A Novel Robotic System for Ultrasound-guided Peripheral Vascular Localization

Guangshen Ma*,[1], Siobhan R. Oca*,[1], Yifan Zhu[2],
Patrick J. Codd, MD[1,3] and Daniel M. Buckland, MD, PhD[1,4]

***Abstract*— In this paper, we present an autonomous RGB-D and 2D ultrasound-guided robotic system for collecting 3D localized volumes of peripheral vessels. This compact design, with available commercial components, lends itself to platform utility throughout the human body. The fully integrated system works with force limits for future safety in human use. We propose a PID force controller for smooth and safe robot scanning following *a priori* 3D trajectory generated from a surface point cloud. System calibration is implemented to determine transformations among sensors, end-effector and robot base. A vascular localization pipeline that consists of detection and tracking is proposed to find the 3D vessel positions in real-time. Precision tests are performed with both predesignated and autonomously selected areas in an arm phantom. The average variance of the autonomously collected ultrasound images (to construct 3D volumes) between repeated tests is shown to be around 0.3 mm, similar to the theoretical spatial resolution a clinical ultrasound system. This fully integrated system demonstrates the capability of autonomous collection of peripheral vessels with built-in safety measures for future human testing.**

I. Introduction

A crucial step in the diagnosis or treatment of any patient is access to the blood in their veins. For vessels that are hard to identify at the skin surface, clinicians use an ultrasound (US) to identify and track veins for catheter placement. Autonomous robotic vessel localization, including detection and tracking via US guidance, can provide a more accurate and reliable process for patients, which does not require advanced training for the clinician.

Robotic controlled US has been widely employed in research, but differs greatly in intended use, autonomy from the human controller, and flexibility of use case. US guidance was used for multiple robot assisted medical procedures such as prostatectomy, breast biopsies and carotid artery tracking [1], [2]. Most of these platforms differ from the proposed system because they rely significantly on the controller and the clinician for expertise and input. Many systems either require a controller to identify the key targets of the system or utilize teleoperation with shared control placement of the probe. Therefore, the autonomy of these robotic systems are limited, as the presence of a clinician is necessary to navigate and identify correct structures. For intervention without a clinician, the system must be able to perform certain tasks autonomously. This is only possible if the robot-controlled US can safely navigate the surface of the body without injury-causing force on the patient's body, while accomplishing the objectives of the procedure.

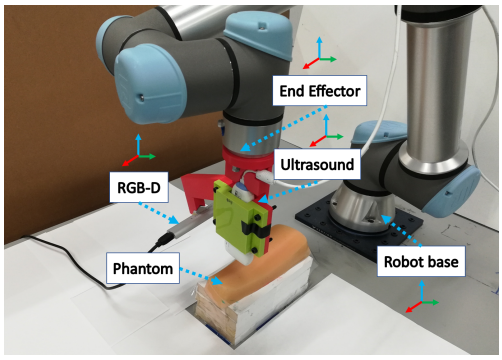


Fig. 1: System Setup: Labelled frames of RGB-D camera, ultrasound probe, UR5 robot base frame and end-effector.

Robotic controlled US systems that are autonomously force modulated have been developed [2], [3], [4], but many of these systems are specialized for a specific and limited surface of the patient. With the range and accuracy of current commercially available robotic arms [5], a more flexible system may be possible for different vision-based control tasks such as finding and tracking various vessels in the body through precise control of the US probe. Force modulation can be achieved with specific sensors on the end effector (EE) or as part of the system, as with the Universal Robotics UR5e [5]. For example, Mathaissen et al. demonstrated that it is possible to use the UR5e robotic arm for US guided tele-robotic procedures [6].

Building on an effective robot controlled US, autonomous vascular localization mainly includes two tasks: vessel detection and vessel tracking [7], [8]. Vessel detection aims to search for a vessel, determine its location, and use this proposed region to initialize the vessel tracking process. Given a vessel candidate, the goal of vessel tracking is to find the 3D coordinates of the vessel contour in each frame and transform to the global coordinates, based on the unique features of the vessels in the US image. The main visual characteristics of the vasculature is lower pixel intensity inside the vessel boundary and higher intensity outside the edge, as the echoes reflected back to the US encounter different material properties between the blood,

Supported by NASA Grant 80NSSC20K14337 and Duke MedX.

∗ Denotes equal contribution.

[1] Department of Mechanical Engineering and Materials Science, Duke University. Email: {guangshen.ma, siobhan.rigby}@duke.edu

[2] Department of Computer Science, University of Illinois at Urbana-Champaign.

[3] Department of Neurosurgery, Duke University.

[4] Department of Surgery, Division of Emergency Medicine, Duke University.

muscle, and fatty tissue. These unique features enable the use of Active Contour and its variants to track the object contours with a level-set formulation [7], [8], [9], [10]. Moreover, recent progress in deep learning shows good performance for various tasks of 2D image analysis, such as U-net, a powerful network for medical image segmentation [11]. Integration of deep learning and Active Contour method shows potential applications for autonomous vascular scanning. Recent work in [12] proposed a Convolutional Long Short Term Memory network to segment the vessel by using temporal features, demonstrating the efficacy of deep learning for vessel analysis in US images.

Related work that found target peripheral vasculature includes a robotic venipuncture prototype which integrates a near-infrared and US imaging with a 7-degree-of-freedom needle insertion subsystem by Chen et al. [3], [13]. However, this system is limited to surface vessels that can be identified by an infrared camera and requires a custom robotic system only used in cannulation of peripheral vessels. Chen et al. also proposed a deep learning strategy for fully autonomous vascular access with B-mode and color doppler image sequences by using a recurrent convolutional encoder-decoder network [13]. This is different from our work since only the B-mode image sequences are used in the study. To address these limitations, the proposed system is comprised of an integrated US probe and RGB-D camera at the robotic arm EE. This system is able to capture an ***a priori*** point cloud, generate a scanning trajectory and perform a fully automatic vessel localization in real-time by adjusting the force with a PID controller. The main contributions of this letter are summarized as:

1) System integration of an ultrasound probe and a RGB-D camera to a 6-DOF robotic arm for performing autonomous peripheral vascular localization.
2) A PID control strategy that enables the robot to move smoothly with a safe force threshold on an arm phantom.
3) A 3D vessel localization system for automated vessel detection, tracking and contour reconstruction.

II. Method

A. System Hardware

This system includes three main components: a UR5e robotic arm (Universal Robots, Denmark), Interson US probe (Interson Corporation, CA), and a Realsense SR305 RGB-D camera (Intel Corporation, CA), as shown in Fig. 1. The US probe and RGB-D camera are attached to the robotic arm's EE with a 3D-printed stage, which has fixed the physical positions of the sensors. The built-in UR5e Force-torque is utilized to measure the robot force contacts.

The surface image is retrieved by the Intel Realsense SR305. With a range of 0.2-1.5 m and up to 640×480 resolution at 60 frames per second (fps), it sufficiently depicts the stationary surface of the tissue. This surface determines the orientation of the US probe for smooth and safe movement across the tissue for a given trajectory. An initial image is taken to capture the full surface, assuming no movement of the tissue throughout the procedure. This is required as the camera, connected to the robot arm EE, cannot have good visualization and be 0.2m from the tissue surface during US collection. The US image data is collected with Interson's (SP-101) USB US Imaging Probe, with 7.5 MHz and 5 cm depth range. The low frequency (7.5 MHz) enables penetration to deeper tissue and it is widely used for vessel detection in human arm. Low frequency (LF) B-scan US video is converted to a sequence of images at 30 fps. Each LF image has dimension as 754×494 pixels. The resolution and speed of the US probe enable fast image analysis during the robot scanning procedure.

B. System Calibration

For successful vascular localization, system calibration aims to find precise transformations between various coordinate frames. The base of robotic arm is defined as the global world frame denoted as $\{W\}$ (Fig. 3.(f)). The frames of robot end-effector and ultrasound probe center tip are described as $\{EE\}$ and $\{US\}$ respectively (Fig. 3.(e)). The goal of calibration is to find the frame transformations of $T_{US}^{EE} = (R_{US}^{EE}, t_{US}^{EE})$ and $T_{EE}^{W} = (R_{EE}^{W}, t_{EE}^{W})$, such that a pixel $p(u, v)$ in the ultrasound image plane can be transformed to $\{W\}$:

$$p^W = R_{EE}^W(R_{US}^{EE}[u \cdot \gamma, v \cdot \gamma, 0]^T + t_{US}^{EE}) + t_{EE}^W \quad (1)$$

where R_{EE}^W and R_{US}^{EE} are 3 x 3 rotation matrix, and t_{US}^{EE} and t_{EE}^W are 3 x 1 translation vectors. p^W is the 3D coordinate of a pixel defined in $\{US\}$. The coefficient γ is a ratio between the pixel distance and millimeter for US images. T_W^{EE} can be calculated by forward kinematics of the robot arm.

1) RGB-D Camera Calibration: RGB-D camera calibration is a common procedure in robotics with the aim of finding the optimal transformation between the depth camera and the robotic EE. The RGB-D camera calibration was implemented with an iterative approach to solve the Hand-Eye calibration problem [14] [15].

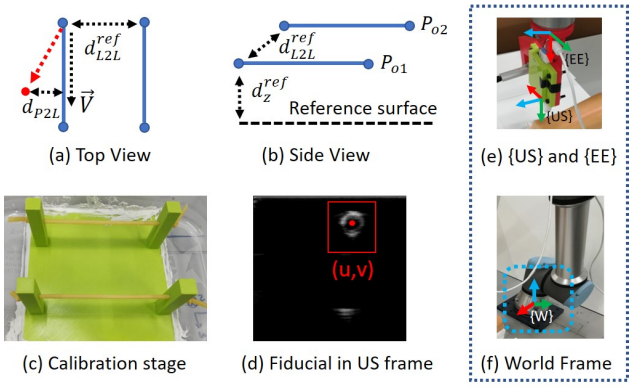


Fig. 3: Calibration stage and coordinate frames.

2) Ultrasound Calibration: Ultrasound-guided robotic system requires spatial calibration between the robot end-effector and the ultrasound image plane. We refer to the N-wire US calibration method proposed in [16], [17] and made

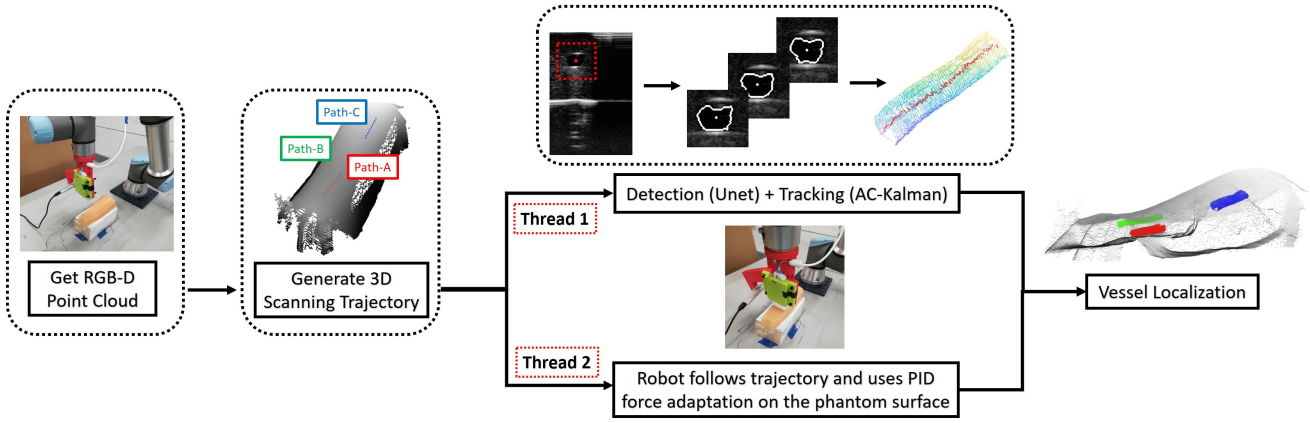


Fig. 2: The pipeline of the ultrasound and robotic guided vascular localization. The labelled Path-A, Path-B and Path-C are the individual scanning regions for each precision test. Thread 1 and 2 represent a parallel procedure of robot control and vision tracking. AC-Kalman denotes the Active-Contour-Kalman-Filter vision framework for vessel localization. The last image on the right side shows the surface point cloud and 3D reconstructed vascular contours.

improvements through design of a new calibration stage. The N-wire calibration uses a phantom with N cross-section wires that can be detected by a freehand US probe. However, this method is limited by a requirement of an external optical tracker as well as a complex design of a calibration stage. We introduce an improved US calibration method with a simpler design of the calibration stage (Fig. 3 (c)).

US calibration aims to find $\{EE\}$ to $\{US\}$ transformation such that the configuration of the reconstructed tube locations in the US image (Fig. 3 (d)) matches the geometry of the calibration stage. To solve the US calibration problem, we formulate an optimization problem by defining three geometric error constraints: point-to-line E_{P2L}, point-to-plane E_{P2Z} and line-to-line E_{L2L}. E_{P2L} penalizes the distance between the reprojection point and the tube center line; since the two tubes in the calibration stage are parallel to the table surface (reference plane), E_{P2Z} measures the length between the detected fiducial (at the tube) and the reference plane; E_{L2L} is the geometric constraint of two parallel tubes where the line-to-line distances should be fixed. US calibration requires a collection of sampled data from various robot configurations in 3D for the generality of the method. The optimization aims to minimize the summation of E_{P2Z}, E_{P2L} and E_{L2L} among all the sampled data:

$$\begin{aligned}\underset{T_{US}^{EE},\overrightarrow{V},P_{o,1},P_{o,2}}{\arg\min} \quad & \sum_{i=1,2}\sum_{j=1}^{m} \underbrace{||d_z(P_{i,j}^W) - d_z^{ref}||_2}_{E_{P2Z}} + \\ & \sum_{i=1,2}\sum_{j=1}^{m} \underbrace{||(P_{i,j}^W - P_{o,i}) - [(P_{i,j}^W - P_{o,i})\cdot\overrightarrow{V}]\cdot\overrightarrow{V}||_2}_{E_{P2L}} + \\ & \underbrace{||(P_{o,2} - P_{o,1}) - [(P_{o,2} - P_{o,1})\cdot\overrightarrow{V}]\cdot\overrightarrow{V}||_2 - d_{L2L}^{ref}}_{E_{L2L}},\end{aligned} \tag{2}$$

Where $P_{o,i}$ is the initial position of the i-th tube. $P_{i,j}^W$ is computed by Eqn. 1 based on the fiducial pixel center (u, v) at the i-th tube and the j-th point (Fig. 3 (b) and (d)). $\overrightarrow{V}$ represents the same direction vector of the two parallel tubes. $d_z(P_{i,j}^W)$ is the Z-axis coordinate to denote the point-to-plane distance from the reference surface. d_z^{ref} describes the ground truth distance between the reference surface and the height of the tube. Similar to d_z^{ref}, d_{L2L}^{ref} depicts the fixed reference distance between the two parallel tubes. The formulation in (2) includes 15 unknown parameters and the derivative of the objective function can be approximated by Finite-difference. The Python Scipy Optimization Package is used to find the optimal solution via Broyden–Fletcher–Goldfarb–Shanno (BFGS) method [18]. With a reasonable initial guess, a local minimum can be determined.

C. Robot Control Strategy

Autonomous vascular localization requires a robust robot control strategy to move the US probe on a deformable surface. The robot must minimize jolting movement and continuously adapt the force to maintain constant pressure on the surface. Physical robot movement follows a surface trajectory, with proposed positions and normals generated from the *a priori* point cloud data, and force control, which adapts dynamically to the surface-based forces in the EE force torque sensor.

The adjustment of the trajectory via force input is critical for safe scanning to avoid potential deformation of the vessel in the arm. A robust PID force controller can maintain adequate contact to the surface based on the point cloud data with noise and bias in visual perception. Therefore, we propose a vessel scanning approach using a PID controller and an *a priori* surface point cloud.

1) Trajectory from Point Cloud: The system first moves to an initial position and takes a RGB-D image of the arm phantom. The depth data is filtered to eliminate outliers and segmented to select points that lie above the table surface by at least 5 cm. The processed point cloud is denoted as $\overrightarrow{P} = \{p_1, p_2, \cdots, p_n\}, p_i \in \mathbb{R}^3$. The surface normals are estimated

from $\overrightarrow{P}$ and referred as $\overrightarrow{V_N} = \{\overrightarrow{v_1}, \overrightarrow{v_2}, \cdots, \overrightarrow{v_n}\}, \overrightarrow{v_i} \in \mathbb{R}^3$. A pre-defined 3D trajectory is applied to the surface point cloud data. To eliminate jerky motion and generate a smooth trajectory, a basis spline (B-spline) function is made from the surface positions $\overrightarrow{P}$ and normals $\overrightarrow{V_N}$. The interpolated curves are denoted as $f_P(t) : \mathbb{R} \rightarrow \mathbb{R}^3$ and $f_N(t) : \mathbb{R} \rightarrow \mathbb{R}^3$, with $t \in [0, 1]$ as the ratio of the time step.

2) Force PID control: The force is modulated by adjusting the position of the US probe towards the surface normal $f_N(t)$ at the 3D trajectory $f_P(t)$, and we refer this offset distance as penetration depth which is denoted as Δd. The target force is set as $F_{target} = 3.5N$ with a danger force threshold defined as $5.0N$. A PID controller takes the force error as input and outputs the penetration depth for the next move (Fig. 4). To determine the current force with sensory noise, a 5th order low-pass Butterworth filter is applied to continuously utilize the past 50 measurements for force control.

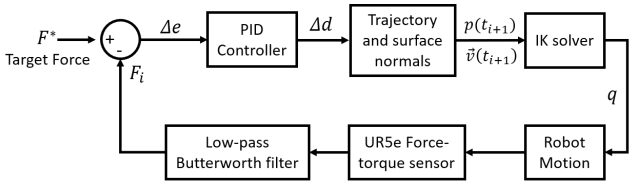
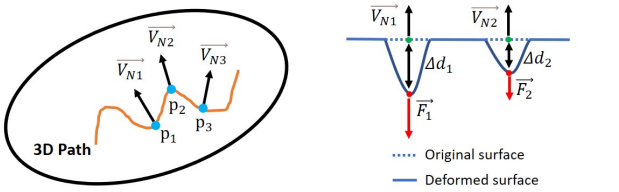


Fig. 4: Block diagram of the PID force controller. $p(t_{i+1})$ and $\overrightarrow{v}(t_{i+1})$ are the new position and orientation for the IK solver, and q is the corresponding robot configuration.

For the consecutive motions in the same trajectory, we have $p(t_{i+1}) = f_P(t_{i+1}) + \Delta d \cdot \overrightarrow{v}(t_{i+1})$ with $\overrightarrow{v}(t_{i+1}) = f_N(t_{i+1})$. Δd is the control output from the PID controller based on force error. The controller adjusts the error between the current and the target force by modulating Δd at each time step. For example, Fig. 5 (b) illustrates a simple 2D projection model for the PID controller where the force can be adjusted by controlling the penetration depth.



(a) 3D surface positions and normal p_i and $\overrightarrow{V_{Ni}}$ from B-spline functions

(b) 2D model for PID force control with ($\|\overrightarrow{F_1}\| > \|\overrightarrow{F_2}\|$, $\Delta d_2 < \Delta d_1$)

Fig. 5: (a): 3D trajectory with surface points and normals at each time step. (b): An example 2D model for the force control. $\Delta d_2 < \Delta d_1$ depicts the adjustment based on the force input. $\overrightarrow{V_{N1}}$ and $\overrightarrow{V_{N2}}$ are surface normals from the original surface. $\overrightarrow{F_1}$ and $\overrightarrow{F_2}$ are force measurements.

Furthermore, the position of the inverse kinematic target satisfies the constraint that the US probe aligns with the surface normal and the orientation follows the guided direction, i.e. parallel to the world frame's Y-Z plane. With the updated trajectory, the inverse kinematic solver of Klampt Software Toolbox [19] is used to generate the joint configurations for robot movements.

D. Vessel Localization System

An effective robot control strategy can usually ensure a successful vascular localization by keeping adequate contacted areas and smooth movements. The US frames are consistently collected during robot scanning and the vessel localization pipeline mainly consists of two problems: vessel detection and tracking (Fig. 6). These problems can be modelled as image segmentation tasks with the vessel contours as targets. U-net network architecture has shown great success for image segmentation tasks [11] and is implemented to detect the first vessel candidate from the raw US frames. The output of U-net is a masked image with clustered pixel regions and the centroid of the segment can be computed by averaging the connected pixel coordinates.

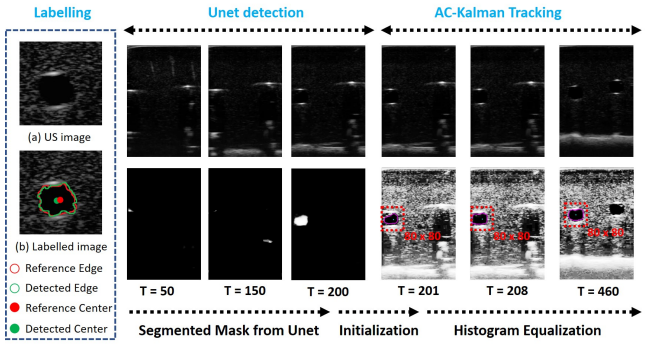


Fig. 6: Vessel detection, tracking and labelling in 2D US frame. The reference center and contour are labeled and compared with the predicted results. T is the time step and six sampled images are selected for visualization.

Given a vessel candidate for tracking initialization, the image is first pre-processed by histogram equalization to increase the contrast of the grayscale intensity inside and outside the vascular structure. An 80×80 region of interest (ROI) is centered around the detected vessel centroid, which is empirically determined in order to cover the region of the vascular contour. After tracking initializes, the next task is to segment the vessel boundary in the ROI via the Active Contour method. Chan-Vese model for active contours is a powerful tool to localize the boundary for the US images that cannot be easily processed by simple threshold-based or gradient-based methods [20]. In the ROI, The Chan-Vese active contour model is able to flexibly identify the vessel boundary for vessel tracking. With the segment pixels, the center of the vessel can be tracked in real-time and a temporal Kalman filter is employed to trace the centers, which performs an accurate state estimation under inaccurate vessel tracking measurements [8], [20]. The Active-Contour-Kalman-Filter framework is referred as "AC-Kalman" in this study. Since the image processing is only analyzed in a local ROI and the change of boundary between the consecutive

frames is small, the number of iteration for the active contour model is set as 5 for real-time application.

III. Experiments and Results

A. Testing of Sensor Calibration

1) Camera Calibration: The evaluation metric of the camera calibration was based on the reprojection errors of optimal transformation. 30 different configurations were defined to collect the images with the RGB-D camera on the robot controlled EE. Average reprojection error was 2.5 mm.

2) Ultrasound Calibration: The optimal transformation between $\{US\}$ and $\{EE\}$ was obtained by US calibration. The ultrasound calibration was evaluated by applying the optimal T_{US}^{EE} to new data points and calculating the reprojection errors based on E_{P2L}, E_{P2Z} and E_{L2L}. The optimal value of E_{P2L} should be zero since all the center points are localized at the center line of the tube. E_{P2Z} and E_{L2L} are compared with the reference values for validation.

To evaluate the US calibration, 12 datasets were generated by moving the calibration stage to different 3D positions and used for calculating the reprojection errors in E_{P2L}, E_{P2Z} and E_{L2L}. These data were collected from each tube by moving the phantom to 6 different positions (different X, Y, Z coordinates), including 10 data points for each tube. The Root-mean-square errors of E_{P2L}, E_{P2Z} and E_{L2L} are reported as 0.61 mm, 0.33 mm and 0.76 mm, with the maximum errors as 1.24 mm, 0.76 mm and 1.20 mm. This shows that the calibration method can find the precise transformation between sensor frames with good accuracy.

B. Vessel Detection and Tracking Experiments

Detection and tracking have the same tasks of tracing the closed vessel boundary and estimating the centers. Based on the testing dataset, the results of the vision pipeline are compared with the labelled reference contours and centers. DICE similarity coefficient [21] and the offset of the center error were used to evaluate the vision performance (Fig. 6). The center error measures the distance between the detected and reference centers, while the DICE evaluates the similarity between two regions formed by the traced vessel boundary and the labelled contour.

Testing data used for vessel detection and tracking were different. For the detection dataset, the robotic system was controlled to scan the arm phantom multiple times to search many areas of the phantom. A small dataset was sufficient for the phantom study since the image features were similar in different frames. Then, 100 images were randomly sampled from the collected data and shuffled, which include zero or more vessels. For the tracking dataset, three specific regions in the phantom, which show only one vessel, were chosen manually and the consecutive US image sequences were utilized for the tracking experiments. For each region, more than 30 images in different time steps were selected from the consecutive image sequences. This ensured all the sampled images were from the same tracking process. A total of 100 images were used for evaluation of the tracking method.

Table. I shows the results of average DICEs and center errors. The success rate is defined by counting the ratio of the correct detection or tracking case with a center error less than 1.0 mm. The image processing time for vessel detection is 12 fps with GPU processing and 32 fps for the vessel tracking with CPU. The tracking speed with 32 fps demonstrates that the vision prototype can perform real-time vessel localization ($32\ fps > 24\ fps$) in a phantom.

TABLE I: Vision testing results with US phantom images

	Detection	Tracking
Average center error (mm)	0.46	0.51
Average DICE coefficient	0.88	0.85
Success Rate	96% (96/100)	98% (98/100)
Image Processing Time (fps)	12	32

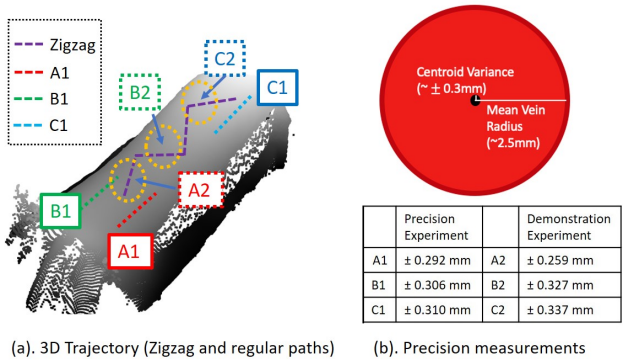
C. Precision Experiment

The goal of the precision experiment was to test the system precision on an arm phantom with the proposed vascular localization system. First, a surface map was captured by a RGB-D camera and utilized to manually determine 3 locations for US scanning, referred as A_1, B_1 and C_1, as shown in Fig. 7 (a). The arm phantom includes vessels with multiple structures, e.g some vessels are fused in a single vessel and others have limited length. Therefore, sections were chosen with only one vessel for the precision experiment. A surface trajectory was generated based on these selections for repeated testing (4 times for each 3 locations shown in Fig. 7 (a)).

For each scanning location, the robot moved along the predefined path and automatically adapted the force on the surface to minimize deformation and maintain adequate contacted area. The vessel centers and contours were tracked simultaneously during the real-time robot movement. This procedure was repeated for 4 trials for each scanning location with 12 cases conducted. The system performance was evaluated by the average variances among all 4 trials centroid locations collected. Fig. 7 (b) illustrates the system precision results on each trajectory. The mean radius of the vessel in the phantom is around 2.5 mm and the variances are ± 0.3 mm.

D. Demonstration Experiment

Demonstration experiment aimed to validate the utility of the integrated system to develop a map of the phantom's vessels and perform a fully automatic vessel localization. Another goal was to show that the proposed PID controller can modulate the force safely on a curved surface. To scan the whole phantom surface, a zigzag trajectory was generated from the surface map with minimal **a priori** user settings, by defining the confined region of the set point at the path. Similar to the precision experiment, this demonstration was conducted with repeated measurements to ensure that the resulting map was repeatable and effective at defining the vessels on the automated trajectory.



	Precision Experiment		Demonstration Experiment
A1	± 0.292 mm	A2	± 0.259 mm
B1	± 0.306 mm	B2	± 0.327 mm
C1	± 0.310 mm	C2	± 0.337 mm

(a). 3D Trajectory (Zigzag and regular paths) (b). Precision measurements

Fig. 7: (a) Visualization of RGB-D point cloud with surface trajectories created by manually selecting 3 regions (A_1, B_1 and C_1) and a zigzag trajectory generated automatically. Specifically, sections A_2, B_2 and C_2 are three different sections defined in the zigzag path. (b) Precision and demonstration variance results shown relative to vein radius in red.

To find the vessel centers for evaluation, US images were collected simultaneously during zigzag scanning. As the trajectory covered various surface regions and it could not guarantee that only one vessel always appear in the US images, three trajectory sections were manually selected so that at least one vessel appears consistently in the US images, referred as A_2, B_2, C_2 (Fig. 7 (a)). The four repeated trials were conducted with the same experimental setting and it was assumed that the small difference between each trial did not change the index of the image in each dataset. Therefore, the sample dataset should be the representation of the same trajectory section for the zigzag path. The centroids of these vessels were compared between demonstrations to ensure repeatability. The results of variance analysis is shown in Fig. 7 (b).

IV. Discussion and Conclusion

The precision experiment and demonstration validate the functionality of the proposed system including the PID controller as well as the vascular localization pipeline. The results of the precision experiment show a variability per location of approximately $\pm 0.3\ mm$. As the pose repeatability of the UR5e robot is 0.03 mm, it is expected that the error in measurements comes primarily from US system and automated vision selection. Error due to the ultrasound probe itself is expected to be approximately 0.25 mm [22]. Target blood vessels for procedures, such as peripheral vessel cannulation, would be around 4 mm. The error of the centroid of $\pm 0.3\ mm$ is acceptable within several margins of error for safety. If the cannulation or other device were part of the robot EE, this error would only be in addition to a needle insertion device. Precision error combined with calibration results would be maximum $\pm 1.8\ mm$ of the complete system. Therefore, these results indicate the system has sufficient precision for 3D reconstruction of critical peripheral vasculature and could be used in conjunction with another intervention that requires this level of accuracy.

The demonstration aimed to mimic *in vivo* requirements with a safety threshold of 5N and target force as 3.5N, phantom curvature like that of a forearm and vessels of similar size (∼5mm). Additionally, the system works autonomously as there is no human intervention from RGBD picture capture until vessel reconstruction. In total, the system demonstrates the capture of the phantom surface to perform a zigzag pattern of searching and reconstruction, at safe force, of multiple vessels in the phantom. This is done in real-time using a commercial robotic arm, linear probe and RGB-D camera. As shown in Fig. 7 (b), the precision of around 0.31 mm of the autonomously found vessels is similar to the manually directed precision experiment.

For the vision testing, the success rates for detection and tracking are 96% (96/100) and 98% (98/100). This demonstrates the proposed vascular pipeline can precisely localize the center and the contour of the phantom vascular structure. In addition, the average center error of 0.46 mm corresponds to a 7.0 pixel distance in the US image. This provides adequate error threshold for the detection task with a 80×80 image ROI. This indicates that if the detected center is located in the ROI, the AC-Kalman method can robustly track the position of the center in real-time. The average center error for tracking (0.51 mm) shows that the deviation contributes to 10% of the vessel dimension (the diameter of the vessel is about 5.0 mm), which is an acceptable range for successful vessel tracking.

Limitations of this study include using a phantom versus a human subject in the capture of vessels. Although the phantom is similar to the human arm in curvature, vasculature size (4mm vs 5mm), and elasticity- visually (in US) vessels are easier to identify in the phantom, versus a human, because humans have more heterogeneous tissue structures. The image features in the phantom US images are very different from the image data collected from a real human arm. The proposed vessel detection and tracking can only confirm the feasibility in phantom data. Additionally, this study assumes that the phantom or subject is static throughout the experiment, relying on the depth image taken at the beginning to represent normals for the probe throughout the procedure. If used for humans, they would have to be secured effectively for similar results.

In summary, this system demonstrates precise reconstruction of small tubes and their centroids deep below the curved surface, like the arm, within safe force limits. The vascular scanning procedure is performed with safety precautions that will be applicable to future human use. Future work could explore the exact error tolerance of the force control system, accuracy of the vision system in human subjects, and integrated interventions built into the EE including IV insertions or central line placement.

Acknowledgment

We thank Professor Kris Hauser for guidance in the initial concepts of the project, Iris Chang for her work in developing the robot end effector and Jing Zhang for his work in helping with the computer vision algorithm development.

REFERENCES

[1] T. Costantino, A. Parikh, W. Satz, and J. Fojtik, "Ultrasonography-guided peripheral intravenous access versus traditional approaches in patients with difficult intravenous access," *Annals of emergency medicine*, vol. 46, pp. 456–61, 12 2005.

[2] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 1, pp. 11–23, 2002.

[3] M. L. Balter, A. I. Chen, T. J. Maguire, and M. L. Yarmush, "The system design and evaluation of a 7-dof image-guided venipuncture robot," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 1044–1053, 2015.

[4] Z. Jiang, M. Grimm, M. Zhou, J. Esteban, W. Simson, G. Zahnd, and N. Navab, "Automatic normal positioning of robotic ultrasound probe based only on confidence map optimization and force measurement," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1342–1349, 2020.

[5] *Universal Robotics E-series User Manual: UR5e*, Universal Robotics, 2018, r 5.0.2.

[6] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle, "An ultrasound robotic system using the commercial robot ur5," *Frontiers in Robotics and AI*, vol. 3, p. 1, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/frobt.2016.00001

[7] T. S. Mathai, L. Jin, V. Gorantla, and J. Galeotti, "Fast vessel segmentation and tracking in ultra high-frequency ultrasound images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 746–754.

[8] J. Guerrero, S. E. Salcudean, J. A. McEwen, B. A. Masri, and S. Nicolaou, "Real-time vessel segmentation and tracking for ultrasound imaging applications," *IEEE transactions on medical imaging*, vol. 26, no. 8, pp. 1079–1090, 2007.

[9] E. Smistad and F. Lindseth, "Real-time automatic artery segmentation, reconstruction and registration for ultrasound-guided regional anaesthesia of the femoral nerve," *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 752–761, 2016.

[10] A. Crimi, M. Makhinya, U. Baumann, C. Thalhammer, G. Szekely, and O. Goksel, "Automatic measurement of venous pressure using b-mode ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 288–299, 2016.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[12] T. S. Mathai, V. Gorantla, and J. Galeotti, "Segmentation of vessels in ultra high frequency ultrasound sequences using contextual memory," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 173–181.

[13] A. I. Chen, M. L. Balter, T. J. Maguire, and M. L. Yarmush, "Deep learning robotic guidance for autonomous vascular access," *Nature Machine Intelligence*, vol. 2, no. 2, pp. 104–115, 2020.

[14] A. Tabb and K. M. A. Yousef, "Solving the robot-world hand-eye (s) calibration problem with iterative methods," *Machine Vision and Applications*, vol. 28, no. 5-6, pp. 569–590, 2017.

[15] R. L. Hirsh, G. N. DeSouza, and A. C. Kak, "An iterative approach to the hand-eye and base-world calibration problem," in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, vol. 3. IEEE, 2001, pp. 2171–2176.

[16] G. Carbajal, A. Lasso, Á. Gómez, and G. Fichtinger, "Improving n-wire phantom-based freehand ultrasound calibration," *International journal of computer assisted radiology and surgery*, vol. 8, no. 6, pp. 1063–1072, 2013.

[17] A. Ahmad, M. C. Cavusoglu, and O. Bebek, "Calibration of 2d ultrasound in 3d space for robotic biopsies," in *2015 International Conference on Advanced Robotics (ICAR)*. IEEE, 2015, pp. 40–46.

[18] A. S. Berahas, R. H. Byrd, and J. Nocedal, "Derivative-free optimization of noisy functions via quasi-newton methods," *SIAM Journal on Optimization*, vol. 29, no. 2, pp. 965–993, 2019.

[19] K. Hauser, "Robust contact generation for robot simulation with unstructured meshes," in *Robotics Research*. Springer, 2016, pp. 357–373.

[20] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[21] A. Crimi, M. Makhinya, U. Baumann, C. Thalhammer, G. Szekely, and O. Goksel, "Automatic measurement of venous pressure using b-mode ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 288–299, 2015.

[22] A. Ng and J. Swanevelder, "Resolution in ultrasound imaging," *Continuing Education in Anaesthesia, Critical Care Pain*, vol. 11, pp. 186–192, 09 2011.