*, pages 156–163. Springer, 2002.

[33] Rüdiger Göbl, Salvatore Virga, Julia Rackerseder, Benjamin Frisch, Nassir Navab, and Christoph Hennersperger. Acoustic window planning for ultrasound acquisition. *International Journal of Computer Assisted Radiology and Surgery*, 12(6):993–1001, 2017.

[34] Jared J. Grantham, Vicente E. Torres, Arlene B. Chapman, Lisa M. Guay-Woodford, Kyongtae T. Bae, Bernard F. King, Louis H. Wetzel, Deborah A. Baumgarten, Phillip J. Kenney, Peter C. Harris, Saulo Klahr, William M. Bennett, Gladys N. Hirschman, Catherine M. Meyers, Xiaoling Zhang, Fang Zhu, and John P. Miller. Volume progression in polycystic kidney disease. *New England Journal of Medicine*, 354(20):2122–2130, 2006. PMID: 16707749.

[35] Christoph Graumann, Bernhard Fuerst, Christoph Hennersperger, Felix Bork, and Nassir Navab. Robotic ultrasound trajectory planning for volume of interest coverage. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 736–741, 2016.

[36] Ibrahim A Hameed. Coverage path planning software for autonomous robotic lawn mower using dubins' curve. In *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pages 517–522. IEEE, 2017.

[37] Peter C Harris and Vicente E Torres. Polycystic kidney disease, autosomal dominant. *GeneReviews*, Jan 2022.

[38] Christoph Hennersperger, Bernhard Fuerst, Salvatore Virga, Oliver Zettinig, Benjamin Frisch, Thomas Neff, and Nassir Navab. Towards mri-based autonomous robotic us acquisitions: a first feasibility study. *IEEE transactions on medical imaging*, 36(2):538–548, 2016.

[39] Karin Horsthuis, Shandra Bipat, Roelof J Bennink, and Jaap Stoker. Inflammatory bowel disease diagnosed with us, mr, scintigraphy, and ct: meta-analysis of prospective studies. *Radiology*, 247(1):64–79, 2008.

[40] Samantha Horvath, John Galeotti, Bo Wang, Matt Perich, Jihang Wang, Mel Siegel, Patrick Vescovi, and George Stetten. Towards an ultrasound probe with vision: structured light to determine surface orientation. In *Augmented Environments for Computer-Assisted Interventions: 6th International Workshop, AE-CAI 2011, Held in Conjunction with MICCAI 2011, Toronto, ON, Canada, September 22, 2011, Revised Selected Papers 6*, pages 58–64. Springer, 2012.

[41] F Hottier and A Collet Billon. 3d echography: status and perspective. In *3D Imaging in Medicine: Algorithms, Systems, Applications*, pages 21–41. Springer, 1990.

[42] Qinghua Huang, Bowen Wu, Jiulong Lan, and Xuelong Li. Fully automatic three-dimensional ultrasound imaging based on conventional b-scan. *IEEE transactions on biomedical circuits and systems*, 12(2):426–436, 2018.

[43] Qinghua Huang and Zhaozheng Zeng. A review on real-time 3d ultrasound imaging technology. *BioMed research international*, 2017, 2017.

[44] Shuya Ito, Koichi Ito, Takafumi Aoki, Jun Ohmiya, and Satoshi Kondo. Probe localization using structure from motion for 3d ultrasound image reconstruction. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 68–71. IEEE, 2017.

[45] Vesa M Järvinen, Markku M Kupari, Pauli E Hekali, and Veli-Pekka Poutanen. Right atrial mr imaging studies of cadaveric atrial casts and comparison with right and left atrial volumes and function in healthy subjects. *Radiology*, 191(1):137–142, 1994.

[46] Hyeong Ryeol Kam, Sung-Ho Lee, Taejung Park, and Chang-Hun Kim. Rviz: A toolkit for real domain data visualization. *Telecommun. Syst.*, 60(2):337345, oct 2015.

[47] Jakub T Kaminski, Khashayar Rafatzand, and Haichong K Zhang. Feasibility of robot-assisted ultrasound imaging with force feedback for assessment of thyroid diseases. In *Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 11315, pages 356–364. SPIE, 2020.

[48] Anna Kotronen and Hannele Yki-Jarvinen. Fatty liver: a novel component of the metabolic syndrome. *Arteriosclerosis, thrombosis, and vascular biology*, 28(1):27–38, 2008.

[49] Alexandre Krupa, Gabor Fichtinger, and Gregory D Hager. Real-time tissue tracking with b-mode ultrasound using speckle and visual servoing. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007: 10th International Conference, Brisbane, Australia, October 29-November 2, 2007, Proceedings, Part II 10*, pages 1–8. Springer, 2007.

[50] Oliver Kutter, Ramtin Shams, and Nassir Navab. Visualization and gpu-accelerated simulation of medical ultrasound from ct images. *Computer methods and programs in biomedicine*, 94(3):250–266, 2009.

[51] Gerhard Lakemeyer and Bernhard Nebel. *Exploring artificial intelligence in the new millennium*. Morgan Kaufmann, 2003.

[52] Jason A Levy and Richard G Bachur. Bedside ultrasound in the pediatric emergency department. *Current opinion in pediatrics*, 20(3):352–242, 2008.

[53] Keyu Li, Yangxin Xu, and Max Q-H Meng. An overview of systems and techniques for autonomous robotic ultrasound acquisitions. *IEEE Transactions on Medical Robotics and Bionics*, 3(2):510–524, 2021.

[54] Seong-Jae Lim, Yong-Yeon Jeong, and Yo-Sung Ho. Automatic liver segmentation for volume measurement in ct images. *Journal of Visual Communication and Image Representation*, 17(4):860–875, 2006.

[55] Xi-Zhang Lin, Yung-Nien Sun, Yuh Hwan Liu, Bor-Shyang Sheu, Bin Nan Cheng, Chiung-Yu Chen, Hong-Ming Tsai, and Ching Liang Shen. Liver volume in patients with or without chronic liver diseases. *Hepato-gastroenterology*, 45(22):1069–1074, 1998.

[56] Shengfeng Liu, Yi Wang, Xin Yang, Baiying Lei, Li Liu, Shawn Xiang Li, Dong Ni, and Tianfu Wang. Deep learning in medical ultrasound analysis: a review. *Engineering*, 5(2):261–275, 2019.

[57] O John Ma, Jeffrey G Norvell, and Srikala Subramanian. Ultrasound applications in mass casualties and extreme environments. *Critical care medicine*, 35(5):S275–S279, 2007.

[58] D Mahieu-Caputo, P Sonigo, M Dommergues, JC Fournet, JC Thalabard, C Abarca, A Benachi, F Brunelle, and Y Dumez. Fetal lung volume measurement by magnetic resonance imaging in congenital diaphragmatic hernia. *British Journal of Obstetrics and Gynaecology*, 108(8):863–868, 2001.

[59] Giulio Marchesini, Simona Moscatiello, Silvia Di Domizio, and Gabriele Forlani. Obesity-associated liver disease. *The Journal of Clinical Endocrinology & Metabolism*, 93(11_supplement_1):s74–s80, 2008.

[60] Nina A Mayr, Toshiaki Taoka, William TC Yuh, Leah M Denning, Weining K Zhen, Arnold C Paulino, Robert C Gaston, Joel I Sorosky, Sanford L Meeks, Joan L Walker, et al. Method and timing of tumor volume measurement for outcome prediction in cervical cancer using magnetic resonance imaging. *International Journal of Radiation Oncology* Biology* Physics*, 52(1):14–22, 2002.

[61] Rafik Mebarki, Alexandre Krupa, and François Chaumette. 2-d ultrasound probe complete guidance by visual servoing using image moments. *IEEE Transactions on Robotics*, 26(2):296–306, 2010.

[62] Kanta Miura, Koichi Ito, Takafumi Aoki, Jun Ohmiya, and Satoshi Kondo. Probe localization from ultrasound image sequences using deep learning for volume reconstruction. In *International Forum on Medical Imaging in Asia 2021*, volume 11792, pages 133–138. SPIE, 2021.

[63] Mehrdad Moghbel, Syamsiah Mashohor, Rozi Mahmud, and M Iqbal Bin Saripan. Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography. *Artificial Intelligence Review*, 50:497–537, 2018.

[64] Ammar Safwan Bin Mustafa, Takashi Ishii, Yoshiki Matsunaga, Ryu Nakadate, Hiroyuki Ishii, Kouji Ogawa, Akiko Saito, Motoaki Sugawara, Kiyomi Niki, and Atsuo Takanishi. Development of robotic system for autonomous liver screening using ultrasound scanning device. In *2013 IEEE international conference on robotics and biomimetics (ROBIO)*, pages 804–809. IEEE, 2013.

[65] Ryu Nakadate, Jorge Solis, Atsuo Takanishi, Eiichi Minagawa, Motoaki Sugawara, and Kiyomi Niki. Implementation of an automatic scanning and detection algorithm for the carotid artery by an assisted-robotic measurement system. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 313–318. IEEE, 2010.

[66] Myrte K Neijenhuis, Wietske Kievit, Stef MH Verheesen, Hedwig M DAgnolo, Tom JG Gevers, and Joost PH Drenth. Impact of liver volume on polycystic liver disease-related symptoms and quality of life. *United European Gastroenterology Journal*, 6(1):81–88, 2018.

[67] Thomas R Nelson and Dolores H Pretorius. Interactive acquisition, analysis, and visualization of sonographic volume data. *International Journal of Imaging Systems and Technology*, 8(1):26–37, 1997.

[68] Dick Oepkes, P Gareth Seaward, Frank PHA Vandenbussche, Rory Windrim, John Kingdom, Joseph Beyene, Humphrey HH Kanhai, Arne Ohlsson, and Greg Ryan. Doppler ultrasonography versus amniocentesis to predict fetal anemia. *New England Journal of Medicine*, 355(2):156–164, 2006.

[69] Ryutarou Ohbuchi, David Chen, and Henry Fuchs. Incremental volume reconstruction and rendering for 3-d ultrasound imaging. In *Visualization in Biomedical Computing'92*, volume 1808, pages 312–323. SPIE, 1992.

[70] Jerry L Old, Reginald W Dusing, Wendell Yap, and Jared Dirks. Imaging for suspected appendicitis. *American family physician*, 71(1):71–78, 2005.

[71] Arinc Ozturk, Joseph R Grajo, Michael S Gee, Alex Benjamin, Rebecca E Zubajlo, Kai E Thomenius, Brian W Anthony, Anthony E Samir, and Manish Dhyani. Quantitative hepatic fat quantification in non-alcoholic fatty liver

disease using ultrasound-based techniques: a review of literature and their diagnostic performance. *Ultrasound in medicine & biology*, 44(12):2461–2475, 2018.

[72] William D OBrien Jr, John W Erdman Jr, and Tammy B Hebner. Ultrasonic propagation properties (@ 100 mhz) in excessively fatty rat liver. *The Journal of the Acoustical Society of America*, 83(3):1159–1166, 1988.

[73] Mi Suk Park, Jeong Sik Yu, Young Hwan Kim, Myeong Jin Kim, Ji Hyung Kim, Sunah Lee, Nariya Cho, Dong Guk Kim, and Ki Whang Kim. Acute cholecystitis: comparison of mr cholangiography and us. *Radiology*, 209(3):781–785, 1998.

[74] York Pei and Terry Watnick. Diagnosis and screening of autosomal dominant polycystic kidney disease. *Advances in chronic kidney disease*, 17(2):140–52, 2010.

[75] Norbert J Pelc. Recent and future directions in ct imaging. *Annals of biomedical engineering*, 42:260–268, 2014.

[76] Richard W Prager, Umer Z Ijaz, AH Gee, and Graham M Treece. Three-dimensional ultrasound imaging. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(2):193–223, 2010.

[77] Raphael Prevost, Mehrdad Salehi, Simon Jagoda, Navneet Kumar, Julian Sprung, Alexander Ladikos, Robert Bauer, Oliver Zettinig, and Wolfgang Wein. 3d freehand ultrasound without external tracking using deep learning. *Medical image analysis*, 48:187–202, 2018.

[78] Hedyeh Rafii-Tari, Purang Abolmaesumi, and Robert Rohling. Panorama ultrasound for guiding epidural anesthesia: A feasibility study. In *Information Processing in Computer-Assisted Interventions: Second International Conference, IPCAI 2011, Berlin, Germany, June 22, 2011. Proceedings 2*, pages 179–189. Springer, 2011.

[79] Uma M Reddy, Roy A Filly, and Joshua A Copel. Prenatal imaging: ultrasonography and magnetic resonance imaging. *Obstetrics and gynecology*, 112(1):145, 2008.

[80] Hassan Rivaz, Richard Zellars, Gregory D. Hager, Gabor Fichtinger, and Emad M. Boctor. 9c-1 beam steering approach for speckle characterization and

out-of-plane motion estimation in real tissue. *2007 IEEE Ultrasonics Symposium Proceedings*, pages 781–784, 2007.

[81] Steven Rogers, Joao Carreira, Ruth Thompson, Ana Morais, Christopher Miller, Wolfgang Wein, Jonathan Ghosh, and Charles McCollum. An ex vivo evaluation of tomographic 3-d ultrasound, b-mode ultrasound, ct and mr imaging to measure artery diameter, length and wall volume. *Ultrasound in Medicine & Biology*, 45(10):2819–2829, 2019.

[82] Robert Rohling, Andrew Gee, and Laurence Berman. A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical image analysis*, 3(4):339–359, 1999.

[83] Radu Bogdan Rusu. *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. PhD thesis, Computer Science department, Technische Universitaet Muenchen, Germany, October 2009.

[84] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In *2011 IEEE International Conference on Robotics and Automation*, pages 1–4, 2011.

[85] Naveed Sattar, Ewan Forrest, and David Preiss. Non-alcoholic fatty liver disease. *Bmj*, 349, 2014.

[86] Julia Schwaab, Yago Diez, Arnau Oliver, Robert Martí, Jan van Zelst, Albert Gubern-Mérida, Ahmed Bensouda Mourri, Johannes Gregori, and Matthias Günther. Automated quality assessment in three-dimensional breast ultrasound images. *Journal of Medical Imaging*, 3(2):027002–027002, 2016.

[87] William R Scott, Gerhard Roth, and Jean-François Rivest. View planning for automated three-dimensional object reconstruction and inspection. *ACM Computing Surveys (CSUR)*, 35(1):64–96, 2003.

[88] H Tutkun Sen, Muyinatu A Lediju Bell, Iulian Iordachita, John Wong, and Peter Kazanzides. A cooperatively controlled robot for ultrasound monitoring of radiation therapy. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3071–3076. IEEE, 2013.

[89] Hasan Tutkun Şen, Alexis Cheng, Kai Ding, Emad Boctor, John Wong, Iulian Iordachita, and Peter Kazanzides. Cooperative control with ultrasound guidance for radiation therapy. *Frontiers in Robotics and AI*, 3:49, 2016.

[90] Shi Sherebrin, Aaron Fenster, Richard N Rankin, and David Spence. Freehand three-dimensional ultrasound: implementation and applications. In *Medical Imaging 1996: Physics of Medical Imaging*, volume 2708, pages 296–303. SPIE, 1996.

[91] Stanford Artificial Intelligence Laboratory et al. Robotic operating system.

[92] Shih-Yu Sun, Matthew Gilbertson, and Brian W Anthony. Probe localization for freehand 3d ultrasound by tracking skin features. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part II 17*, pages 365–372. Springer, 2014.

[93] Daniel R Swerdlow, Kevin Cleary, Emmanuel Wilson, Bamshad Azizi-Koutenaei, and Reza Monfaredi. Robotic arm–assisted sonography: Review of technical developments and potential clinical applications. *American Journal of Roentgenology*, 208(4):733–738, 2017.

[94] Konstantinos A Tarabanis, Peter K Allen, and Roger Y Tsai. A survey of sensor planning in computer vision. *IEEE transactions on Robotics and Automation*, 11(1):86–104, 1995.

[95] Sebastian Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71, 1998.

[96] Shidong Tong, H Neale Cardinal, Raymond F McLoughlin, Dónal B Downey, and Aaron Fenster. Intra-and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging. *Ultrasound in medicine & biology*, 24(5):673–681, 1998.

[97] Jason W Trobaugh, Darin J Trobaugh, and William D Richard. Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323, 1994.

[98] J. Irving Vasquez-Gomez, L. Enrique Sucar, Rafael Murrieta-Cid, and Efrain Lopez-Damian. Volumetric next-best-view planning for 3d object reconstruction with positioning error. *International Journal of Advanced Robotic Systems*, 11(10):159, 2014.

[99] G Vernon, A Baranova, and ZM Younossi. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic

steatohepatitis in adults. *Alimentary pharmacology & therapeutics*, 34(3):274–285, 2011.

[100] Salvatore Virga, Oliver Zettinig, Marco Esposito, Karin Pfister, Benjamin Frisch, Thomas Neff, Nassir Navab, and Christoph Hennersperger. Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms. In *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 508–513. IEEE, 2016.

[101] Shuangyi Wang, Davinder Singh, Devapriyan Johnson, Kaspar Althoefer, Kawal Rhode, and Richard James Housden. Robotic ultrasound: View planning, tracking, and automatic acquisition of transesophageal echocardiography. *IEEE Robotics & Automation Magazine*, 23(4):118–127, 2016.

[102] Wolfgang Wein, Shelby Brunke, Ali Khamene, Matthew R Callstrom, and Nassir Navab. Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical image analysis*, 12(5):577–585, 2008.

[103] JJ Wild and JM Reid. Diagnostic use of ultrasound. *The British journal of physical medicine: including its application to industry*, 19(11):248, 1956.

[104] Sylvia Wong. *Qualitative topological coverage of unknown environments by mobile robots*. PhD thesis, ResearchSpace@ Auckland, 2006.

[105] Sylvia C Wong and Bruce A MacDonald. Complete coverage by mobile robots using slice decomposition based on natural landmarks. In *PRICAI 2004: Trends in Artificial Intelligence: 8th Pacific Rim International Conference on Artificial Intelligence, Auckland, New Zealand, August 9-13, 2004. Proceedings 8*, pages 683–692. Springer, 2004.

[106] Lingyun Wu, Jie-Zhi Cheng, Shengli Li, Baiying Lei, Tianfu Wang, and Dong Ni. Fuiqa: fetal ultrasound image quality assessment with deep convolutional networks. *IEEE transactions on cybernetics*, 47(5):1336–1349, 2017.

[107] Guang-Zhong Yang, James Cambias, Kevin Cleary, Eric Daimler, James Drake, Pierre E Dupont, Nobuhiko Hata, Peter Kazanzides, Sylvain Martel, Rajni V Patel, et al. Medical roboticsregulatory, ethical, and legal considerations for increasing levels of autonomy, 2017.

[108] Jesse T Yen, Jordan P Steinberg, and Stephen W Smith. Sparse 2-d array design for real time rectilinear volumetric imaging. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 47(1):93–110, 2000.

[109] Alan S.L. Yu, Chengli Shen, Douglas P. Landsittel, Peter C. Harris, Vicente E. Torres, Michal Mrug, Kyongtae T. Bae, Jared J. Grantham, Frederic F. Rahbari-Oskoui, Michael F. Flessner, William M. Bennett, and Arlene B. Chapman. Baseline total kidney volume and the rate of kidney growth are associated with chronic kidney disease progression in autosomal dominant polycystic kidney disease. *Kidney International*, 93(3):691–699, 2018.

[110] Cheng Zhao, Richard Droste, Lior Drukker, Aris T Papageorghiou, and J Alison Noble. Visual-assisted probe movement guidance for obstetric ultrasound scanning using landmark retrieval. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 670–679. Springer, 2021.

[111] Yanong Zhu, D Magee, R Ratnalingam, and D Kessel. A virtual ultrasound imaging system for the simulation of ultrasound-guided needle insertion procedures. In *Proceedings of Medical Image Understanding and Analysis*, pages 61–65, 2006.