RobUSt–An Autonomous Robotic Ultrasound System for Medical Imaging

FILIP ŠULIGOJ[1], **CHRISTOFF M. HEUNIS**[1], (Student Member, IEEE),
JAKUB SIKORSKI[1], (Member, IEEE), AND **SARTHAK MISRA**[1,2], (Senior Member, IEEE)

[1]Surgical Robotics Laboratory, Department of Biomechanical Engineering, University of Twente, 7500 Enschede, The Netherlands
[2]Surgical Robotics Laboratory, Department of Biomedical Engineering, University Medical Center Groningen, University of Groningen, 9713 Groningen, The Netherlands

Corresponding author: Filip Šuligoj (f.suligoj@utwente.nl)

ABSTRACT Medical ultrasound (US) systems are widely used for the diagnosis of internal tissues. However, there are challenges associated with acquiring and interpreting US images, such as incorrect US probe placement and limited available spatial information. In this study, we expand the capabilities of medical US imaging using a robotic framework with a high level of autonomy. A 3D camera is used to capture the surface of an anthropomorphic phantom as a point cloud, which is then used for path planning and navigation of the US probe. Robotic positioning of the probe is realised using an impedance controller, which maintains stable contact with the surface during US scanning and compensates for uneven and moving surfaces. Robotic US positioning accuracy is measured to be 1.19±0.76mm. The mean force along US probe z–direction is measured to be 6.11±1.18N on static surfaces and 6.63±2.18N on moving surfaces. Overall lowest measured force of 1.58N demonstrates constant probe-to-surface contact during scanning. Acquired US images are used for the 3D reconstruction and multi-modal visualization of the surface and the inner anatomical structures of the phantom. Finally, K-means clustering is used to segment different tissues. Best segmentation accuracy of the jugular vein according to Jaccard similarity coefficient is measured to be 0.89. With such an accuracy, this system could substantially improve autonomous US acquisition and enhance the diagnostic confidence of clinicians.

INDEX TERMS Biomedical imaging, medical robots, object segmentation, path planning, robot control.

I. INTRODUCTION

Robots are gradually transforming medical procedures by facilitating the earlier diagnosis of diseases, improved operational efficiency, and implementation of advanced surgical techniques [1]. Medical imaging modalities are used to acquire data and create a visual representation of the human anatomy, that can be utilized for diagnostics or during surgical procedures [2]. Specifically, medical ultrasound (US) is widely used for the diagnosis of internal tissues due to its beneficial features, including low cost of operation, non-invasive nature, and real-time imaging. However, there are challenges associated with acquiring and interpreting US images. The acquisition of US images requires correct placement and consistent force between the US probe and the

patient. Variables such-as probe-to-surface orientation, probe position, and tissue contact force affect the quality of the US images [3]. Once the US images are acquired, an analysis follows, whereby the clinician localizes and inspects the tissues. In order to enhance this process, image processing and analysis techniques are used for extraction of additional diagnostic information from the US images [4].

Volumetric data is useful to the clinicians to visualize and inspect internal tissues. Such data is generated from 2D and 3D US images using reconstruction techniques [5] and knowledge of the US probe pose. The pose relation between the US probe and the frame of reference is obtained using electromagnetic [6], visual [7] or mechanical [8] spatial sensing devices. Based on that notion a robotic system can be used to manipulate the US probe and obtain accurate spatial data. For example, Zhang *et al.* demonstrated the application of a 6-DOF robot with an in-hand 2D US probe that facilitates 3D

US volume reconstruction [9]. Aalamifar *et al.* introduced a calibrated robot-assisted US system, which employs visual tracking of movements to follow and align to a manually operated probe [10]. US probe calibration has been investigated by incorporating automatic calibration using a 6 degrees-of-freedom (DOF) robot, a marker-based optical tracking system, and a US calibration phantom [11], [12]. Employing a robot is also useful to provide contact force information and control the end-effector position in relation with the scanned surface [13]. As shown by Jiang *et al.*, a combination of US images and force feedback from a 7-DOF robot can be used to control the alignment of the US probe in reference to the scanning surface [14]. Furthermore, Jiang *et al.* demonstrated the importance of the US probe-to-surface alignment when acquiring images of human tissue [15].

Despite the aforementioned large body of work in the subject, the planning and execution of trajectories ensuring stable robot contact with an anthropomorphic body remains a vital research problem. This problem is catalyzed by high anatomical variations in human population, including uneven geometries of skin surfaces and varying skin elasticity. Robotic path planning is intensively researched in the fields of mobile robotics [16] and non-destructive testing [17]. Translation of these techniques to medical applications involving robotic US and the use of point clouds acquired from the 3D camera for path planning of a probe in contact with the body surface is studied in [18]–[20]. For example, Graumann *et al.* gave only a rudimentary account of using impedance control for stable contact. Next, Huang *et al.* proposed the planning of one line surface path and used two single-axis force sensors for correcting US probe poses. Finally, Virga *et al.* studied autonomous robotic US acquisition to automate 3D camera-based single line trajectory planning, and reported on force-based position compensation during breathing motions. Nevertheless, planning of paths that contain multiple scanning sweeps on curved surfaces and motion control necessary to execute these trajectories while compensating for both the position and the orientation of the US probe in presence of physiological movement, was not demonstrated. To address this issue, we propose a novel approach, whereby the planned path is transformed into a series of equilibrium poses for the impedance controller. With this approach we maintain constant contact with both deformable and moving surfaces during US scanning.

US image analysis is a well-researched topic within the context of manual US equipment without robotic assistance [21]. While conventional US image analysis involves a clinician and requires experience, automatic segmentation methods can improve accuracy and clinical outcome [22] or enable real-time systems to control medical instruments like needles [23]–[25], catheters [26], or drills [27]. However, US images inherently contain many imaging artifacts and have low signal-to-noise ratio so filtering of noise from the signal is one of the main perceived challenges [28].

Furthermore, due to the aforementioned anatomical variances and US machine characteristics, the intensity, density,
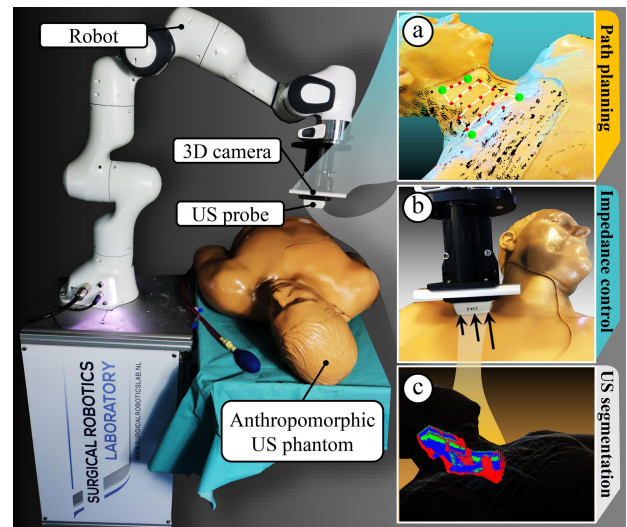

FIGURE 1. RobUSt - "Robotic Ultrasound System" (a) A 3D camera is used to obtain a point cloud of an anthropomorphic ultrasound (US) phantom surface. This point cloud is used for planning of an US probe scanning path. (b) US images are acquired using a probe that is manipulated, along the planned path, with the impedance controller that ensures constant surface contact. (c) A US volume is reconstructed and unsupervised learning K-means clustering method is used for the segmentation of scanned tissues and multi-modal visualization.

and speckle content of tissues can be perceived differently in US images. This occurrence requires the use of machine learning methods, which segment such tissues automatically [29], [30]. In this study we demonstrate that the robotic-aided reconstruction of US volumes and the utilization of unsupervised learning method K-means enable the robust and automatic segmentation of sub-dermal tissues. The resulting accurate volumetric data of segmented tissue can be used for diagnostics, computer-assisted surgery, and can be readily integrated with other automated interventional systems.

This study presents a novel integrated medical imaging solution, combining the 3D US volume reconstruction and image analysis with autonomous path planning and robot control. RobUSt is a "**Rob**otic **U**ltrasound **S**ystem" that performs autonomous path planning and implements impedance control during US scanning (Fig. 1). Such an approach facilitates the capture of US images while compensating for tissue deformation, body motions, and positioning errors. Our system improves on diagnostic capabilities by performing 3D US reconstruction, tissue segmentation, and registration of the multi-modal data. To the best of our knowledge, this study on robotic US is the first to combine path-to-motion robot control and the generation of multi-modal data. The paper is outlined as follows: In Section II, we present both the hardware of RobUSt as well as its key software components, comprising of novel algorithms for autonomous path planning, robotic impedance control, 3D US reconstruction, and tissue segmentation. The proposed solutions are validated through a series of experiments in Section III. Conclusions and directions for future studies are provided in Section IV.
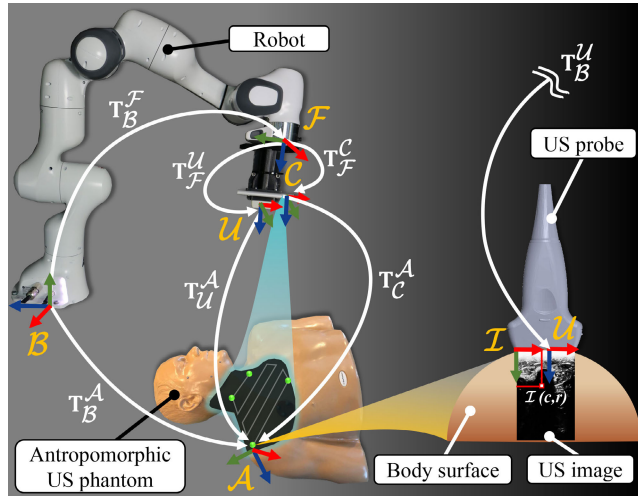
FIGURE 2. System overview featuring major components and reference frames used in this study: The robot base ($\mathcal{B}$), robot flange ($\mathcal{F}$), camera ($\mathcal{C}$), ultrasound (US) probe ($\mathcal{U}$) and US image frames ($\mathcal{I}$). Homogeneous transformation between any two frames ($\mathcal{Y}$) and ($\mathcal{X}$) is denoted as $\mathbf{T}_{\mathcal{Y}}^{\mathcal{X}} \in \mathrm{SE}(3)$.

II. THE RobUSt SYSTEM

The main components of the RobUSt system are as follows: a 7-DoF robot (Panda, Franka Emika GmbH) with integrated torque sensors (force resolution < 0.05N and relative force accuracy 0.8N), a diagnostic medical US machine (Sonix-Touch Q+, BK Medical), and a 3D camera (Intel Realsense SR305). The end-effector of the robot is used to manipulate the attached 3D camera and the US probe. The 3D camera uses a projected light pattern to capture a 3D-RGB point cloud of the body surface (resolution of 640×480 pixels), which is used for planning of the US scanning path. During 2D B-mode image acquisition the robot moves the US probe along the planned path while keeping contact with the surface using the impedance control.

US images are used for 3D volume reconstruction and segmentation by employing Point Cloud Library (PCL v1.9.1) for the handling of point clouds, and Open Source Computer Vision Library (OpenCV v4.4.0) for image processing. Algorithms for robot control and machine vision are integrated into a C++ application on a portable workstation (Intel Core i7-9750H, 2.60Ghz, and 32 GB RAM) running Ubuntu 18.04 with Kernel 5.4, and RT-PREEMPT for real-time robot control capabilities. Experiments are performed using an anthropomorphic US mannequin (Gen II US Central Line Training Model, CAE Healthcare, USA), which we will refer to as anthropomorphic US phantom.

A. SYSTEM CALIBRATION

The robot base frame ($\mathcal{B}$) is chosen to serve as a global reference for position control of the robot (Fig. 2). Homogeneous transformation matrix between any two coordinate frames ($\mathcal{Y}$ and $\mathcal{X}$) is denoted as $\mathbf{T}_{\mathcal{Y}}^{\mathcal{X}} \in \mathrm{SE}(3)$. It is composed of the rotation matrix $\left(\mathbf{R}_{\mathcal{Y}}^{\mathcal{X}} \in \mathrm{SO}(3)\right)$ that describes the

orientation, and the position vector $\left(\mathbf{p}_{\mathcal{Y}}^{\mathcal{X}} \in \mathbb{R}^3\right)$ of frame ($\mathcal{X}$) with respect to the reference frame ($\mathcal{Y}$). Robot flange pose is the transformation $\left(\mathbf{T}_{\mathcal{B}}^{\mathcal{F}}\right)$. 3D camera and the US probe are rigidly fixed to the robot flange frame ($\mathcal{F}$), making $\mathbf{T}_{\mathcal{F}}^{\mathcal{U}}$ and $\mathbf{T}_{\mathcal{F}}^{\mathcal{C}}$ constant. $\mathbf{T}_{\mathcal{F}}^{\mathcal{C}}$ is estimated using calibration method implemented in Visual Servoing Platform (ViSP v3.3.0) [31]. The least square error solution for $\mathbf{p}_{\mathcal{F}}^{\mathcal{U}}$ is estimated by retrieving $\left(\mathbf{T}_{\mathcal{B}}^{\mathcal{F}}\right)$ while moving the robot in several configurations while keeping constant position of the US probe. Positioning of the 3d camera and the US probe on the robot end-effector ensures that $\mathbf{R}_{\mathcal{F}}^{\mathcal{U}} = \mathbf{R}_{\mathcal{F}}^{\mathcal{C}}$.

Transformation of frame ($\mathcal{A}$) which is originally observed in camera frame ($\mathcal{C}$) to robot base frame ($\mathcal{B}$) is valid under the assumption that $\mathbf{T}_{\mathcal{B}}^{\mathcal{F}}$ is retrieved at the same time as $\mathbf{T}_{\mathcal{C}}^{\mathcal{A}}$:

$$\mathbf{T}_{\mathcal{B}}^{\mathcal{A}} = \mathbf{T}_{\mathcal{B}}^{\mathcal{F}} \mathbf{T}_{\mathcal{F}}^{\mathcal{C}} \mathbf{T}_{\mathcal{C}}^{\mathcal{A}}. \tag{1}$$

During US scanning, the probe is in contact with the surface and captures B-mode US images. Frame ($\mathcal{I}$) is assigned to the US image matrix and any pixel is transformed to the position coordinates in frame ($\mathcal{B}$) using the equation:

$$\begin{bmatrix} \mathbf{p}_{\mathcal{B}}^{\mathcal{I}(\mathbf{c},\mathbf{r})} \\ 1 \end{bmatrix} = \mathbf{T}_{\mathcal{B}}^{\mathcal{F}} \mathbf{T}_{\mathcal{F}}^{\mathcal{U}} [\, cw - \frac{v}{2} \quad 0 \quad rh \quad 1 \,]^T, \tag{2}$$

where $c \in N$ and $r \in N$ are the c-th column and r-th row in the US image matrix, v is the width of the probe and h and w are the height and width of any pixel, respectively.

B. PATH PLANNING

The goal of the path planning algorithm is to automatically generate a sequence of poses on the body surface that are used to image and reconstruct the underlying volume beneath this area. These poses are based on user-defined outer point limits of the scanning area. The developed path planning algorithm is described in Fig. 3. First, the 3D camera is used to capture 3D-RGB point cloud of the anthropomorphic US phantom surface. The point cloud ($\{\kappa\}$) is used in the algorithm that autonomously plans a path that the robot carrying the US probe needs to follow when preforming US scanning. Only positional data (without the colors) are used to define the point cloud ($\{\kappa\}$) in the ($\mathcal{C}$) frame, comprised of f number of points $\kappa_j \in \mathbb{R}^3$ with $j = 1, \ldots, f$. Outer limits of the scanning area are defined by a set of points ($\{\iota\}$), where $\iota_k \in \mathbb{R}^3$ and $k = 1, \ldots, 4$. The number of the generated grid points (s) and the spacing between them depend on the distance between the limit points and the user set variable (g) which is based on US probe size. A set of grid points ($\{\sigma\}$), shown in Fig. 3(a), is generated such that minimal number of points (σ_i) with $i = 1, \ldots, s$ are linearly spaced, and g is the largest allowed distance between any two points. All the grid points (σ_i) are spaced in straight 3D lines between the points of the set ($\{\iota\}$). Thus, they need to be projected to the actual surface to compensate for curvatures. Each grid point (σ_i) is projected along the direction of normal ($\mathbf{n}$) to the least square fitted grid plane. The nearest κ_j found on the surface of the body to the
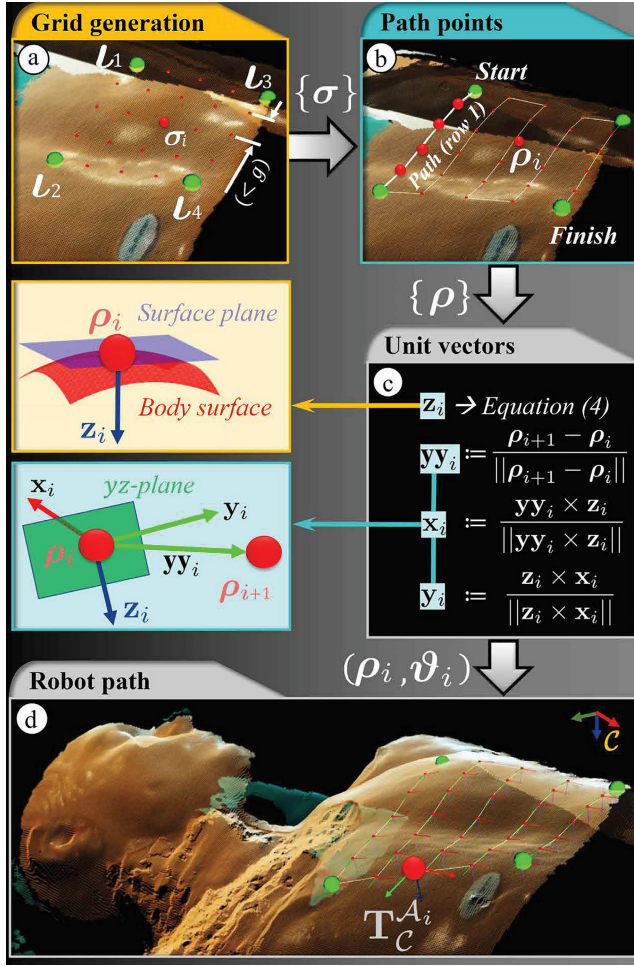
FIGURE 3. Path planning algorithm: (a) A set of grid points ($\{\sigma\}$) (denoted red) is generated between a set of four limit points ($\{\iota\}$) (denoted green). (b) A surface path point (ρ_i) (red points), with $i = 1, \ldots, s$ and s equal to number of grid points (σ_i) and point cloud set ($\{\kappa\}$) using (3). The white lines connecting consecutive path points denote the pass schedule. (c) The orientation (ϑ_i) of the ultrasound (US) probe in a path point (ρ_i) is defined using the z-axis unit vector as a normal to local surface according to (4) and a vector (yy_i) that connects consecutive path points (ρ_i) and (ρ_{i+1}) and lies in the yz-plane. It should be noted that for the last path point in every row, yy_i from a previous point is used. When the probe is returning on even row sweeps, $yy_i := -yy_i$, to avoid the turning of the US probe. Unit vectors (x_i) and (y_i) for orientation (ϑ_i) are derived from z_i and yy_i. (d) The robot path in frame ($\mathcal{C}$) is a set of poses ($\{T_\mathcal{C}^{\mathcal{A}_1}, \ldots, T_\mathcal{C}^{\mathcal{A}_s}\}$), where $T_\mathcal{C}^{\mathcal{A}_i}$ is based on position (ρ_i) and orientation (ϑ_i).

ray ($\overrightarrow{\sigma_i n}$) then becomes a surface path point (ρ_i):

$$\rho_i = \underset{\kappa_j \in \{\kappa\}}{argmin}(||\overrightarrow{\sigma_i \kappa_j} \times n||). \tag{3}$$

Orientation of the US probe (ϑ_i) is calculated for path point (ρ_i) using surface points around it and the position of the next path point (ρ_{i+1}). The direction of z-axis unit vector for every path point (ρ_i) is calculated based on the local neighborhood of points in a radius of size (g) around it:

$$z_i := normal\{\forall t | t \in \{\kappa\} \wedge ||t - \rho_i|| < g\}, \tag{4}$$

where t is any point in point cloud ($\{\kappa\}$). Therefore, value of variable (g), the maximal distance between neighboring path points, is used for calculation of local surface normal. Normal to surface for a set of points is estimated using the PCL function, denoted as *normal*{} in (4), that solves the eigenvalues and eigenvectors of a given covariance matrix and estimates the least-squares plane normal. Unit vectors (x_i) and (y_i) are calculated in Fig. 3(c). Path point position (ρ_i) and orientation (ϑ_i) together define a transformation ($T_\mathcal{C}^{\mathcal{A}_i}$) that is transformed to frame ($\mathcal{B}$) using (1). A complete set of robot end-effector path poses in frame ($\mathcal{B}$) is denoted as $\Gamma = \{T_\mathcal{B}^{\mathcal{A}_1}, \ldots, T_\mathcal{B}^{\mathcal{A}_s}\}$.

C. IMPEDANCE CONTROL FOR PROBE-TO-SURFACE CONTACT

The autonomous positioning of the RobUSt probe in contact with the body surface necessitates the use of a control algorithm, which adheres to a number of requirements. First of all, the control of the probe must involve contact forces which are deemed safe. Second, at all times, the distal surface of the probe must remain completely in contact with the surface of the body for imaging. Finally, the positioning accuracy on a plane tangent to the surface of the body must be sufficient to ensure that the entire path (Γ) is mapped. We fulfil these requirements by using a two-stage algorithm (Fig. 4) comprising of a motion generator and a closed-loop impedance controller. As an input, the algorithm takes the Cartesian position vector ($p_\mathcal{B}^{\mathcal{A}_i}$) and orientation quaternion ($q_\mathcal{B}^{\mathcal{A}_i} \in \mathbb{H}$) of the desired robot pose ($T_\mathcal{C}^{\mathcal{A}_i}$). Operating continuously, the algorithm generates reference joint torques ($\tau^r \in \mathbb{R}^7$) fed as an input to the low-level built-in controller of the robot.

The motion generator contains an internal model of the Panda robot, described by joint angles ($\theta^m \in \mathbb{R}^7$). The generator uses an inverse kinematics scheme, based on the damped Moore-Penrose pseudoinverse (†) [32]

$$J_\mathcal{B}^\dagger = J_\mathcal{B}^T \left(J_\mathcal{B} J_\mathcal{B}^T + \psi I\right)^{-1} \tag{5}$$

of the space robot Jacobian ($J_\mathcal{B} \in \mathbb{R}^{6\times7}$), where $\psi \in \mathbb{R}_{>\epsilon_0}$ is the damping coefficient and $I \in \mathbb{R}^{6\times6}$ is an identity matrix. The generator provides a continuous motion trajectory comprising of intermediate US reference (ref) poses, described by the position vector ($p_\mathcal{B}^{ref} \in \mathbb{R}^3$) and orientation quaternion ($q_\mathcal{B}^{ref} \in \mathbb{H}$), between the discrete poses of the set ($\{\Gamma\}$), thereby smoothing the input to the controller.

The impedance control (Fig. 4) takes place in the reference frame ($\mathcal{U}$) of the US probe. We use the control law

$$\tau_r = J_\mathcal{U}^T \left(K_c e_\mathcal{U}^c + K_d \begin{bmatrix} \dot{p}_\mathcal{U} \\ \omega_\mathcal{U} \end{bmatrix}\right) + \tau_c, \tag{6}$$

where $e_\mathcal{U}^c \in \mathbb{R}^6$ is the error between the reference and current configuration of the robot expressed in frame ($\mathcal{U}$), $\dot{p}_\mathcal{U} \in \mathbb{R}^3$ and $\omega_\mathcal{U} \in \mathbb{R}^3$ are linear and rotational velocities of $T_\mathcal{B}^\mathcal{U}$ expressed in frame ($\mathcal{U}$) [33], and $\tau_c \in \mathbb{R}^7$ are torques compensating for robot dynamics. These torques are computed
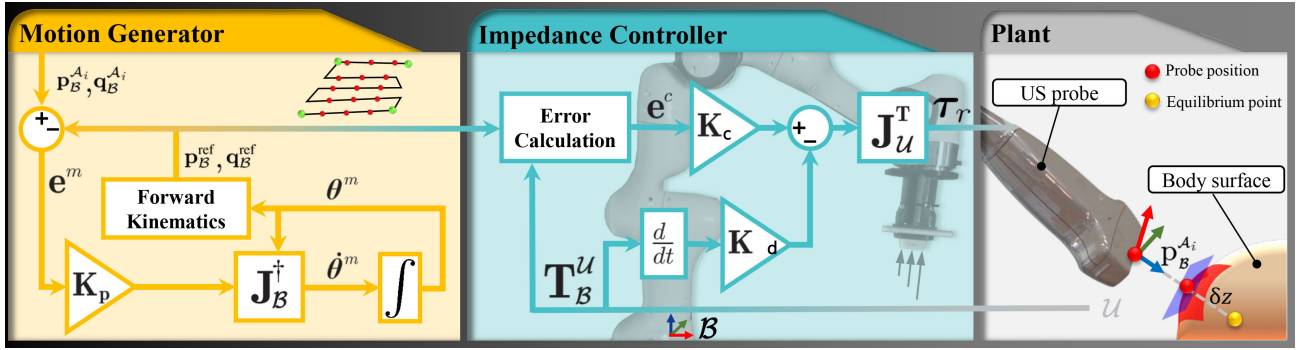
FIGURE 4. The control algorithm is used to steer the ultrasound (US) probe along discrete consecutive poses defined by frames $(\mathcal{A}_i)$. Each pose is converted to a Cartesian position vector $\left(\mathbf{p}_{\mathcal{B}}^{\mathcal{A}_i}\right)$ and orientation quaternion $\left(\mathbf{q}_{\mathcal{B}}^{\mathcal{A}_i}\right)$. A smooth reference trajectory (ref), described by the position $\left(\mathbf{p}_{\mathcal{B}}^{\text{ref}}\right)$ and orientation quaterion $\left(\mathbf{q}_{\mathcal{B}}^{\text{ref}}\right)$ is generated with an internal kinematic model, described by joint angles (θ^m). This model is used in an inverse-kinematics motion generator based on damped pseudo-inverse of the robot space Jacobian $(\mathbf{J}_{\mathcal{B}})$. The generator minimizes the error $(\mathbf{e}^m)$ between the end-effector pose of the model and the path pose. The smooth trajectory is then used as an equilibrium reference in an impedance controller. The control input comprises of reference torques (τ_r) for the low-level torque controllers of individual robot joints. The input is based on the actual displacement of the US probe from the equilibrium, simulating a spring-dampener system with parameters defined by stiffness matrix $(\mathbf{K}_c)$ and damping $(\mathbf{K}_d)$.

using a dynamic model of the Panda robot provided by the manufacturer. We use the gain matrices $\mathbf{K}_c \in \mathbb{R}_{\geq 0}^{6 \times 6}$ and $\mathbf{K}_d \in \mathbb{R}_{\geq 0}^{6 \times 6}$ to define the stiffness and the damping of the force-position relation. The stiffness is defined as follows:

$$\mathbf{K}_c = diag\left(\begin{bmatrix} K_x & K_y & K_z & K_i & K_j & K_k \end{bmatrix}\right), \quad (7)$$

while

$$\mathbf{K}_d = 2\,\mathbf{K}_c^{\circ(1/2)}, \quad (8)$$

such that the system is critically damped.

The force-position relation provided by impedance control enables the robot compliance in the presence of contact force. By introducing a constant offset $(\delta z \in \mathbb{R}_+)$ to the instantaneous position reference $\left(\mathbf{p}_{\mathcal{U}}^{\text{ref}}\right)$ along the axis $(\widehat{\mathbf{z}}_{\mathcal{U}})$, we use this phenomenon to achieve a sustained probe-to-surface contact with the resulting normal force $\left(F_{\mathcal{U}}^z \in \mathbb{R}^+\right)$. The steady-state components of that force can be modelled as follows:

$$F_{\mathcal{U}}^z = K_z(\delta z - e_z^r), \quad (9)$$

where $\left(e_z^r \in \mathbb{R}\right)$ represents the error between the location surface as registered using the point-cloud and the actual one. As long as $\delta z > e_z^r$, the probe will remain in contact with the body. By tuning the coefficient (K_z), the range of contact forces can be selected. Additionally, we choose K_j such that the resulting torsional compliance compensates for uneven contact of the probe surface with the body [18]. The rest of the coefficients of $\mathbf{K}_c$ are $K_x, K_y \gg K_z$, and $K_i, K_k \gg K_j$ to achieve high stiffness, which improves the precision of the motion. The values for all coefficients are specified in Section III.

D. ULTRASOUND VOLUME RECONSTRUCTION AND SEGMENTATION

The 3D camera and impedance control are used to realise the autonomous US surface scanning for volume reconstruction
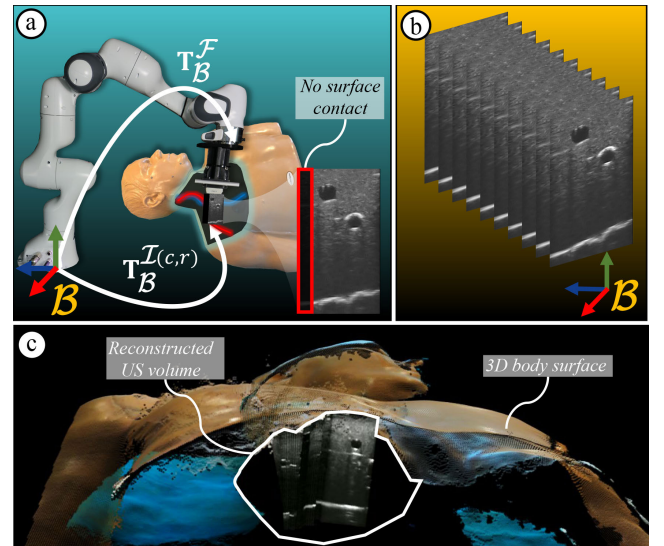

FIGURE 5. Ultrasound (US) volume reconstruction and multi-modal registration: (a) Acquisition of US images with no probe-to-surface contact on the left side of the image; (b) US volume reconstruction based on 2D US images and corresponding robot poses in base frame $(\mathcal{B})$; (c) Visualization of the registered multi-modal data set composed of 3D-RGB surface point cloud and the reconstructed 3D-intensity US point cloud.

and segmentation. During US scanning 2D-intensity images are retrieved and inspected to detect potential loss of contact between the US probe and the body. In case of uncertainties (e.g., variation in tissue stiffness or body surface motion), part of the probe can lose contact with the surface, resulting in zero value pixels in columns of the US image (Fig. 5(a)). The intensity check algorithm is developed to detect such columns and prevent the conversion of their pixels to 3D points in the reconstructed US volume. Redundant control using contact force and US image inspection is used to detect unexpected situations during US scanning.
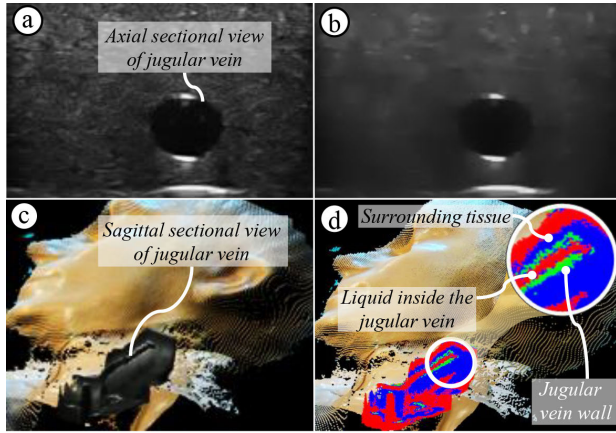
FIGURE 6. Volumetric segmentation: (a) A 2D Ultrasound (US) image of the jugular vein; (b) A 2D US image of the jugular vein after processing with an Edge-Preserving-Filter; (c) The reconstructed volume with a cross-section of the jugular vein; (d) The K-means segmentation results showing the jugular vein wall as green points, liquid within the jugular vein as red points, and the surrounding tissue as blue points.

FIGURE 7. Positioning accuracy measurements are performed using a series of 5mm diameter cylindrical shafts fitted inside the phantom filled with a gelatin-water mixture. Fabricated dimensions between the shafts are considered as the ground truth since their accuracy is at least an order of the magnitude higher than measured accuracy. Shafts localized in US images are transformed to the robot base frame ($\mathcal{B}$) using the corresponding robot poses and (2).

The acquired 2D US image and corresponding transformation $\left(\mathbf{T}_{\mathcal{B}}^{\mathcal{F}}\right)$ of the robot flange are used with (2) to map all the valid pixels to the 3D points that represent the reconstructed US volume (Fig. 5(b)). Reconstructed US volume and the surface previously captured with the 3D camera are transformed to the robot base coordinate system using (1) and (2). The result is a multi-modal data set registered in frame ($\mathcal{B}$), as shown in Fig. 5(c). The advantage of this data set over a standard US data set is that visualisation of inner tissues is shown relative to the surface of the body.

Segmentation is used to visualize different tissues within the reconstructed 3D US volume. The segmentation process starts with speckle reduction and intensity homogenization in 2D US images using the Edge-Preserving-Filter (EPF) implemented in OpenCV (v4.4.0). EPF avoids edge shifting which characterizes simpler methods, such as Mean or Gaussian filters, and has a lower computational cost in comparison to other edge-preserving algorithms, like bilateral texture filter and anisotropic diffusion [34]. Before EPF filtering, the image is transformed from the grey-scale to RGB color representation. The filtering results are shown in Fig. 6(b). Following the image processing, the pixels are transformed to 3D space using (2). Due to anatomical variability and changes in the US machine settings, identical tissues can be observed differently in US images with regards to intensity, density, and speckle content. K-means clustering method [35], implemented in the PCL (v1.9.1), is used to separate differently observed intensities of tissues in the reconstructed US volume. K-means clustering is an unsupervised learning method that identifies a set number of centroids in any given data set and then appoints every data point to the nearest cluster by reducing the in-cluster sum of squares. K-means is chosen for the segmentation task because it requires no prior knowledge and performs well in cases with large variability in the input data. In our implementation, the complete
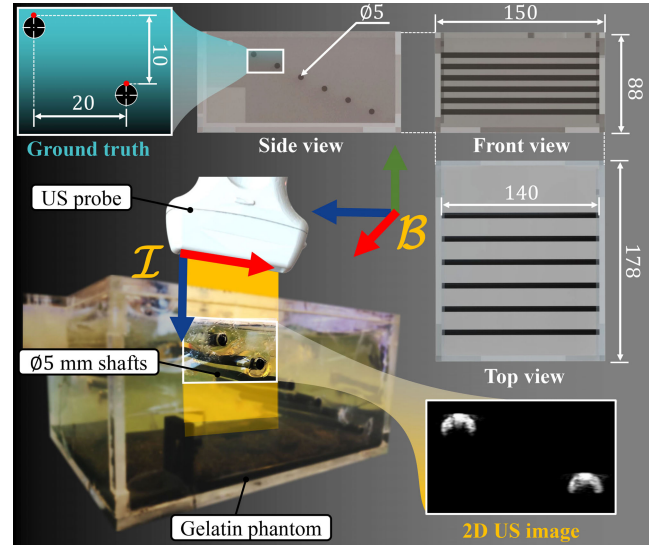
3D-RGB type point cloud is used as an input for K-means algorithm with three clusters. This results in clustering that is predominantly influenced by color values while the impact of point position grows with the distance between clustered points. Segmentation of the neck area is shown in Fig. 6(d).

III. EXPERIMENTAL RESULTS

We validate the RobUSt system in a series of studies. First, we assess the performance of the robotic system, by measuring the positioning accuracy and the contact forces during US scanning. Secondly, the segmented US volumes are evaluated based on robustness to changes in US system parameters. The protocol for all the experiments involves aligning the robot end-effector with the attached 3D camera towards the surface of interest and capturing the point cloud. The outer limits of the scanning area are then selected interactively on the visualization window. Based on those inputs the algorithm autonomously plans the motion trajectory. The robot moves the US probe with the motion controller while simultaneously capturing the US images, end-effector poses and forces.

The positioning accuracy of the robot with the US probe is defined as the ability of the robot to localize 3D points in the gelatine phantom (Fig. 7). The positioning error between localised points is a cumulative error composed of robot positioning errors, robot-to-US probe calibration errors, US image reconstruction parameters, and in-image localization errors. Based on the localization of 20 points uniformly distributed inside the phantom, the mean Euclidean error is 1.19±0.76mm. However, authors expect larger positioning errors when scanning uneven surfaces, due to changes in orientation of the robot end-effector [36].
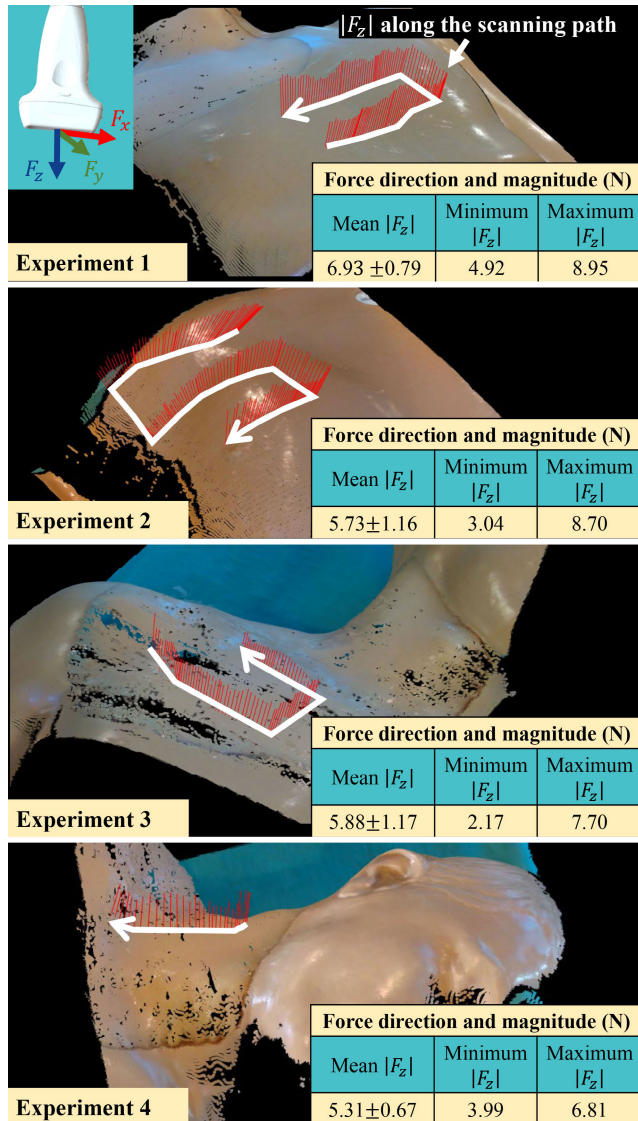
FIGURE 8. Force feedback measurements from the robot (using built-in torque sensors) are recorded at 20Hz during ultrasound (US) scanning on a static anthropomorphic US phantom. The parameters of the robot compliance matrix are set to $K_c = diag([3000\ 3000\ 200\ 150\ 10\ 150])$. Maximum distance between grid points (σ_i) is set to $g = 30mm$ and each equilibrium point is translated along the z-axis of orientation (ϑ_i) from every path point position (ρ_i) for $\delta z = 40mm$. The reference pose changes from frame ($\mathcal{A}_i$) to frame ($\mathcal{A}_{i+1}$) once $\sqrt{(x_{\mathcal{U}}^{\mathcal{A}_i})^2 + (y_{\mathcal{U}}^{\mathcal{A}_i})^2} < 3mm$.

The probe-to-surface forces are measured during scanning of four surfaces with different geometry on the static anthropomorphic US phantom. Measured forces relative to z-axis of frame ($\mathcal{U}$) are reported in Fig. 8. The experiment demonstrated the ability of the RobUSt to keep stable contact with surfaces of different geometries. The mean normal ($|F_z|$) forces exerted on the surface are between 5.31N and 6.93N, which falls within the range advised for acquiring clear US images [13], [37].

Next, the experiment with the moving surface is performed. The probe position in frame ($\mathcal{B}$) and the force in frame ($\mathcal{U}$) are measured during path execution while the
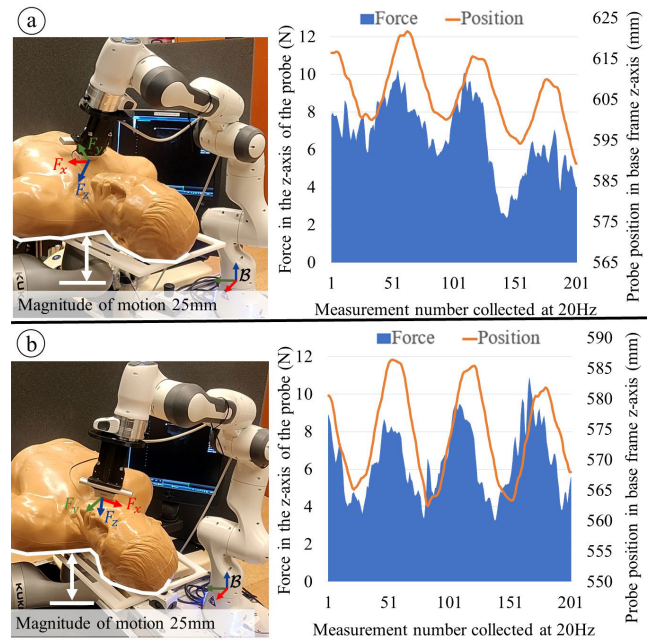

FIGURE 9. Force and position feedback measurements from the robot (using built-in torque sensors) during ultrasound (US) scanning on a moving anthropomorphic US phantom. The parameters of the robot compliance matrix are set to $K_c = diag([3000\ 3000\ 100\ 150\ 150\ 150])$. Each equilibrium point is translated along the z-axis of orientation (ϑ_i) from every path point position (ρ_i) for $\delta z = 80mm$. (a) Scanning path located in the chest area; (b) Scanning path located in the neck area. Please refer to accompanying video for a demonstration of these experiments.

anthropomorphic US phantom is moved using the second robot with cycle time of 3s and vertical position amplitude of 25mm (Fig. 9). The experiment is performed 4 times on the chest and 2 times on the neck area, each with a duration of 22s. The mean normal ($|F_z|$) forces exerted on the surface are 6.63±2.18N, with overall lowest and highest force measured to be 1.58N and 11.86N, respectively. This experiment demonstrates that probe-to-surface contact is consistent on a moving body.

Robustness of the segmentation algorithm is validated by changing US transducer frequency (f), dynamic range (DR, ratio of the largest to the smallest signal), and US image gain (K_u, the overall brightness) during three separate scanning experiments on the collum (neck) area of the static anthropomorphic US phantom as shown in Fig. 10 (1a, 2a, and 3a). Changing the values of US settings results in capturing US images that differ with regard to intensity, speckle content, and the level of tissue homogeneity. During robotic US scanning, images of size 256 × 460 pixels (38 × 70mm) are retrieved and filtered with EPF. The parameters of the EPF algorithm are determined heuristically with the size of the neighborhood being set to 60 and the coefficient of the edge-preserving threshold set to 0.15. The reconstructed US volume is segmented using the algorithm described in Section II.D. A qualitative segmentation analysis is performed based on the visual examination of the featured segmentation results shown in column (b) of Fig. 10.
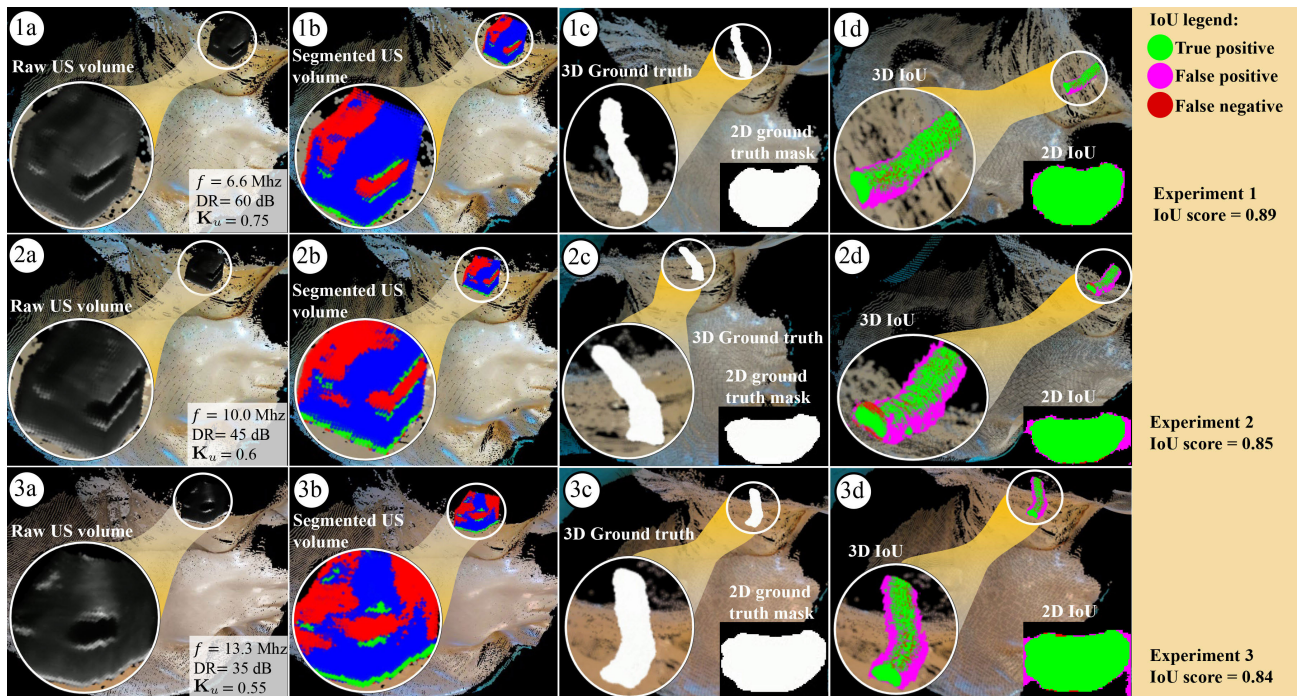
FIGURE 10. Evaluation of segmentation results using Intersection over Union (IoU) on the static anthropomorphic phantom: (1a) First reconstructed Ultrasound (US) volume; (1b) Segmentation of the first US volume; (1c) Ground truth of the jugular vein; (1d) First volume IoU; (2a) Second reconstructed US volume; (2b) Segmentation of the second US volume; (2c) Ground truth of the jugular vein; (2d) Second volume IoU; (3a) Third reconstructed US volume; (3b) Segmentation of third US volume; (3c) Ground truth of the jugular vein; (3d) Third volume IoU. Please refer to accompanying video for a demonstration of these experiments.

We conclude that, even with different US system parameters, the algorithm can differentiate between the jugular vein wall (green points), the liquid within the jugular vein (red points), and the tissue surrounding the vein (blue points). For the quantitative evaluation–the ground truth mask of the jugular vein is manually denoted in 2D images and the 3D ground truth volume is reconstructed. The pixels in the rectangular area around every 2D ground truth mask are compared to their corresponding segmented 3D points. As a measure of similarity between the masks, the Jaccard similarity coefficient, i.e., the Intersection over Union (IoU) is calculated in column (d) of Fig. 10. In conclusion, the unsupervised learning algorithm demonstrated that it can robustly segment a variety of tissues with significant accuracy, with no prior training data.

IV. CONCLUSION AND FUTURE WORK

In this study, we present a framework for RobUSt: a robotic US system with high level of autonomy. RobUSt is capable of performing autonomous path planning that accounts for anthropomorphic surface geometry, by utilizing the 3D camera data. The positioning accuracy measured using the gelatin phantom is 1.19±0.76mm. The impedance control enables the positioning of the US probe while compensating for tissue deformation or surface movement. This is shown in by measuring probe-to-surface force during US scanning of a static body and moving body with 6.17±1.19N and 6.63±2.18N, respectively. Furthermore, overall lowest measured force of 1.58N demonstrates constant probe-to-surface contact is kept. The acquired US images and corresponding probe poses are used for the 3D US reconstruction and registration of multi-modal images consisting of the US volume and the body surface point cloud. The resulting data is used for the segmentation of the jugular vein in three scenarios with varying US machine settings. Ground truth and segmentation volumes of the jugular vein are used to the calculate the best case IoU to be 0.89. Using K-means clustering, we can reliably differentiate between the inside of the vein, vein walls, and the surrounding tissue. In conclusion, with added task-specific functionalities, the RobUSt system framework could be utilized for autonomous medical examinations, clinical diagnosis and automated surgical interventions.

One of the challenges observed when registering the surface point cloud to the US volume is the decreased accuracy between the two due to tissue deformations caused by contact with the US probe. To account for this in our future work, we plan to acquire 3D point-clouds of the probe-surface contact area in real-time. Based on the obtained US images and measured forces, a model can be created that describes and predicts the internal tissue deformations. This approach will enable estimation and reconstruction of the undeformed US volume and improve the registration accuracy with the surface point cloud.

Furthermore, we will utilize the RobUSt framework in specific applications e.g., endovascular interventions in clinically relevant scenarios.

REFERENCES

[1] J. Troccaz, G. Dagnino, and G.-Z. Yang, "Frontiers of medical robotics: From concept to systems to clinical translation," *Annu. Rev. Biomed. Eng.*, vol. 21, no. 1, pp. 193–218, Jun. 2019.

[2] K. Cleary and T. M. Peters, "Image-guided interventions: Technology review and clinical applications," *Annu. Rev. Biomed. Eng.*, vol. 12, no. 1, pp. 119–142, Jul. 2010.

[3] A. Scorza, S. Conforto, C. D'Anna, and S. A. Sciuto, "A comparative study on the influence of probe placement on quality assurance measurements in B-mode ultrasound by means of ultrasound phantoms," *Open Biomed. Eng. J.*, vol. 9, no. 1, pp. 164–178, Jul. 2015.

[4] K. Saini, M. L. Dewal, and M. Rohit, "Ultrasound imaging and image segmentation in the area of ultrasound: A review," *Int. J. Adv. Sci. Technol.*, vol. 24, pp. 41–60, Nov. 2010.

[5] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. N. Hernes, "Freehand 3D ultrasound reconstruction algorithms—A review," *Ultrasound Med. Biol.*, vol. 33, no. 7, pp. 991–1009, 2007.

[6] Z. Chen and Q. Huang, "Real-time freehand 3D ultrasound imaging," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 1, pp. 74–83, Jan. 2018.

[7] Q. Cai, C. Peng, J. C. Prieto, A. J. Rosenbaum, J. S. A. Stringer, and X. Jiang, "A low-cost camera-based ultrasound probe tracking system: Design and prototype," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 997–999.

[8] M. R. Morgan, J. S. Broder, J. J. Dahl, and C. D. Herickhoff, "Versatile low-cost volumetric 3-D ultrasound platform for existing clinical 2-D systems," *IEEE Trans. Med. Imag.*, vol. 37, no. 10, pp. 2248–2256, Oct. 2018.

[9] H. K. Zhang, T. Y. Fang, R. Finocchi, and E. M. Boctor, "High resolution three-dimensional robotic synthetic tracked aperture ultrasound imaging: Feasibility study," *Proc. SPIE*, vol. 10139, Mar. 2017, Art. no. 1013914.

[10] F. Aalamifar, R. Khurana, A. Cheng, R. H. Taylor, I. Iordachita, and E. M. Boctor, "Enabling technologies for robot assisted ultrasound tomography: System setup and calibration," *Proc. SPIE*, vol. 9040, Apr. 2014, Art. no. 90401X.

[11] A. Ahmad, M. C. Cavusoglu, and O. Bebek, "Calibration of 2D ultrasound in 3D space for robotic biopsies," in *Proc. Int. Conf. Adv. Robot. (ICAR)*, Jul. 2015, pp. 40–46.

[12] L. E. Bø, E. F. Hofstad, F. Lindseth, and T. A. N. Hernes, "Versatile robotic probe calibration for position tracking in ultrasound imaging," *Phys. Med. Biol.*, vol. 60, no. 9, pp. 3499–3513, May 2015.

[13] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle, "An ultrasound robotic system using the commercial robot UR5," *Frontiers Robot. AI*, vol. 3, p. 1, Feb. 2016.

[14] 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 Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1342–1349, Apr. 2020.

[15] Z. Jiang, M. Grimm, M. Zhou, Y. Hu, J. Esteban, and N. Navab, "Automatic force-based probe positioning for precise robotic ultrasound acquisition," *IEEE Trans. Ind. Electron.*, early access, Nov. 16, 2020, doi: 10.1109/TIE.2020.3036215.

[16] B. K. Patle, G. B. Loganathan, A. Pandey, D. R. K. Parhi, and A. Jagadeesh, "A review: On path planning strategies for navigation of mobile robot," *Defence Technol.*, vol. 15, no. 4, pp. 582–606, Aug. 2019.

[17] C. Mineo, S. G. Pierce, P. I. Nicholson, and I. Cooper, "Robotic path planning for non-destructive testing—A custom MATLAB toolbox approach," *Robot. Comput.-Integr. Manuf.*, vol. 37, pp. 1–12, Feb. 2016.

[18] Q. Huang, J. Lan, and X. Li, "Robotic arm based automatic ultrasound scanning for three-dimensional imaging," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1173–1182, Feb. 2019.

[19] C. Graumann, B. Fuerst, C. Hennersperger, F. Bork, and N. Navab, "Robotic ultrasound trajectory planning for volume of interest coverage," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2016, pp. 736–741.

[20] S. Virga, O. Zettinig, M. Esposito, K. Pfister, B. Frisch, T. Neff, N. Navab, and C. Hennersperger, "Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms," in *Proc. IEEE Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 508–513.

[21] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[22] K. M. Meiburger, U. R. Acharya, and F. Molinari, "Automated localization and segmentation techniques for B-mode ultrasound images: A review," *Comput. Biol. Med.*, vol. 92, pp. 210–235, Jan. 2018.

[23] J. Chevrie, N. Shahriari, M. Babel, A. Krupa, and S. Misra, "Flexible needle steering in moving biological tissue with motion compensation using ultrasound and force feedback," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2338–2345, Jul. 2018.

[24] B. Takacs and T. Haidegger, "Autonomous applied robotics: Ultrasound-based robot-assisted needle insertion system concept and development," in *Proc. IEEE 15th Int. Conf. Syst. Syst. Eng. (SoSE)*, Jun. 2020, pp. 487–492.

[25] O. Zettinig, B. Frisch, S. Virga, M. Esposito, A. Rienmüller, B. Meyer, C. Hennersperger, Y.-M. Ryang, and N. Navab, "3D ultrasound registration-based visual servoing for neurosurgical navigation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 9, pp. 1607–1619, Sep. 2017.

[26] C. M. Heunis, Y. P. Wotte, J. Sikorski, G. P. Furtado, and S. Misra, "The ARMM system–autonomous steering of magnetically-actuated catheters: Towards endovascular applications," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 705–712, Apr. 2020.

[27] P. J. S. Gonçalves, P. M. B. Torres, F. Santos, R. António, N. Catarino, and J. M. M. Martins, "A vision system for robotic ultrasound guided orthopaedic surgery," *J. Intell. Robot. Syst.*, vol. 77, no. 2, pp. 327–339, Feb. 2015.

[28] N. Gupta, A. P. Shukla, and S. Agarwal, "Despeckling of medical ultrasound images: A technical review," *Int. J. Inf. Eng. Electron. Bus.*, vol. 8, no. 3, p. 11, 2016.

[29] R.-M. Menchón-Lara and J.-L. Sancho-Gómez, "Fully automatic segmentation of ultrasound common carotid artery images based on machine learning," *Neurocomputing*, vol. 151, pp. 161–167, Mar. 2015.

[30] H. Cui, Y. Xia, and Y. Zhang, "Supervised machine learning for coronary artery lumen segmentation in intravascular ultrasound images," *Int. J. Numer. Methods Biomed. Eng.*, vol. 36, no. 7, Jul. 2020, Art. no. e3348.

[31] E. Marchand, F. Spindler, and F. Chaumette, "ViSP for visual servoing: A generic software platform with a wide class of robot control skills," *IEEE Robot. Autom. Mag.*, vol. 12, no. 4, pp. 40–52, Dec. 2005.

[32] A. S. Deo and I. D. Walker, "Overview of damped least-squares methods for inverse kinematics of robot manipulators," *J. Intell. Robot. Syst.*, vol. 14, no. 1, pp. 43–68, Sep. 1995.

[33] N. Hogan, "Impedance control: An approach to manipulation: Part II—Implementation," *J. Dyn. Syst., Meas., Control*, vol. 107, no. 1, pp. 8–16, Mar. 1985.

[34] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," in *Proc. ACM SIGGRAPH Papers*, 2011, pp. 1–12.

[35] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.

[36] F. Šuligoj, B. Jerbić, B. Šekoranja, J. Vidaković, and M. Švaco, "Influence of the localization strategy on the accuracy of a neurosurgical robot system," *Trans. FAMENA*, vol. 42, no. 2, pp. 27–38, Jun. 2018.

[37] M. W. Gilbertson and B. W. Anthony, "Force and position control system for freehand ultrasound," *IEEE Trans. Robot.*, vol. 31, no. 4, pp. 835–849, Aug. 2015.

FILIP ŠULIGOJ received the M.Sc. degree in mechanical engineering and the Ph.D. degree in robotics and automation from the Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia, in 2009 and 2018, respectively. In 2020, he joined the Surgical Robotics Laboratory, University of Twente, The Netherlands, as a Postdoctoral Researcher. His research interests include medical robotics, machine vision with an emphasis on medical image analysis, and artificial intelligence methods. His current research interests include medical image processing methods and control of robots in surgical procedures.

CHRISTOFF M. HEUNIS (Student Member, IEEE) received the M.Sc. degree in mechatronic engineering from the University of Stellenbosch, Stellenbosch, South Africa, in 2016. He is currently pursuing the Ph.D. degree with the Surgical Robotics Laboratory, University of Twente, Enschede, The Netherlands. His work focuses on adapting and integrating embedded sub-systems for health care applications, with a specific focus on surgical robotics, magnetic actuation systems, and medical image processing. His research interest includes the design of clinical equipment for the treatment of endovascular, cardiovascular, and oncological disorders.

JAKUB SIKORSKI (Member, IEEE) received the M.Eng. degree in biomedical engineering from The University of Glasgow, Glasgow, U.K., in 2015, and the Ph.D. degree in biomechanical engineering from the University of Twente, Enschede, The Netherlands, in 2020. He currently works as a Postdoctoral Researcher with the Surgical Robotics Laboratory, University of Twente. His work focuses on the development of novel robotic instruments for minimally-invasive surgical procedures. His research interests include continuum and soft robotics, magnetic devices, mechatronic design, and control of biomedical systems.

SARTHAK MISRA (Senior Member, IEEE) received the master's degree in mechanical engineering from McGill University, Canada, and the Ph.D. degree in mechanical engineering from Johns Hopkins University, USA, in 2009. He is currently a Full Professor with the Department of Biomechanical Engineering, University of Twente, Enschede, The Netherlands, and also affiliated with the Department of Biomedical Engineering, University of Groningen and University Medical Center Groningen, The Netherlands. Prior to commencing his Ph.D. studies, he was a Dynamics and Controls Analyst on the International Space Station Program. His research interests include surgical robotics and medical microrobotics. He was a recipient of the European Research Council Starting, Proof-of-Concept, and Consolidator Grants and the Netherlands Organization for Scientific Research VENI and VIDI Awards. He is the Co-Chair of the Robotics and Automation Society Technical Committee on Surgical Robotics and the International Federation of Automatic Control Technical Committee on Biological and Medical Systems.

• • •