Optimisation based trajectory planning using robot perception

Undergraduate Thesis

Submitted in partial fulfillment of the requirements of
BITS F421T Thesis

By

Ashutosh Gupta
ID No. 2019AATS0223G

Under the supervision of:

Prof. Howie Choset
&
Dr. Lu Li

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI, GOA CAMPUS
September 2025

Declaration of Authorship

I, Ashutosh Gupta, declare that this Undergraduate Thesis titled, 'Optimisation based trajectory planning using robot perception' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

Certificate

This is to certify that the thesis entitled, "*Optimisation based trajectory planning using robot perception*" and submitted by Ashutosh Gupta ID No. 2019AATS0223G in partial fulfillment of the requirements of BITS F421T Thesis embodies the work done by him under my supervision.

Supervisor	*Co-Supervisor*
Prof. Howie Choset	Dr. Lu Li
Professor,	Project Scientist,
Carnegie Mellon University	Carnegie Mellon University
Date:	Date:

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI, GOA CAMPUS

Abstract

Bachelor of Engineering (Hons.)

Optimisation based trajectory planning using robot perception

by Ashutosh Gupta

Ultrasound scanning is an efficient imaging modality preferred for quick medical procedures. However, due to the lack of skilled sonographers, researchers have developed many Robotic Ultrasound System (RUS) prototypes for various procedures. Most of these systems have a human-in-the-loop and require an expert to point the robot to the region of the subject to be scanned. Only a few systems try to incorporate some knowledge from the exterior shape of the subject for ultrasound scanning. Accurate 3D surface reconstruction of a patient's exterior can enable an RUS to perceive subjects more like a clinician would. It can help localize the subject for the robot while eliminating input from an expert. Ultrasound scanning trajectories can be better planned if the RUS first detects critical regions on the surface of the subject and corresponding curvatures. In this paper, we use an RGB-D sensor to acquire point clouds representing the 3D surface of the subject, which in the present work is for a lower-torso leg phantom. We present a consolidated pipeline for creating an optimized 3D surface reconstruction of a subject and use it to autonomously identify a region of interest for scanning femoral vessels with an ultrasound probe. To make our system more robust to inter-subject variations in shape and size, we incorporate a trajectory optimization module of the RUS-mounted RGB-D sensor. To this end, we introduce a comprehensive evaluation score to quantify the quality of point cloud reconstructions. The resulting improvements in 3D surface scanning and reconstruction enable near-automation in generating ultrasound scanning trajectories for femoral vessels. Our pipeline produces ultrasound images with an average ZNCC score of 0.86 and our 3D point cloud reconstructions are accurate up to 1e-5 m based on the Chamfer distance from a ground-truth high-resolution CT scan.

Acknowledgements

I would like to acknowledge the continuous support of my parents throughout my undergraduate thesis and in fact, my entire academic life. Thanks to Professor Howie Choset for being such a helpful advisor and giving constructive feedback all along. I would also like to mention Andrew Orekhov, Lu Li, Anamya Bal and Abhimanyu whose guidance and inspiration all throughout the 6 months at The Robotics Institute of Carnegie Mellon University made it possible for me to complete this thesis. I am also quite grateful to my on-campus advisor Prof. Abhijit Pethe for all the help. I also place on record, my sense of gratitude to one and all, who directly or indirectly, have lent their hand in this venture.

Contents

Declaration of Authorship i

Certificate ii

Abstract iii

Acknowledgements iv

Contents v

List of Figures vii

List of Tables viii

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Our Work . . . 2

2 Related Work 4

3 Methods 6
- 3.1 Robot Setup . . . 6
- 3.2 Architecture Overview . . . 6
- 3.3 Point Clustering for Leg Localization . . . 7
- 3.4 Start and End Point Generation for RGB-D Scanning . . . 8
- 3.5 Composite Surface Reconstruction . . . 9
- 3.6 Evaluation Score Computation . . . 9
  - 3.6.1 Ratio of Inliers . . . 10
  - 3.6.2 Ratio of Continuous Area to Total Area . . . 10
  - 3.6.3 Normalized Standard Deviation of Neighborhood Normals . . . 11
- 3.7 RGB-D Scanning Trajectory Optimization . . . 11
- 3.8 RoI Detection and US Trajectory Generation . . . 13

4 Experiments and Results 14
- 4.1 Clustering . . . 14

4.2 RGB-D Trajectory Optimization 14
4.3 Ultrasound Scanning 18

5 Conclusion **19**

Bibliography **20**

List of Figures

1.1 Our RUS setup, similar to [6], with novelties in evaluation score and optimization for automated US trajectory generation. 2

3.1 Surface Reconstruction Pipeline with novel contributions highlighted 7
3.2 Results from color-based region growing clustering - a) Single instance point cloud; b) Identified clusters; c) Selected cluster with selection criteria (Eqn. 1) corresponds to the leg 8
3.3 Steps in start and end point detection - a) Selected cluster point cloud along with the centroid, principal axis (red) and 2D convex hull (green); b) Detected start and end points for the RGB-D scanning trajectory along the principal axis. . . . 9
3.4 (Left) Top view of the cluster point cloud, (Right) Computed inliers highlighted in red. 10
3.5 Top and side view of a cluster point cloud with a concavity. 11
3.6 Leg cluster with detected RoI in purple and US trajectory in green. 13

4.1 Multiple scenarios with occlusions on the leg phantom and presented reconstructions with selected leg clusters. 15
4.2 Plot of the evaluation score vs epochs for the optimization Llop a) with all three metrics; b) without $Metric_1$; c) without $Metric_2$; d) without $Metric_3$; in the evaluation score. The scores in b, c, d are scaled between 0 and 1 for equal comparison. 16
4.3 Improving reconstructed point clouds with increase in epochs in the optimization loop. 17
4.4 a) RoI based US trajectory overlaid on the 3D reconstructed surface; b) and c) Femoral vessels captured in our generated US trajectory (blue) 17

List of Tables

3.1 Cylinder Radii, Optimized Height and Evaluation Scores 12

4.1 ZNCC score comparison for various trajectories 18

Chapter 1

Introduction

1.1 Background

Ultrasound (US) examination is an indispensable tool in quick diagnostic and medical intervention procedures [6]. It has several advantages such as high portability, zero ionizing radiation, and low costs which make it suitable for emergency interventions such as Extracorporeal Membrane Oxygenation (ECMO), and Resuscitative Endovascular Balloon Occlusion (REBOA) [18] [10]. Despite its strengths, ultrasound scanning has a strong dependence on the skills of trained human professionals (Sonographers) [3]. Well-trained sonographers are not present everywhere, and so a Robotic Ultrasound System (RUS) could provide medical assistance when professionals are absent [3].

To get a robot to perform ultrasound scanning, we need to replicate a wide variety of steps typically performed by a sonographer like - (1) finding an appropriate area on the patient to scan, (2) moving the ultrasound probe within the region of interest while making subtle corrections to the probe's pose, especially over regions of high curvature, and (3) providing safe, significant and accurate forces through the probe to maintain diagnosable image quality while preventing patient injury. In a typical RUS, the Region of Interest (RoI) to scan is defined by human input from a skilled operator, e.g. a sonographer or doctor, based on internal as well as superficial anatomical landmarks on the skin [2] [19]. Identifying these landmarks requires a 3D understanding of the skin of the subject. An accurate 3D reconstruction can also help provide normals for controlling the ultrasound probe.

Reconstructions in the form of 3D point clouds provide multiple features such as surface normals, curvature etc, that can be analyzed for the above-mentioned purposes. RGB-D sensing provides an accessible modality for capturing 3D point cloud data and is increasingly being used in RUSs. Having an accurate point cloud representation of the objects present in the field of view of the

sensor, can help us localize the subject and identify based on superficial landmarks, where to position the US probe. However, surface reconstruction pipelines usually require human input to infer RGB-D sensor trajectories. In the case of medical applications, this input is required from experts. Additionally, surface reconstruction is prone to noise stemming from incorrect sensor placement and changes in environmental factors like lighting, sterile casings etc. Hence, elimination of human input for automation and improvement of reconstruction quality are both open areas in research.

1.2 Our Work

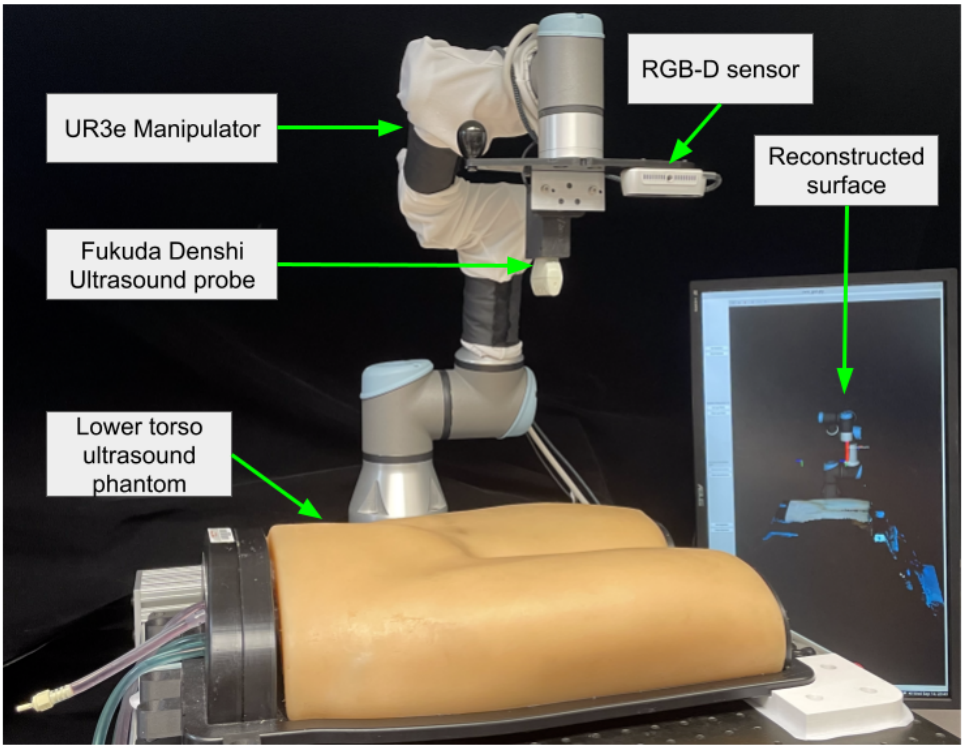

FIGURE 1.1: Our RUS setup, similar to [6], with novelties in evaluation score and optimization for automated US trajectory generation.

In this thesis, we present a novel pipeline for 3D Surface Reconstruction with an RGB-D sensor trajectory optimization module for automatic generation of ultrasound scanning trajectories. We eliminate human inputs by automatically finding the start and end scanning points for our RGB-D sensor. Then, we optimize the height of the sensor to produce the highest fidelity point cloud reconstruction of the subject. Lastly, we use the identified RoI and surface normals of the best 3D point cloud reconstruction from our optimization to generate an ultrasound scanning trajectory. This work is done in the context of scanning femoral vessels and we have designed the pipeline focusing on scanning the leg region. We have tested our pipeline on a medical phantom and have achieved near automation in optimized RGB-D and ultrasound scanning.

Main contributions are:

- A novel 3D surface reconstruction pipeline which feeds into automated ultrasound scanning trajectory generation.
- A comprehensive evaluation score for quantifying the quality of reconstructed point clouds.
- An evolutionary algorithm-based real-time optimization module that determines the best trajectory for the RGB-D sensor for surface reconstruction.

The rest of the thesis is divided as follows - Chapter 2 discusses related work in the areas of RUS, RGB-D based 3D reconstruction, sensor position optimization, and RoI for ultrasound scanning. Chapter 3 discusses our methods, followed by our results in Chapter 4. Finally, we present our conclusions and future work in Chapter 5.

Chapter 2

Related Work

In, [11], an RUS is presented with an automatic vessel tracking strategy. The pipeline was applied to provide 3D internal volume results of the lower limb arteries. The system can calculate the distance between the center of the vessel and the center element of the probe. However, the system performance heavily depends upon the detection and tracking of the vessels. An RUS developed by [5] autonomously generates trajectories based on the points selected by the physician marked in an MRI or CT scan. This system enables autonomy but is dependent on MRI/CT scans which are expensive and not always available. We work on improving both these aspects.

A Kinect RGB-D sensor was used in [6] to obtain 3D contours of a lumbar phantom. They are able to obtain a scanning region but this is heavily reliant on heuristic rules and user experience. They also mention parameters which can be varied by the user for various cases. Our work differs from this by using a minimally heuristic approach. Another RGB-D-based RUS is presented by [8] to detect sarcopenia in legs. They perform a piece-wise curve fitting to understand the leg surface from 8 point clouds. Their system has only 4 DoF and takes 12 minutes to perform RGB-D scanning and curve fitting for US probe positioning. We do not use computationally expensive math operations such as curve-fitting.

For estimating the optimal trajectory of the RGB-D sensor, there exist methods like [17], which uses filtering of multiple RGB frames to optimize the trajectory. This method only shown to work in simulated conditions. Similarly, [7] also presents results only in simulation. These methods therefore can't be used for real-world scenarios especially surgical settings where there are multiple occlusions present.

Most RUS systems depend on human operators to define the path that the ultrasound probe should trace. [4] uses an operator defined volume of interest. [20] presents an RUS system to enable needle insertion. This system needs user defined points on the surface to define the

needed trajectory for maximum vessel coverage. Some works like [3] automate RoI detection with only ultrasound images and do not take into consideration the exterior shape of the subject. Very few systems like [6] automate RoI detection for ultrasound scanning using external surface information.

Chapter 3

Methods

3.1 Robot Setup

Our system uses 2 sensors: an Intel Realsense D435i RGB-D sensor and a Fukuda Denshi portable point-of-care ultrasound scanner. Both these sensors are mounted on a Universal Robot UR3e manipulator. The experiments were conducted on a lower torso ultrasound training model BPF1500-HP. See Fig. 1.1 for our RUS Setup. In this work, we assume that the subject is in a supine position on a flat surface and is in the field-of-view of the RGB-D sensor. Our pipeline has been implemented in C++ and Python, with functions from Point Cloud Library [15] and Robot Operating System (ROS) [13] was used to combine all of the components.

3.2 Architecture Overview

Our pipeline, seen in Fig. 3.1, has the following modules

1. Clustering for leg localization.
2. Generation of start and end points for RGB-D scanning trajectory.
3. Composite surface reconstruction as we read point cloud data from the RGB-D sensor.
4. Computation of an evaluation score based on the quality of reconstruction.
5. RGB-D sensor trajectory optimization.
6. RoI detection and ultrasound trajectory generation.

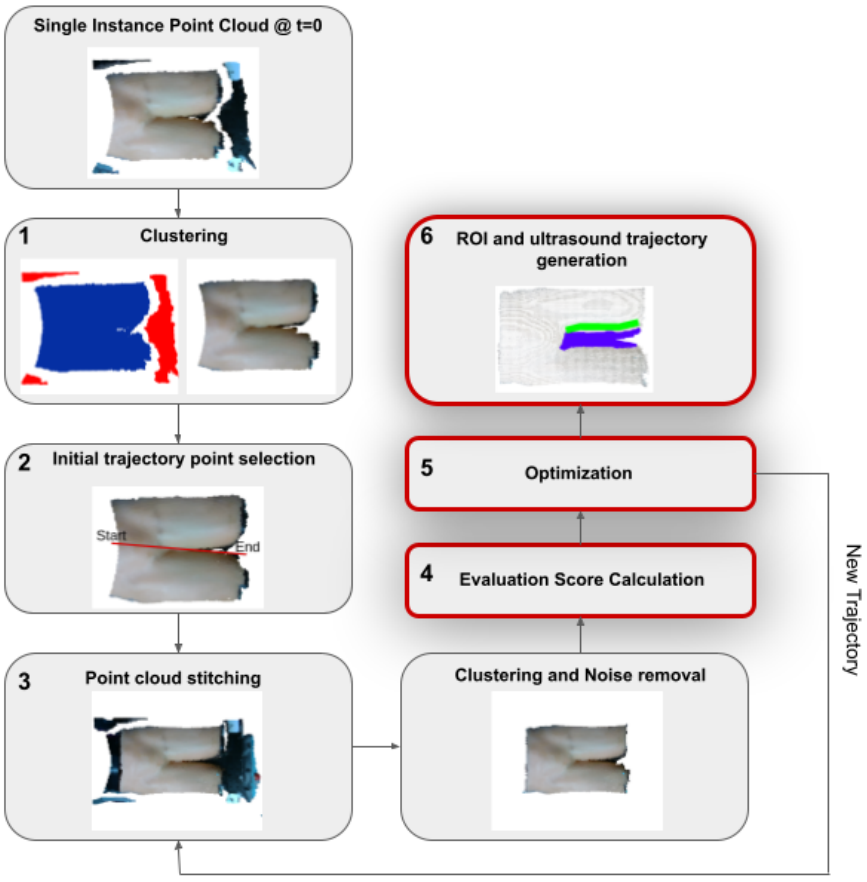

FIGURE 3.1: Surface Reconstruction Pipeline with novel contributions highlighted

We first take a single sensor reading and obtain the point cloud of the subject, along with the base and other objects, if any. Our clustering module finds the points corresponding to the skin of the subject and finds an axis to perform the RGB-D scan along. We generate a trajectory along this axis, scan along this trajectory and compute an evaluation score. This score is utilized by an optimizer to find the optimized trajectory producing the highest evaluation score. Once we find the optimized trajectory, we perform one last RGB-D scan along this trajectory and use the reconstruction from this scan to find the RoI. The RoI and its corresponding surface normals are then used to generate an ultrasound scanning trajectory.

3.3 Point Clustering for Leg Localization

Point clustering is required to identify all points in the point cloud corresponding to the skin of the subject being scanned. Once the subject is positioned in the sensor field-of-view, we position the RGB-D sensor at the highest possible height. We then capture a single instance point cloud from this height representing the subject on a relatively flat base. This point cloud is downsampled with a 5 mm voxel size. The color-based region growing clustering algorithm from [21] is used for generating clusters from this downsampled point cloud. We use a distance

threshold value of 5 and a point color threshold value of 10 for the region growing. The region color threshold is 10 for the merging process. We use a high point-color threshold to compensate for variations in skin color for a single subject. This method is also agnostic to skin tone as we do not look for a specific skin color. We then localize the leg by finding the cluster which corresponds to the skin of the subject. This is done by applying a selection criteria to the clusters. Given that the leg has higher curvature than the base and covers the majority of field-of-view, the cluster corresponding to the leg will have both the highest average curvature and no. of points.

$$C = argmax_i(\alpha * (\frac{\sum_i N}{\sum_n N})) + (1 - \alpha) * (\frac{\sum_i \kappa}{\sum_n \kappa})) \tag{1}$$

i is the index of the selected cluster where $\sum_i N$ is the total number of points in a cluster C_i and $\sum_i \kappa$ is the sum of curvature of the cluster i. There are n such clusters to choose from. Also, α and $(1 - \alpha)$ are the weights given to the respective criteria. Based on experimental observations, we set $\alpha = 0.5$. The single instance point cloud, identified clusters, and the selected cluster are visualized in Fig. 3.2.

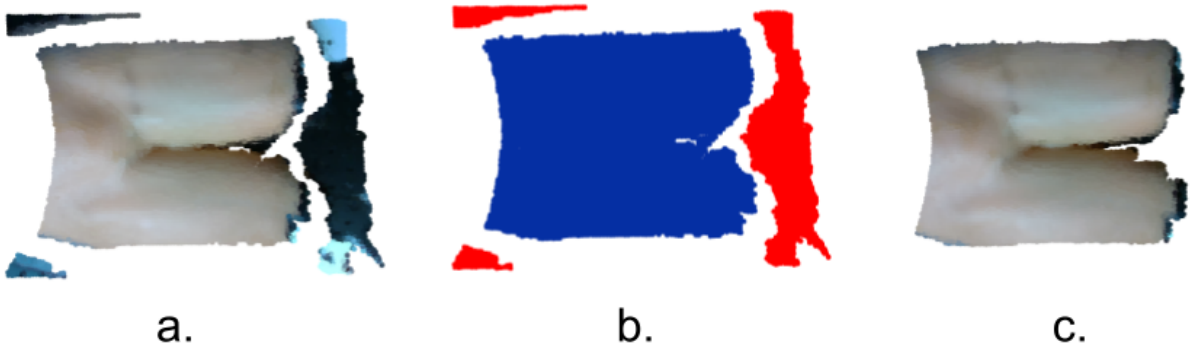

FIGURE 3.2: Results from color-based region growing clustering - a) Single instance point cloud; b) Identified clusters; c) Selected cluster with selection criteria (Eqn. 1) corresponds to the leg

3.4 Start and End Point Generation for RGB-D Scanning

Once the leg is localized, we generate a trajectory that the RGB-D sensor will follow for a detailed, closer scan. To obtain the start and end points of this trajectory, we use the principal axis of the selected cluster. For this, we flatten the selected cluster point cloud from the previous module along the z axis and find the 2D convex hull for the flattened point cloud. A covariance matrix is computed for the flattened point cloud and the eigen vectors for this matrix are computed. The largest eigen vector corresponds to the direction of the principal axis. We extend the principal axis vector in both directions from the centroid of the flattened point cloud. The intersection points of the principal axis with the edges of the 2D convex hull are found. The

(x, y) coordinates of the start and end points of the RGB-D scanning trajectory are the (x, y) coordinates of these intersection points. The points are visualized in Fig. 3.3. The z coordinates (height) of the start and end points are obtained from the optimization module.

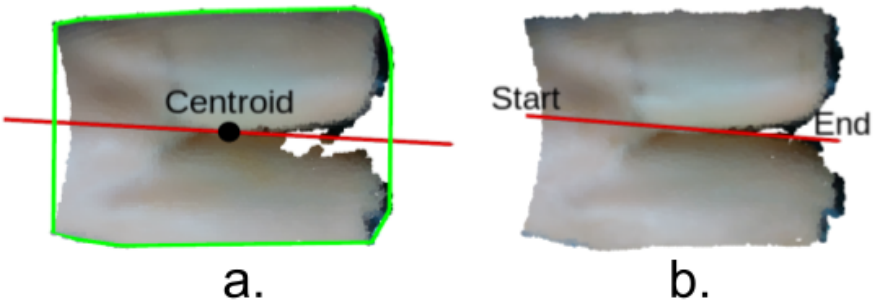

FIGURE 3.3: Steps in start and end point detection - a) Selected cluster point cloud along with the centroid, principal axis (red) and 2D convex hull (green); b) Detected start and end points for the RGB-D scanning trajectory along the principal axis.

3.5 Composite Surface Reconstruction

This module discusses the steps for processing frame-wise point clouds and generating a composite surface reconstruction for every RGB-D scanning trajectory. The RGB-D sensor was calibrated by obtaining the accurate transform from the sensor's optical frame to the robot base frame using an extensive CAD model of the robot and the camera mount. Frame-wise point clouds from the RGB-D sensor are read and transformed from the sensor frame to the robot base frame using parameters from the sensor calibration. Once in the base frame, they are stitched together as the sensor moves and newer points are read. The stitched point cloud is downsampled with voxel size 5 mm. This stitched point cloud is the composite point cloud from a given scan trajectory and is clustered again to obtain the cluster corresponding to the leg of the subject. This cluster point cloud, and the larger composite point cloud are used for the remaining pipeline.

3.6 Evaluation Score Computation

We have identified certain metrics to quantify the quality of the surface point cloud. In this module, we compute a score based on these metrics for evaluating the quality of the surface reconstruction and use this score to optimize the RGB-D scanning trajectory. The metrics are:

3.6.1 Ratio of Inliers

The quality of point clouds is dependent on the presence of outliers. Outliers signify regions that were not reconstructed properly, whereas the remaining points (inliers) signify well reconstructed regions. For quantifying the quality of our point cloud, we use $Metric_1$ where

$$Metric_1 = \frac{\text{No. of inlier points}}{\text{Total No. of points}} \tag{2}$$

in the evaluation score. Outliers are detected by the application of the statistical outlier detection algorithm [16] to the selected cluster. We consider 50 nearest neighbours and a standard deviation of 0.75 for our implementation. This gives us distinct outliers on the cluster point cloud. The inliers after the removal of detected outliers are visualized in Fig. 3.4. Additionally, we consider the points above and below the surface which are removed during clustering to be outliers as well while determining $Metric_1$.

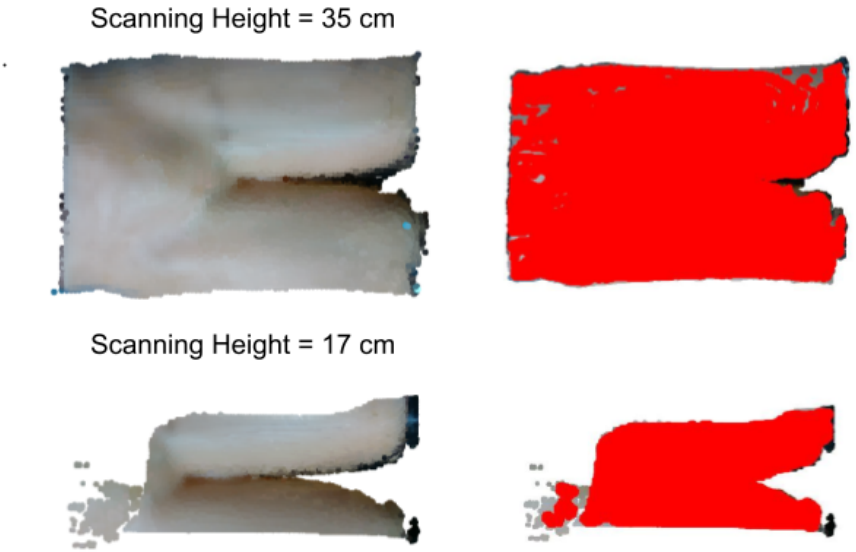

FIGURE 3.4: (Left) Top view of the cluster point cloud, (Right) Computed inliers highlighted in red.

3.6.2 Ratio of Continuous Area to Total Area

Different regions on the scanning subject could be reconstructed with different point densities. While most regions have uniform point density and are continuous, some regions lack points altogether and form hole-like regions on the surface resulting in concavities, as seen in Fig. 3.5.

To quantify our surface reconstruction as a function of continuous regions, we use a ratio of continuous area to the total area of the reconstruction. We obtain the continuous 2D projected area of the cluster point cloud by discounting discontinuous cavity regions and divide this by

FIGURE 3.5: Top and side view of a cluster point cloud with a concavity.

the total 2D projected area of the cluster point cloud (Eqn. 3). We obtain these quantities by computing the area of the concave and convex hulls after 2D projection respectively.

$$Metric_2 = \frac{\text{Area of 2D Concave Hull}}{\text{Area of 2D Convex Hull}} \tag{3}$$

3.6.3 Normalized Standard Deviation of Neighborhood Normals

Distortions present in reconstructions, especially when the sensor is too far from the subject, produce inconsistent surface normals which signify poor reconstructions. This characteristic needs to be penalized in the trajectory optimization. We quantify this by calculating the normalized standard deviation in surface normals for sample points in the cluster point cloud. t points are selected on the cluster using the farthest point sampling [12]. For each of these sampled points, we select k-neighbouring points and calculate variance (Var) of the surface normals in terms of the angle between them for each of these clusters. The final metric, which is the normalized standard deviation of all the chosen normals is given by

$$Metric_3 = \sqrt{\frac{\sum_t Var(cos^{-1}(n_i.n_j)_t)}{\sum_t Mean(cos^{-1}(n_i.n_j)_t)}} \tag{4}$$

where n_i and n_j are the two neighbouring normals on the surface at t^{th} point selected using farthest point sampling. For our experiments $t = 50$ and $k = 20$. Finally -

$$EvaluationScore = Metric_1 + Metric_2 - Metric_3 \tag{5}$$

3.7 RGB-D Scanning Trajectory Optimization

We hypothesized that inter-subject variations in shape and size require changes in RGB-D sensor positioning for good 3D reconstructions. We performed a simple experiment to test this

TABLE 3.1: Cylinder Radii, Optimized Height and Evaluation Scores

Radius of Cylinder (cm)	Optimal Height (cm)	Evaluation Score at Optimal Height	Evaluation Score at height 26.5 cm
4.5	25.28	0.5794	0.5457
5.5	23.48	0.5338	0.4947
8	30.52	0.5448	0.5313

hypothesis. Considering an over-simplified geometric primitive for a leg, we use cylinders of different radii and perform RGB-D scans using trajectories with constant heights. The results, as seen in Table 3.1, show different optimal heights obtained by using the evaluation score defined in 3.6. We observe that the optimal height differs with different radii and the data confirms our hypothesis. We observe that the evaluation score is significantly lower at a close-to-mean height for all the three cylinders than their optimal height. Additionally, there exists no linear relationship between the radius of the cylinder and the optimal height. Even if this non-linear relationship was to be estimated for the cylinders, there is very low probability that it would hold to varying human anatomies where we observe regions of differing curvatures. This justifies a per-subject RGB-D scanning trajectory optimization. Further, we test varying heights for our RGB-D scanning trajectories as we expect the lower torso and the upper leg region to be uneven in curvature.

For the trajectory optimization module, the orientation of sensor is always anti-parallel to the flat surface on which the subject is kept as shown in Fig. 1.1. The initial and the final poses of the sensor come from the method mentioned in Sec. 3.2 as X_i and X_f. The trajectory is discretized into n segments. Intermediate poses of the start of each segment are given by,

$$\begin{aligned} X_n &= n\lambda X_i + (1 - n\lambda)X_f \\ z_n &= a_0 + a_1 sin(k_1 n\lambda) + a_2 sin(k_2 n\lambda) \\ &\quad + a_3 cos(k_1 n\lambda) + a_4 cos(k_2 n\lambda) \end{aligned} \tag{6}$$

where $X = [x, y]^T$. a_0, a_1, a_2, a_3, a_4 are the coefficients of the parameterized equation. k_1 and k_2 are the frequencies of the sin and cos function. $n\lambda$ represents the nth segment. We optimize using the Cross Entropy Maximization (CEM) optimizer[1], an evolutionary algorithm, to find the optimized values for a_0, a_1, a_2, a_3, a_4.

For every epoch of optimization, we run through the steps 3-5 in this chapter for every agent. Based on the evaluation score calculated, the optimizer selects few best agents and updates the mean and variance of each of the coefficients. We use $N_{agents} = 10$, $N_{best} = 3$, $N_{epochs} = 7$.

3.8 RoI Detection and US Trajectory Generation

The optimized reconstructed surface point cloud is used for RoI detection. We look to start at the inguinal fold and scan along and the inner regions of the leg bifurcation along the left leg (highlighted in Fig. 3.6). The boundary of this RoI is used to generate a US scanning trajectory.

To obtain the RoI defined above, we threshold the cluster point cloud from the final optimized RGB-D scan based on curvature. We use the average curvature of the cluster point cloud and retain only points with curvature above this threshold as part of the RoI. We then find 7 points along the boundary of the RoI on the left leg using the principal axis and use these points to define poses of the ultrasound trajectory. We use the hybrid force position controller + SLERP from the work [3] to scan the leg along these trajectory points. The required surface normals for SLERP computation are calculated by applying [14] to the optimized reconstruction.

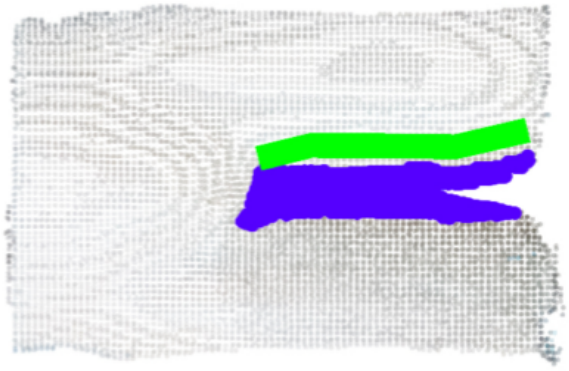

FIGURE 3.6: Leg cluster with detected RoI in purple and US trajectory in green.

Chapter 4

Experiments and Results

The following experiments were conducted on the lower torso ultrasound training phantom.

4.1 Clustering

We test the robustness of the color-based region growing clustering to our application by applying it to the phantom with occlusions on the skin and with multiple objects on the base, near the leg, to replicate real-life medical and surgical scenarios. We provide visual and qualitative results of the clustering performance in Fig. 4.1.

We observe that smaller occlusions, such as clear tubes, needles, catheters etc. are not detected separately and are considered as part of the skin of the subject. This is perhaps due to their small size and/or translucent nature. These occlusions could be considered as part of the scanning surface and be included in RoI and US trajectory generation. This can lead to discontinuous US scans. However, a big positive is the fact that other larger occlusions such as surgical cloth, gloved hands and blood/wounds are identified as not being a part of the subject's skin and excluded from the pipeline for RoI detection and US trajectory generation.

4.2 RGB-D Trajectory Optimization

Our trajectory optimizer converges after the 7th epoch and attains a maximum evaluation score of 0.55. We consider the optimizer to have converged when the maximum standard deviation of the past 5 epochs is below $1.2 * \text{standard deviation of last epoch}$. We check this criterion for the last 6 epochs of our optimization loop.

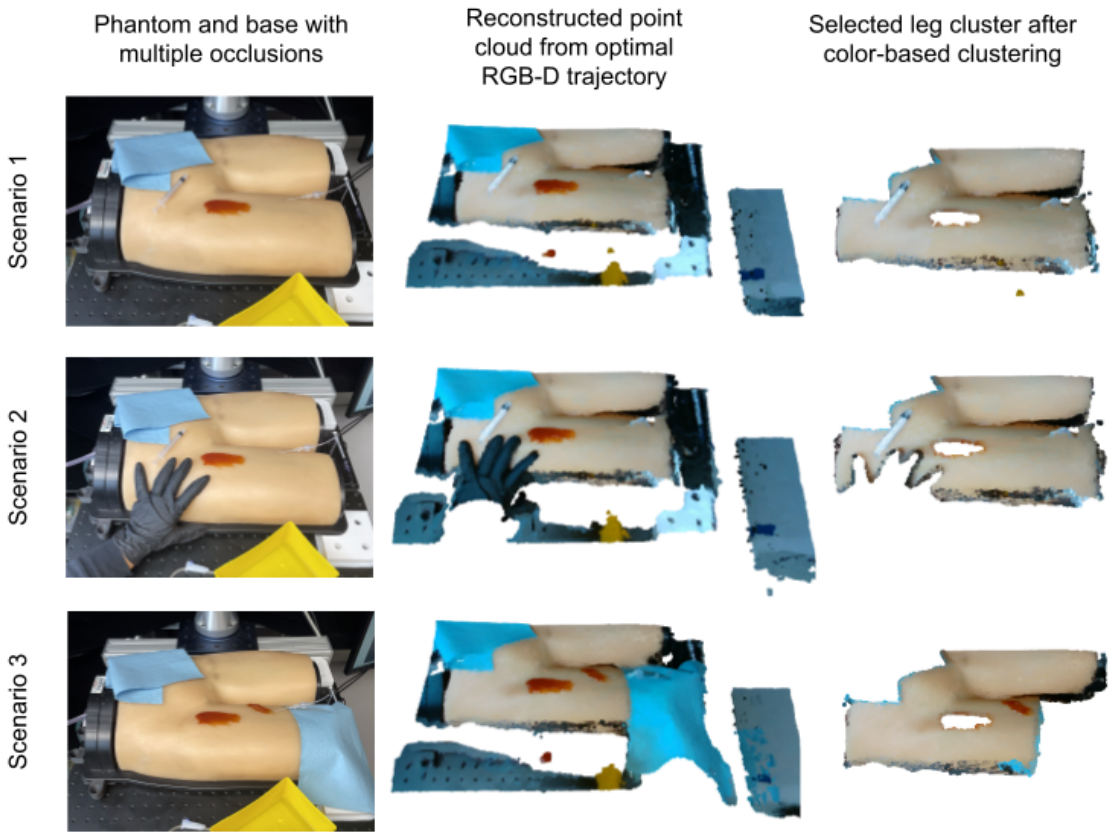

FIGURE 4.1: Multiple scenarios with occlusions on the leg phantom and presented reconstructions with selected leg clusters.

Each epoch takes 2.5 minutes for completion. The plot shown in Fig. 4.2 shows the evaluation score along with its standard deviation at each of the epochs. 4.2 a. shows the rise in evaluation scores with epochs. The standard deviation in evaluation score reduces drastically post epoch 5. It is observed that the height of the optimized RGB-D scanning trajectory is in the range 0.22 - 0.24 m when the minimum and the maximum heights possible are 0.17 m and 0.35 m respectively. We also show the 3D reconstructions generated at epochs 0, 3 and 7 in Fig. 4.3 demonstrating the improvement in point clouds as we reach an optimized RGB-D scanning trajectory.

We also performed an ablation study to find the effect of each of the metrics on the evaluation score. We plot the results in Fig. 4.2 b, c and d. Each metric is important in the evaluation score. Graph b. shows a converging evaluation score with low standard deviation showing that the evaluation score is stable without $Metric_1$. However, there is no clear distinction in evaluation scores over time. This shows that $Metric_1$ is a clear distinguishing factor and and is needed in our evaluation score. Graph c. shows that the evaluation score is unstable without $Metric_2$ as we see varying standard deviations. This signifies that $Metric_2$ is required in the evaluation score calculation. We observe that the evaluation score without $Metric_3$, as seen in graph d. is similar to graph a, marking a lower sensitivity of the evaluation score to $Metric_3$. This is

perhaps due to the skewed penalization in $Metric_3$. However, convergence without $Metric_3$ still takes longer and is seen at Epoch 9.

The reconstructed point cloud from the optimized scan is overlaid on a ground truth point cloud generated from a CT (Computed Tomography) scan of the lower torso ultrasound phantom. We obtained a Chamfer distance CD = 5e-5 m between the reconstructed and the ground truth point clouds.

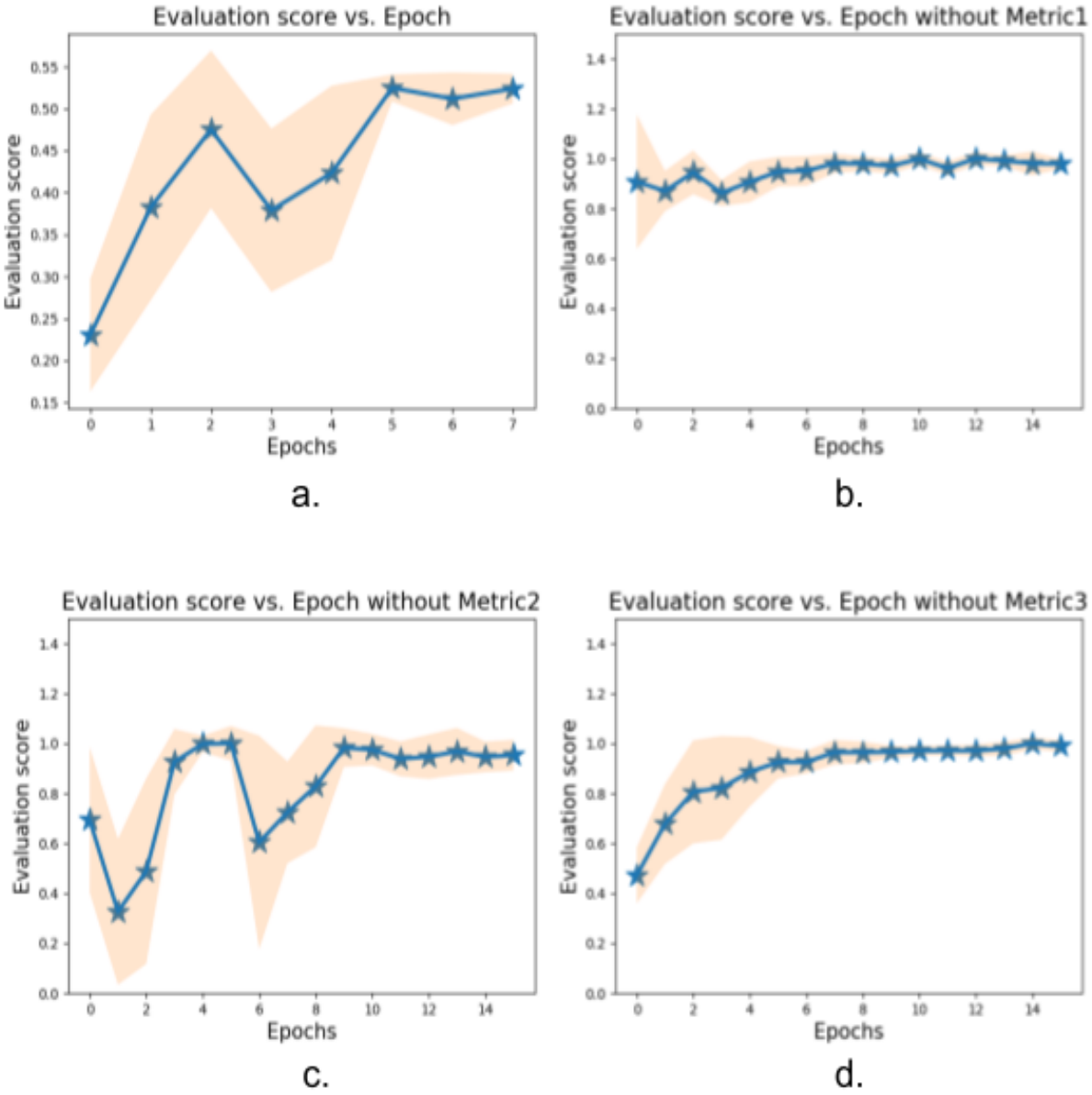

FIGURE 4.2: Plot of the evaluation score vs epochs for the optimization Llop a) with all three metrics; b) without $Metric_1$; c) without $Metric_2$; d) without $Metric_3$; in the evaluation score. The scores in b, c, d are scaled between 0 and 1 for equal comparison.

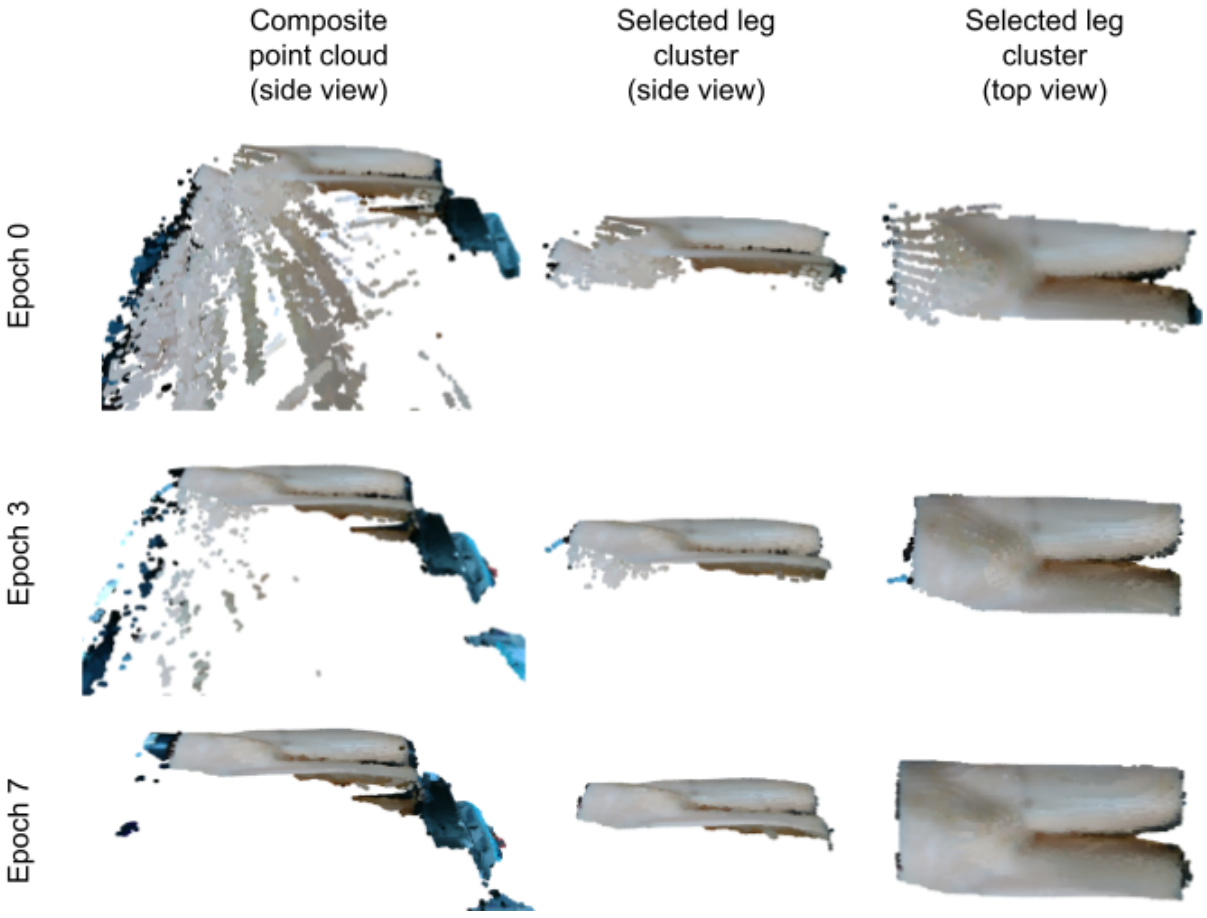

FIGURE 4.3: Improving reconstructed point clouds with increase in epochs in the optimization loop.

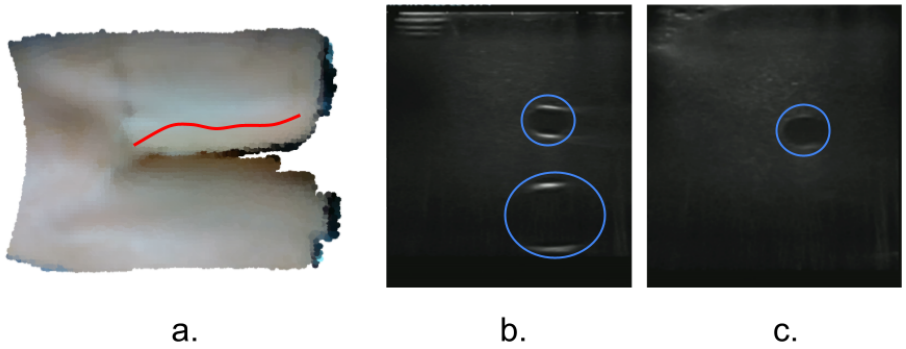

FIGURE 4.4: a) RoI based US trajectory overlaid on the 3D reconstructed surface; b) and c) Femoral vessels captured in our generated US trajectory (blue)

4.3 Ultrasound Scanning

The US trajectory generated from the RoI is used to drive the US probe over the subject. The ultrasound images collected are tested for quality. The trajectory as well as the ultrasound images are shown in Fig 4.4.

We use zero-normalized cross correlation (ZNCC)[9] score to quantify the quality of the ultrasound images. A ZNCC score of 1 represents a perfect scan. The quality of the scan is a direct function of the estimated surface normals which in turn is a function of the reconstructed surface. From the Table 4.1, we see that the ultrasound images collected from the optimized RGB-D scan produce the best ZNCC score compared to the other two RGB-D scans taken at constant heights $h = 0.17m$ and $h = 0.35m$. These heights represents the minimum and maximum scanning height possible for our robot setup. We also present a score from a non-expert user, who was introduced to our setup but had knowledge about the general location of femoral vessels in legs. The femoral vessels are visible in 85% of the ultrasound frames captured from the generated trajectory of our pipeline. Whereas, the femoral vessels are visible in only 70% of the ultrasound frames captured in a trajectory defined by a non-expert user. We provide this comparison as we are aiming for applicability in scenarios where experts are not available.

TABLE 4.1: ZNCC score comparison for various trajectories

US Trajectory From	**ZNCC score of US images**
Optimized RGB-D trajectory	**0.861**
Constant h=17cm RGB-D	0.838
Constant h=35cm RGB-D	0.835
Non-expert	0.791

Chapter 5

Conclusion

Our pipeline successfully localizes the legs of the subject and runs a trajectory optimization module for accurate and optimized RGB-D scanning. This provides us with good RoIs on the surface for US scanning. In the results section, we show that the use of color-based region growing for clustering performs reasonably well in detecting the skin of the subject. Additionally, our evaluation score comprehensively identifies good features in reconstructed point clouds while penalizing bad ones in the optimization process. This helps our RGB-D trajectory optimization module converge the fastest (corroborated by the ablation study). Finally, the RoI detected traces the femoral vessels fairly well while producing good quality US scan images. We have achieved near autonomy in ultrasound scanning using RGB-D sensing

Our pipeline is an important step in the complete automation of an RUS for scanning femoral vessels. Our work ties external anatomy to internal scanning. To test the robustness and generalization of our pipeline in real-world scenarios, live-pig experiments are planned in the near future. We have optimized the height of the RGB-D sensor's scanning trajectory but we plan on optimizing sensor orientation in addition to height. This can provide a higher area coverage of the subject's skin. Currently, our pipeline generates a US RoI based on the external surface and maps femoral vessels 85% of the time. To increase the visibility percentage of the femoral vessels, we plan on incorporating internal anatomy to our pipeline.

Bibliography

[1] Lih-Yuan Deng. "The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning". In: *Technometrics* 48.1 (2006), pp. 147–148. DOI: 10.1198/tech.2006.s353. eprint: https://doi.org/10.1198/tech.2006.s353. URL: https://doi.org/10.1198/tech.2006.s353.

[2] 2018 Siemens Medical Solutions USA Feb 27. *Region of interest placement for quantitative ultrasound imaging.* URL: https://patents.justia.com/patent/11006926.

[3] Raghavv Goel et al. "Autonomous Ultrasound Scanning using Bayesian Optimization and Hybrid Force Control". In: *2022 International Conference on Robotics and Automation (ICRA)*. 2022, pp. 8396–8402. DOI: 10.1109/ICRA46639.2022.9812410.

[4] Christoph Graumann et al. "Robotic ultrasound trajectory planning for volume of interest coverage". In: *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2016, pp. 736–741.

[5] Christoph Hennersperger et al. "Towards MRI-based autonomous robotic US acquisitions: a first feasibility study". In: *IEEE transactions on medical imaging* 36.2 (2016), pp. 538–548.

[6] Qinghua Huang, Jiulong Lan, and Xuelong Li. "Robotic arm based automatic ultrasound scanning for three-dimensional imaging". In: *IEEE Transactions on Industrial Informatics* 15.2 (2018), pp. 1173–1182.

[7] Jan-Philipp Kaiser et al. "Framework for simulation-based Trajectory Planning and Execution of Robots equipped with a Laser Scanner for Measurement and Inspection". In: *Procedia CIRP* 103 (2021). 9th CIRP Global Web Conference â€“ Sustainable, resilient, and agile manufacturing and service operations : Lessons from COVID-19, pp. 292–297. ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2021.10.047. URL: https://www.sciencedirect.com/science/article/pii/S221282712100888X.

[8] Yeoun-Jae Kim et al. "A Sarcopenia Detection System Using an RGB-D Camera and an Ultrasound Probe: Eye-in-Hand Approach". In: *Biosensors* 11.7 (2021), p. 243.

[9] Jerome Martin and James L Crowley. "Experimental comparison of correlation techniques". In: *Int. Conf. on Intelligent Autonomous Systems*. Citeseer. 1995.

[10] Frederick A Masoudi et al. "Cardiovascular care facts: a report from the national cardiovascular data registry: 2011". In: *Journal of the American College of Cardiology* 62.21 (2013), pp. 1931–1947.

[11] Samir Merouche et al. "A robotic ultrasound scanner for automatic vessel tracking and three-dimensional reconstruction of b-mode images". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 63.1 (2015), pp. 35–46.

[12] Carsten Moenning and Neil A Dodgson. "Fast marching farthest point sampling for implicit surfaces and point clouds". In: *Computer Laboratory Technical Report* 565 (2003), pp. 1–12.

[13] Morgan Quigley et al. "ROS: an open-source Robot Operating System". In: *ICRA Workshop on Open Source Software*. 2009.

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

[15] Radu Bogdan Rusu and Steve Cousins. "3D is here: Point Cloud Library (PCL)". In: *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China, 2011.

[16] Radu Bogdan Rusu et al. "Towards 3D Point cloud based object maps for household environments". In: *Robotics and Autonomous Systems* 56.11 (2008). Semantic Knowledge in Robotics, pp. 927–941. ISSN: 0921-8890. DOI: https://doi.org/10.1016/j.robot.2008.08.005. URL: https://www.sciencedirect.com/science/article/pii/S0921889008001140.

[17] Adam Schmidt et al. "Some Remarks on the Optimization-Based Trajectory Reconstruction of an RGB-D Sensor". In: *Image Processing and Communications Challenges 7*. Ed. by Ryszard S. Choraś. Cham: Springer International Publishing, 2016, pp. 223–230. ISBN: 978-3-319-23814-2.

[18] Janet Y Tsui et al. "Placement of a femoral venous catheter". In: *NEW ENGLAND JOURNAL OF MEDICINE* 358.26 (2008), e30.

[19] Philip Yoong, Stephen Duffy, and Thomas J Marshall. "The inguinal and femoral canals: A practical step-by-step approach to accurate sonographic assessment". In: *Indian Journal of Radiology and Imaging* 23.04 (2013), pp. 391–395.

[20] Nico Zevallos et al. "Toward Robotically Automated Femoral Vascular Access". In: *2021 International Symposium on Medical Robotics (ISMR)*. 2021, pp. 1–7. DOI: 10.1109/ISMR48346.2021.9661560.

[21] Q. Zhan, Yubin Liang, and Yinghui Xiao. "Color-based segmentation of point clouds". In: *ISPRS Laser Scanning Workshop* 38 (July 2009).