

# A Curvature and Trajectory Optimization-based 3D Surface Reconstruction Pipeline for Ultrasound Trajectory Generation

Ananya Bal<sup>1</sup>, Ashutosh Gupta<sup>2,\*</sup>, FNU Abhimanyu<sup>1,\*</sup>, John Galeotti<sup>1</sup>, Howie Choset<sup>1</sup>

**Abstract**—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. 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. A consolidated pipeline for creating an optimized 3D surface reconstruction of a subject is presented and is used 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 from a ground-truth high-resolution CT scan.

## I. INTRODUCTION

Ultrasound (US) examination is an indispensable tool in quick diagnostic and medical intervention procedures [1]. 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) [2], [3]. Despite its strengths, US scanning has a strong dependence on the skills of trained human professionals (Sonographers) [4]. Well-trained sonographers are not present everywhere, and so a Robotic Ultrasound System (RUS) could provide medical assistance in such cases [4].

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 corrections to the probe’s pose, and (3) providing safe and accurate forces through the probe to maintain diagnosable image quality. In a typical RUS, the Region of Interest (RoI) to scan is defined by input from a skilled operator, e.g. a sonographer or doctor, based on internal as well as superficial anatomical landmarks on the skin [5], [6]. Identifying these landmarks requires 3D understanding of the skin of the subject. An accurate 3D reconstruction can also help provide normals for controlling the ultrasound probe.

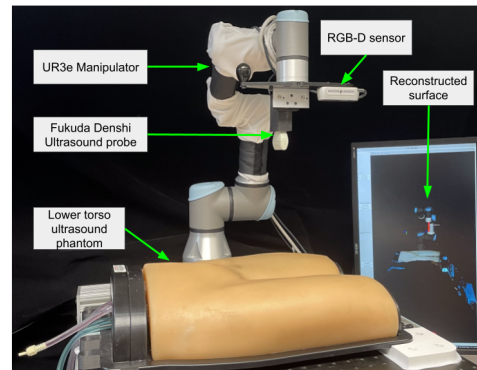


Fig. 1. Our RUS setup, similar to [1], with novelties in evaluation score and optimization for automated US trajectory generation.

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.

In this paper, 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

This work was sponsored in part by US Army Medical contracts W81XWH-19-C0083 and W81XWH-19-C0101

\* Authors of Equal Contribution

<sup>1</sup> are from Robotics Institute, Carnegie Mellon University, Pittsburgh, USA

<sup>2</sup> is from Birla Institute of Technology & Science, Goa Campus, India

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. Our pipeline has been tested on a medical phantom and we have achieved near automation in optimized RGB-D and ultrasound scanning.

Our 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 any RGB-D sensor for surface reconstruction.

The rest of the paper is divided as follows - Section II discusses related work in the areas of RUS, RGB-D based 3D reconstruction, sensor position optimization, and RoI for ultrasound scanning. Section III and IV discuss our methods, results results. Finally, we present our conclusions and future work in Sections V and VI.

## II. RELATED WORK

In, [7], 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 [8] 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 [1] to obtain 3D contours of a lumbar phantom. They obtain a scanning region but are heavily reliant on experience-based rules and physical dimensions from their system and phantom. They also mention parameters which can be varied by the user. Our work is different as we do not depend on the physical dimensions of the scanning region or the RUS components and do not require the user to fine-tune parameters. Another RGB-D-based RUS is presented by [9] to detect sarcopenia in legs. They perform a piece-wise curve fitting to understand the leg surface from 8 point clouds. Their 4 DoF system takes 12 mins to perform RGB-D scanning and curve fitting for US probe positioning. We do not use computation-heavy operations like curve-fitting.

For estimating the optimal trajectory of the RGB-D sensor, there exist methods like [10], which use filtering of multiple RGB frames to optimize the trajectory. This method along with [11] are only shown to work in simulations.

As simulations have controlled environmental factors such as lighting, these methods could have issues transferring to real-world experiments especially in surgical settings. Most RUS systems depend on human operators to define the path that the ultrasound probe should trace. [12] uses an operator defined volume of interest. [13] presents an RUS system to enable needle insertion that needs user defined points on the surface to define the trajectory for maximum vessel coverage. Some works like [4] automate RoI detection with only ultrasound images and do not take into consideration the exterior shape of the subject. Very few systems like [1] automate RoI detection for ultrasound scanning using external surface information.

## III. METHODS

### A. 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 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 [14] and Robot Operating System (ROS) [15] used to combine all of the components.

### B. Architecture Overview

Our pipeline, seen in Fig. 2, has the following modules:

- 1) Clustering for leg localization
- 2) RGB-D trajectory start and end point detection
- 3) Composite surface reconstruction for RGB-D data
- 4) Computation of an evaluation score
- 5) RGB-D sensor trajectory optimization
- 6) RoI detection and ultrasound trajectory generation

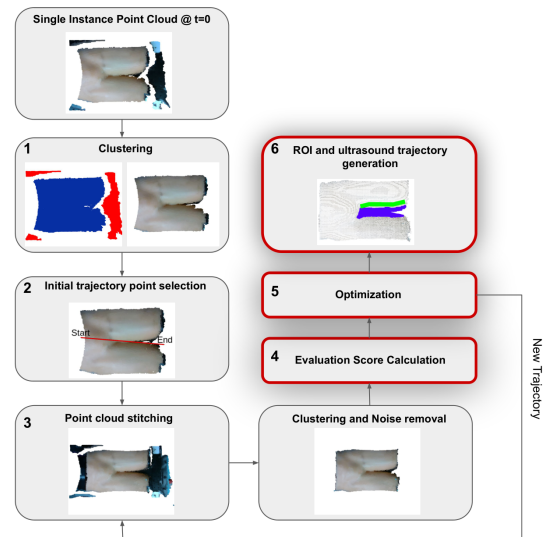


Fig. 2. Surface Reconstruction Pipeline with novel contributions highlighted

We first take a single RGB-D sensor reading and obtain a 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 found, 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.

### C. Point Clustering for Localization of Legs

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. A single instance point cloud, representing the subject on a relatively flat base is then captured. This point cloud is downsampled with a 5 mm voxel size to maintain high processing speed by reasonably reducing the data complexity. The color-based region growing clustering algorithm [16] is used for generating clusters from this downsampled point cloud. We use a distance threshold value of 5, a point color threshold value of 10 for region growing, and a region color threshold value of 10 for merging based on [17], [18]. We use a high point-color threshold to compensate for variations in skin color of 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 legs by finding the cluster  $C^*$  which corresponds to the skin of the subject. This is done by applying a selection criteria to the clusters. Given that the legs have higher curvature than the base and covers the majority of field-of-view, the cluster corresponding to the legs has both the highest average curvature and no. of points.

$$C^* = \underset{i}{\operatorname{argmax}} \left( \alpha \left( \frac{\sum_i N}{\sum_{j=1}^n \sum_j N} \right) + (1 - \alpha) \left( \frac{\sum_i \kappa}{\sum_{j=1}^n \sum_j \kappa} \right) \right) \quad (1)$$

where  $i$  is the index of the selected cluster,  $\sum_i N$  is the total number of points in a cluster  $C_i$  and  $\sum_i \kappa$  is the sum of curvature of the cluster  $C_i$ . There are  $n$  such clusters to choose from.  $\alpha$  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.

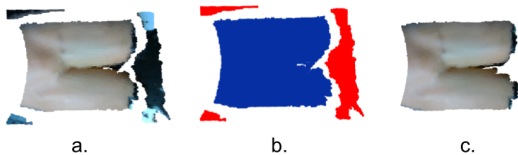


Fig. 3. Results from color-based region growing clustering - a) Single instance point cloud; b) Identified clusters; c) Selected cluster with selection criteria (Eqn. 1) corresponds to the legs

### D. Start and End Point Detection for RGB-D Scanning

Once the legs are 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 (PA) 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 PA. We extend the PA vector in both directions from the centroid of the flattened point cloud. The intersection points of the PA with the edges of the 2D convex hull are found. The  $(x, y)$  coordinates of these points form the coordinates of start and end points of the RGB-D scanning trajectory. The points are visualized in Fig. 4. The  $z$  coordinates (height) of the start and end points are obtained from the optimization module.

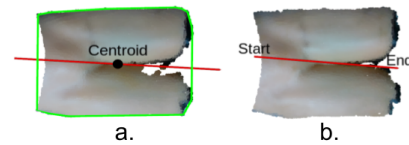


Fig. 4. Steps in start and end point detection - a) Selected cluster point cloud along with the centroid, principal axis (red) and 2D convex hull (green); b) Detected start and end points for RGB-D scanning trajectory along the PA.

### E. 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 accurate transforms from the sensor optical frame to the robot base frame using an extensive CAD model of the system. 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. 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 legs of the subject. This cluster point cloud, and the larger composite point cloud are used for the remaining pipeline.

### F. 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:

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

$$Metric_1 = \frac{\text{No. of inlier points}}{\text{Total No. of points}} \quad (2)$$

in the evaluation score. Outliers are detected by the application of the statistical outlier detection algorithm [19] to the selected cluster. We consider 50 nearest neighbours and a standard deviation of 0.75 for our implementation based on [20], tweaking for more aggressive outlier removal. The inliers are visualized in Fig. 5. Additionally, we consider the points above and below the surface that are removed during clustering to be outliers as well while determining  $Metric_1$ .

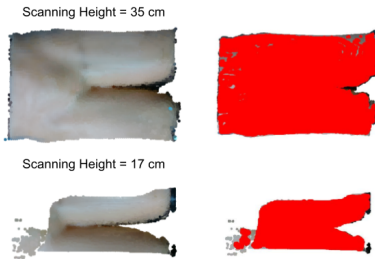


Fig. 5. (Left) Top view of the cluster point cloud, (Right) Computed inliers highlighted in red.

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. 6.



Fig. 6. Top and side view of a cluster point cloud with a concavity.

To quantify our surface reconstruction as a function of continuous regions, we use the 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 the total 2D projected area of the cluster point cloud (Eqn. 3). We obtain these quantities by computing the area of the concave (with alpha shapes) and convex hulls after 2D projection.

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

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 [21]. 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

TABLE I  
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

metric, which is the normalized standard deviation of all the chosen normals is given by

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

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,

$$EvaluationScore = \beta_1 Metric_1 + \beta_2 Metric_2 - \beta_3 Metric_3 \quad (5)$$

Here, each metric is weighted equally at 0.33 as reweighting could disproportionately skew the values of the metrics as they have large variances due to outliers.

### G. RGB-D Scanning Trajectory Optimization

1) *Motivation*: 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 hypothesis. Considering an over-simplified geometric primitive for a leg, we use cylinders of different radii, color and lengths, and perform RGB-D scans using trajectories with constant heights. The results in Table I, show different optimal heights obtained by using the evaluation score defined in III-F. 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. This non-linear relationship between scan quality and distance of sensor from object is documented in [22] and holds true for multiple RGB-D sensors. This further motivates sensor-agnostic scanning trajectory optimization. Even if this non-linear relationship was to be estimated for cylinders, there is very low probability that it would hold to varying human anatomies with 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. The initial and the final poses of the sensor come from the method in Sec. III-B as  $X_i$  and  $X_f$ . The trajectory is discretized into  $n$  segments. Intermediate poses of the start of each segment are given by,

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

where  $X = [x, y]^T$ .  $a_{0-4}$  are the coefficients of the parameterized equation.  $k_1 = 1$  and  $k_2 = 2$  are the frequencies of the sin and cos functions.  $n\lambda$  represents the  $n$ th segment.  $2^{nd}$  and  $3^{rd}$  order polynomials for  $z_n$  in terms  $X_n$  made the coefficients vary in multiple orders of magnitude, making the state space for the optimization very large. We therefore choose the Fourier basis function [23] and represent  $z_n$  independent of  $X_n$ . We optimize using the Cross Entropy Maximization (CEM) optimizer [24], an evolutionary algorithm, to find the optimized values for  $a_{0-4}$ . For every epoch of optimization, we run through the steps 3-5 in Sec III 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  $N_{agents} = 10$ ,  $N_{best} = 3$ ,  $N_{epochs} = 7$ .

#### H. 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. 7). The boundary of this RoI is used to generate a US scanning trajectory.

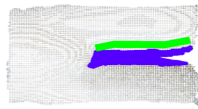


Fig. 7. Leg cluster with detected RoI in purple and US trajectory in green.

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 US trajectory. We use the hybrid force position controller + SLERP from the work [4] to scan the leg along these trajectory points. The required surface normals for SLERP computation are calculated by applying [25] to the optimized reconstruction.

### IV. EXPERIMENTS AND RESULTS

#### A. 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. 8. 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.

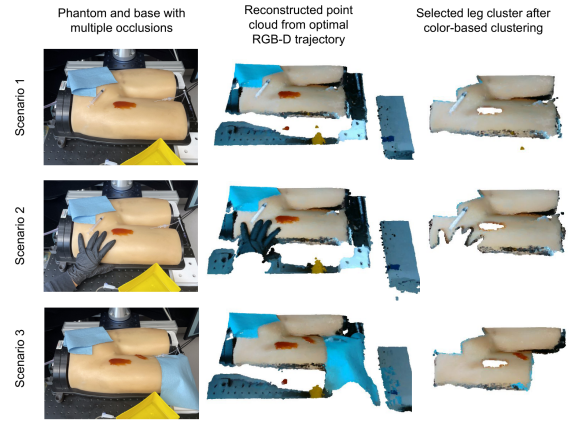


Fig. 8. Multiple scenarios with occlusions on the leg phantom and presented reconstructions with selected leg clusters.

#### B. 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.

Each epoch for scanning the whole region takes 2.5 mins for completion. The plot in Fig. 9 shows the evaluation score along with its standard deviation at each of the epochs. 9 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  $\sim 0.22 - 0.24$  m when the minimum and the maximum heights possible are 0.17 and 0.35 m respectively. We also show the 3D reconstructions generated at epochs 0, 3 and 7 in Fig. 10 demonstrating the improvement in point clouds as we reach an optimized RGB-D scanning trajectory.

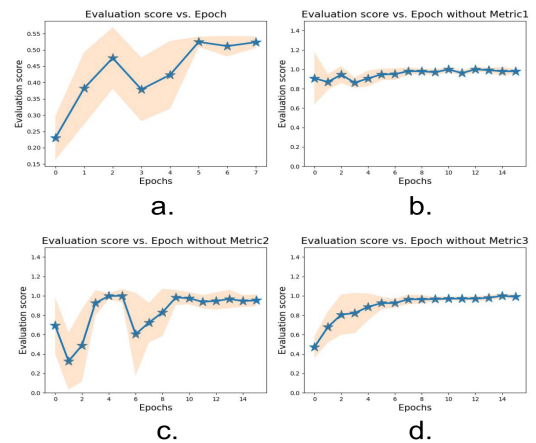


Fig. 9. Evaluation score vs epochs for the optimization loop a) with all three metrics; b) without  $Metric_1$ ; c) without  $Metric_2$ ; d) without  $Metric_3$ ; in the evaluation score. The scores in b, c, d are scaled to  $[0, 1]$  for equal comparison.

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. 9 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 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 onto a ground truth point cloud generated from a CT (Computed Tomography) scan of the phantom. We obtained a Chamfer distance  $CD = 5e-5$  m between the two.

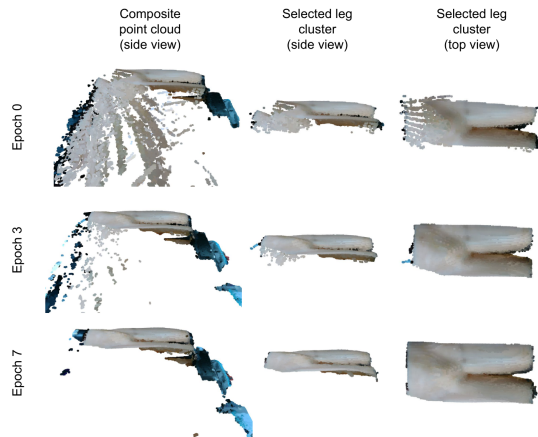


Fig. 10. Improvement in reconstructions within the optimization loop.

### C. Ultrasound Scanning

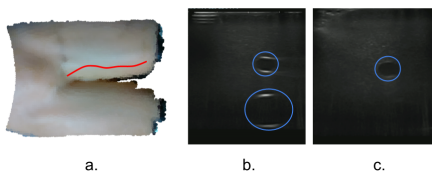


Fig. 11. a) ROI based US trajectory overlaid on the reconstructed surface; b) and c) Femoral vessels captured by generated US trajectory (blue)

The US trajectory generated from the ROI is used to drive the US probe over the subject. The US images collected are tested for quality. The trajectory as well as the ultrasound images are shown in Fig 11. Higher image similarity implies higher consistency in US imaging which implies smoother scan. We use the zero-normalized cross correlation (ZNCC) score [26] as a metric to quantify this consistency as a quality measure for US scanning [4]. 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 II, 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 the minimum and maximum constant heights  $h = 0.17m$  and  $h = 0.35m$ . 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. The femoral vessels are visible in 85% of the US frames captured from our generated trajectory as compared to 70% of the US frames captured from the non-expert defined trajectory. We provide this comparison as we aim for applicability in scenarios where experts are not available.

TABLE II

ZNCC SCORE COMPARISON FOR VARIOUS TRAJECTORIES

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

## V. CONCLUSION

Our pipeline successfully localizes legs of a 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 Section IV, we show that 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 images. We have achieved near autonomy in ultrasound scanning using RGB-D sensing.

## VI. DISCUSSION AND FUTURE WORK

Our pipeline ties external anatomy to internal scanning and is an important step in the complete automation of an RUS for scanning femoral vessels. Since our optimization is not contingent on camera parameters, our method is sensor agnostic and can be applied to other RGB-D sensors. In  $\sim 15$  mins, we can determine an optimal trajectory for RGB-D scanning of legs. While this takes longer than placing the sensor at a few optimal poses and stitching captured point clouds for obtaining a high fidelity 3D reconstruction, the latter is highly subject to the skill of the operator. Our method adds autonomy without assuming any knowledge of the RGB-D sensor or the dimensions of the subject.

To test the generalization of our pipeline to real-world scenarios, live-pig experiments are planned. We plan to optimize the RGB-D sensor orientation in addition to the 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 only. To increase the visibility percentage of the femoral vessels, we plan on incorporating knowledge of internal anatomy.

## REFERENCES

- [1] Q. Huang, J. Lan, and X. Li, "Robotic arm based automatic ultrasound scanning for three-dimensional imaging," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1173–1182, 2018.
- [2] J. Y. Tsui, A. B. Collins, D. W. White, J. Lai, and J. A. Tabas, "Placement of a femoral venous catheter," *NEW ENGLAND JOURNAL OF MEDICINE*, vol. 358, no. 26, p. e30, 2008.
- [3] F. A. Masoudi, A. Ponirakis, R. W. Yeh, T. M. Maddox, J. Beachy, P. N. Casale, J. P. Curtis, J. De Lemos, G. Fonarow, P. Heidenreich *et al.*, "Cardiovascular care facts: a report from the national cardiovascular data registry: 2011," *Journal of the American College of Cardiology*, vol. 62, no. 21, pp. 1931–1947, 2013.
- [4] R. Goel, F. Abhimanyu, K. Patel, J. Galeotti, and H. Choset, "Autonomous ultrasound scanning using bayesian optimization and hybrid force control," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 8396–8402.
- [5] Y. Labyed, "Region of interest placement for quantitative ultrasound imaging," May 18 2021, uS Patent 11,006,926.
- [6] P. Yoong, S. Duffy, and T. J. Marshall, "The inguinal and femoral canals: A practical step-by-step approach to accurate sonographic assessment," *Indian Journal of Radiology and Imaging*, vol. 23, no. 04, pp. 391–395, 2013.
- [7] S. Merouche, L. Allard, E. Montagnon, G. Soulez, P. Bigras, and G. Cloutier, "A robotic ultrasound scanner for automatic vessel tracking and three-dimensional reconstruction of b-mode images," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 63, no. 1, pp. 35–46, 2015.
- [8] C. Hennersperger, B. Fuerst, S. Virga, O. Zettinig, B. Frisch, T. Neff, and N. Navab, "Towards mri-based autonomous robotic us acquisitions: a first feasibility study," *IEEE transactions on medical imaging*, vol. 36, no. 2, pp. 538–548, 2016.
- [9] Y.-J. Kim, J. Choi, J. Moon, K. R. Sung, and J. Choi, "A sarcopenia detection system using an rgb-d camera and an ultrasound probe: Eye-in-hand approach," *Biosensors*, vol. 11, no. 7, p. 243, 2021.
- [10] A. Schmidt, M. Kraft, D. Belter, and A. Kasiński, "Some remarks on the optimization-based trajectory reconstruction of an rgb-d sensor," in *Image Processing and Communications Challenges 7*, R. S. Choraś, Ed. Cham: Springer International Publishing, 2016, pp. 223–230.
- [11] J.-P. Kaiser, S. N. Becker, M. Wurster, N. Stricker, and G. Lanza, "Framework for simulation-based trajectory planning and execution of robots equipped with a laser scanner for measurement and inspection," *Procedia CIRP*, vol. 103, pp. 292–297, 2021.
- [12] C. Graumann, B. Fuerst, C. Hennersperger, F. Bork, and N. Navab, "Robotic ultrasound trajectory planning for volume of interest coverage," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 736–741.
- [13] N. Zevallos, E. Harber, Abhimanyu, K. Patel, Y. Gu, K. Sladick, F. Guyette, L. Weiss, M. R. Pinsky, H. Gomez, J. Galeotti, and H. Choset, "Toward robotically automated femoral vascular access," in *2021 International Symposium on Medical Robotics (ISMR)*, 2021, pp. 1–7.
- [14] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 9-13 2011.
- [15] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, 2009.
- [16] Q. Zhan, Y. Liang, and Y. Xiao, "Color-based segmentation of point clouds," *ISPRS Laser Scanning Workshop*, vol. 38, 07 2009.
- [17] J. Sun, B. Peng, C. C. Wang, K. Chen, B. Zhong, and J. Wu, "Building displacement measurement and analysis based on uav images," *Automation in Construction*, vol. 140, p. 104367, 2022.
- [18] I. I. Bai and R. R. Nidamanuri, "Three dimensional urban building detection using lidar data."
- [19] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3d point cloud based object maps for household environments," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 927–941, 2008, semantic Knowledge in Robotics. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0921889008001140>
- [20] "Removing outliers using a statistical outlier removal filter." [Online]. Available: [https://pcl.readthedocs.io/projects/tutorials/en/latest/statistical\\_outlier.html](https://pcl.readthedocs.io/projects/tutorials/en/latest/statistical_outlier.html)
- [21] C. Moenning and N. A. Dodgson, "Fast marching farthest point sampling for implicit surfaces and point clouds," *Computer Laboratory Technical Report*, vol. 565, pp. 1–12, 2003.
- [22] A. Vit and G. Shani, "Comparing rgb-d sensors for close range outdoor agricultural phenotyping," *Sensors*, vol. 18, no. 12, p. 4413, 2018.
- [23] J. A. Drallmeier, J. B. Siegel, and A. G. Stefanopoulou, "Iterative learning-based trajectory optimization using fourier series basis functions," *IEEE Control Systems Letters*, vol. 6, pp. 2180–2185, 2022.
- [24] L.-Y. Deng, "The cross-entropy method: A unified approach to combinatorial optimization, monte-carlo simulation, and machine learning," *Technometrics*, vol. 48, no. 1, pp. 147–148, 2006. [Online]. Available: <https://doi.org/10.1198/tech.2006.s353>
- [25] R. B. Rusu, "Semantic 3d object maps for everyday manipulation in human living environments," Ph.D. dissertation, Computer Science department, Technische Universitaet Muenchen, Germany, October 2009.
- [26] J. Martin and J. L. Crowley, "Experimental comparison of correlation techniques," in *Int. Conf. on Intelligent Autonomous Systems*. Citeseer, 1995.