Intravascular Tracking of Micro-Agents Using Medical Ultrasound: Towards Clinical Applications

Filip Šuligoj ⬤, Christoff M. Heunis ⬤, *Member, IEEE*, Sumit Mohanty ⬤, and Sarthak Misra ⬤, *Senior Member, IEEE*

***Abstract*—*Objective:* This study demonstrates intravascular micro-agent visualization by utilizing robotic ultrasound-based tracking and visual servoing in clinically-relevant scenarios. *Methods:* Visual servoing path is planned intraoperatively using a body surface point cloud acquired with a 3D camera and the vessel reconstructed from ultrasound (US) images, where both the camera and the US probe are attached to the robot end-effector. Developed machine vision algorithms are used for detection of micro-agents from minimal size of 250μm inside the vessel contour and tracking with error recovery. Finally, real-time positions of the micro-agents are used for servoing of the robot with the attached US probe. Constant contact between the US probe and the surface of the body is accomplished by means of impedance control. *Results:* Breathing motion is compensated to keep constant contact between the US probe and the body surface, with minimal measured force of 2.02 N. Anthropomorphic phantom vessels are segmented with an Intersection-Over-Union (IOU) score of 0.93 ± 0.05, while micro-agent tracking is performed with up to 99.8% success rate at 28–36 frames per second. Path planning, tracking and visual servoing are realized over 80 mm and 120 mm long surface paths. *Conclusion:* Experiments performed using anthropomorphic surfaces, biological tissue, simulation of physiological movement and simulation of fluid flow through the vessels indicate that robust visualization and tracking of micro-agents involving human patients is an achievable goal.**
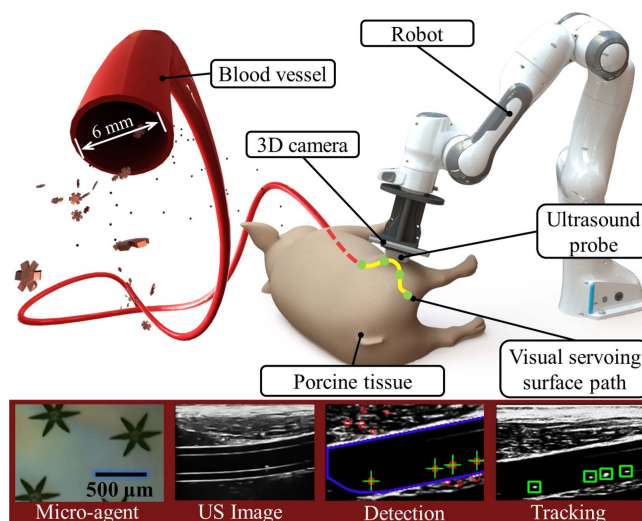
Fig. 1. Robotic ultrasound-based intravascular micro-agent tracking and visual servoing in simulated clinical conditions.

***Index Terms*—Machine vision, medical robotics, microrobotics, ultrasound imaging, visual servoing.**

I. Introduction

THE vision to make clinical interventions less invasive has provided impetus to remotely-actuated micro-agents in the field of microrobotics. Microrobotics entails various technologies that revolve around the design, synthesis and remote actuation of sub-millimeter sized agents. Over the years, these technologies have sought to revolutionize therapeutic and diagnostic procedures while minimizing operation time and risks offered by tethered surgical instruments [1]. A plethora of different micro-agents have been employed in clinical applications such as targeted drug delivery [2], [3], stem cell transplants [4], thrombolysis [5], and microsurgery [6]. Various forms of remote actuation methods such as based on magnetic [7], acoustic [8], optical [9] and chemical [10] principles have been used to deploy micro-agents in the aforementioned clinical applications. Despite the growing outreach of these micro-agents, the numerous challenges involved in deploying them under *in-vivo* conditions have confined most of the studies to lab-based experimental test-beds. Besides the unpredictable conditions offered

by biological environments, a quintessential challenge is reliable localization of micro-agents under such *in-vivo* conditions [11]. This necessitates application of medical imaging modalities in order to reliably deploy these micro-agents in clinically-relevant environments.

Previously, various microrobotic studies have employed intra-operative imaging methods such as magnetic resonance imaging (MRI) [12], clinical ultrasound (US) [13] and fluoroscopy [14] to assess the targeted regions of their deployment. However, MRI provides lower acquisition rates of imaging whereas fluoroscopic imaging may constantly expose a patient to ionizing radiations. As a result, US imaging is the most preferred and ubiquitous form of imaging modality used for clinical applications of micro-agents [5], [7], [8], [13], [15]–[20]. In many of the above clinical studies, tracking of micro-agents under US guidance enables their closed loop control [15]–[17]. Various studies demonstrate US guided visual servoing of micro-agents in workspaces mimicking vasculatures [5], [7], and under *ex-vivo* conditions [6], [20].

Despite numerous microrobotic studies that employ US imaging, their applicability to clinical practice is limited due to the challenging tracking conditions for micro-agents offered by biological environments. This is owing to the inability of tracking methods to adequately resolve fine anatomical features and compensate for mechanical disturbances inside biological specimen. Some of these challenges could be alleviated by obtaining pre-operative US images of targeted environment prior to deployment of micro-agents inside a biological specimen [21]. Moreover, the limited spatial resolution of the US-based tracking methods also applies to micro-agents that are harder to localize in deep seated regions as they possess an overall smaller footprint compared to the targeted region of interest [19]. These aforementioned limitations make US based tracking of micro-agents noise-prone and thus less reliable for long operation times within a biological specimen. In addition, various clinical applications desire deployment and tracking of multiple micro-agents for their effective operation such as in case of targeted drug delivery in a specimen. Here, the inadequate spatial resolution of a single-agent detection compromises with the overall tracking performance of multiple micro-agents and thus their inaccurate prediction [13]. Hence, establishing reliable control of micro-agents in clinical procedures require (a) intraoperative real-time imaging, (b) robust methods to track essential features inside the biological specimen, and (c) visualize multiple micro-agents in tandem.

Besides microrobotics studies, a diverse range of tracking algorithms from general applications can be used for localization of micro-agents in US images. Desirable traits of tracking algorithms for localization of micro-agents are tracking success under occlusions, accuracy and real-time processing. Fiaz *et al.* [22] evaluated performance of numerous benchmark tracking algorithms [23]–[29]. The conclusion of Fiaz's study was that the performance of DCF-CSR algorithm [29] displayed best results when tracking in conditions with background clutter, deformation, illumination changes and low resolution. Aforementioned imaging limitations can be further reduced by producing

adequate robotic response to adjust the probe position while acquiring US images.

While concept of visual servoing is a mature research topic in robotics [30]–[32], its translation to clinical practice requires adaptation for any specific procedure or imaging modality [33]. Based on this notion, a study by Zettinig *et al.* [34] showed that US visual servoing can be used for neurosurgical navigation and needle guidance. Furthermore, in our previous studies, we have demonstrated both optical visual servoing [21] and impedance control [35] used for body motion compensation while performing robotic US scanning along the planned surface path.

As one of the many potential applications for methods presented in this paper, we establish a clinical scenario analogous to Carotid endarterectomy (CEA). CEA is a procedure used to treat carotid disease caused by a buildup of plaque in arteries that deliver blood to the brain. As an alternative to the conventional procedure, with this methodology, micro-agents containing statins can be injected into the radial artery, picked up by the remote actuation system under the guidance of a clinical imaging system, and guided through the axillary artery to reach the site of the plaque and to commence targeted drug delivery. To achieve this, we propose a novel path planning algorithm that uses only intraoperative data, a surface point cloud acquired using a 3D camera that is registered to the underlying vessels segmented from US images. Furthermore, we present a methodology for intravascular ultrasound-based micro-agent tracking and visual servoing in clinically-relevant scenarios (Fig. 1). To the best of our knowledge, this study is the first to demonstrate visual servoing using micro-agents located inside the biological specimen and the anthropomorphic phantom while simulating physiological functions like breathing and blood flow.

In Section II, firstly, artificial vessels which are used of micro-agents transport are segmented to define the region of interest in the US images. In order to ensure robust tracking, our developed algorithm employs object detection alongside the tracking algorithm. Furthermore, it identifies errors associated with the tracking frame, thus ensuring automatic tracking initialization and error recovery. Subsequently, planned surface path and the positions of the tracked micro-agents are used for visual servoing of the robot with the attached US probe while keeping constant contact with the surface of the body being scanned using the impedance control. In Section III, application of the developed algorithms is demonstrated in clinically relevant scenarios which include the anthropomorphic phantom, porcine tissue and simulated physiological functions i.e., breathing motion and fluid flow through the vessels. In Section IV, we conclude with an overview of the results and provide future perspective on this study.

II. Intravascular Micro-Agent Tracking

Robotic micro-agent tracking in 3D space can be divided into two tasks: micro-agent localization in US images and visual servoing of the US probe by means of a robot arm. The first major challenge is to reliably extract 3D coordinates of a micro-agent using 2D US slices of the vessel cross-section. Here, we introduce a tracking algorithm (Fig. 2) specialized for
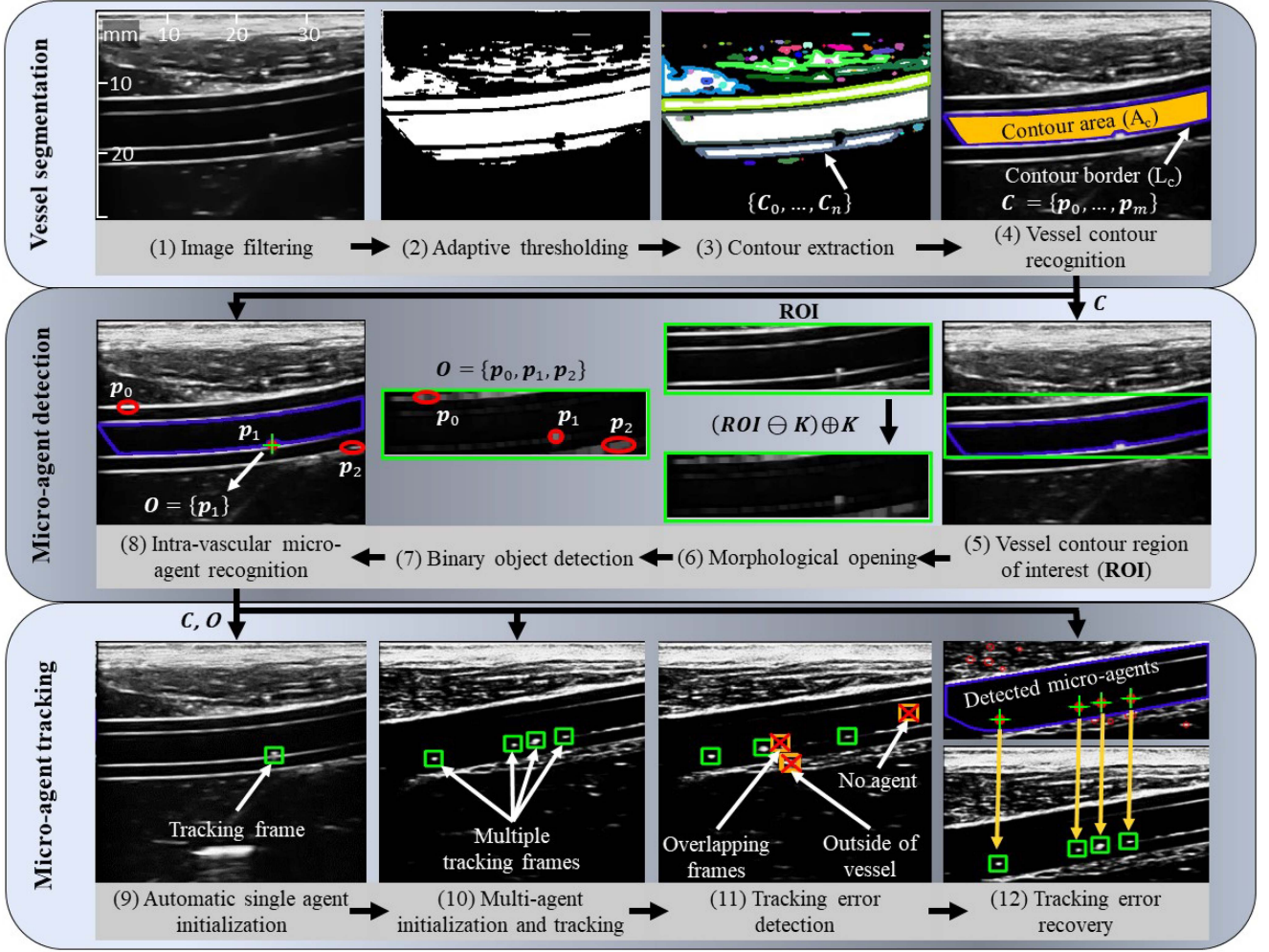
Fig. 2. Machine vision application for micro-agent detection and tracking. (1) Image preprocessing of ultrasound (US) slice for superfluous noise reduction and tissue homogenization. (2) Adaptive thresholding is used for obtaining the binary (black and white) image. (3) A set of n contours ($\{\mathbf{C}_1,\ldots,\mathbf{C}_n\}$) is derived from the binary image, where every contour is defined with m border points ($\{\mathbf{p}_1,\ldots,\mathbf{p}_m\}$) and $\mathbf{p} \in \mathbb{Z}_+^{1\times 2}$. (4) The vessel segmentation algorithm is used to detect the vessel contour ($\mathbf{C}$) based on border length ($L_\mathbf{C}$), surface area ($A_\mathbf{C}$) and shape ($A_\mathbf{C}/L_\mathbf{C}$). (5) Region of interest (**ROI**) is extracted from the original US image as a minimal horizontal bounding box around the vessel contour. (6) Morphological opening i.e., dilatation ($\oplus$) of the erosion ($\ominus$) is applied to **ROI** using the rectangular structuring element $\mathbf{K}$. (7) Binary objects are detected inside the processed **ROI** based on their intensity, size and shape (circularity, convexity, inertia). (8) Detected objects with positions found inside the polygon of the contour $\mathbf{C}$ or inside the user defined distance from its border are recognised as intravascular micro-agents. Micro-agent initial positions are stored in a set $\mathbf{O} = \{\mathbf{p}_1,\ldots,\mathbf{p}_r\}$ with r being the number of detected micro-agents. (9) Single or (10) multiple tracking frames are initialized with center positions in the set $\mathbf{O}$ and default initial frame width (w) and height (h). (11) During continuous tracking, error detection algorithm is used to stop individual trackers in case of overlapping frames, position of the tracking frame outside of vessel contour or if mean intensity value of the tracking frame image has changed significantly from the initial one. (12) In case of tracking error the tracking frames are automatically re-initialized using micro-agent detection algorithm.

localization of the micro-agents and handling of occlusions. In this study we will use the term occlusions for significant change in appearance or loss of the object i.e., micro-agent from the image. For example, body movement can result in misalignment of the US image plane with the micro-agent location, for which reason it is paramount to establish a robust detection to work alongside the tracking algorithm. Secondly, a path planning algorithm is developed that uses the body surface point cloud and the vessel centroids (based on intraoperative US surface scan). Finally, aforementioned planned surface path is used for visual servoing of the US probe to the position of the tracked micro-agent.

A. Ultrasound Based Micro-Agent Tracking

Developed algorithms are used for vessel segmentation and localization of the intravascular micro-agents in the US images. Firstly, vessel segmentation is used both to extract vessel centroids for path planning and to reduce the region of interest around the micro-agent during its localization. Edge preserving filter [36] (EPF) is used to remove noise from the US image. After filtering, the image is binarized using the adaptive threshold algorithm. Each enclosed white area in the binary image is extracted as a set of its outer contour points with each point ($\mathbf{p} \in \mathbb{Z}_+^{1\times 2}$). The result is a set of contours ($\{\mathbf{C}_1,\ldots,\mathbf{C}_n\}$) with n being the number of contours, where the actual vessel contour

$(\mathbf{C} = \{\mathbf{p}_1, \ldots, \mathbf{p}_m\})$ is recognized based on expected size and shape of the vessel.

The object detection algorithm is used to localize the micro-agents inside the vessel, and subsequently, initialize the tracking frame or re-initialize it in cases of tracking errors. The contour $\mathbf{C}$ of the segmented vessel is used to define a rectangular region of interest (**ROI**) inside the original US image in which the micro-agents are detected. Detection of binary objects inside the vessel contour is based on the difference in intensity between the intravascular fluid and the micro-agents in the US image. Additional image processing is used to isolate the micro-agents from the surrounding environment e.g., detection of a micro-agent which is in contact with the vessel wall of similar intensity, as shown in Fig. 2.

Next, a tracking algorithm is used for robust and real-time localization of a frame containing the object of interest in successive images. In comparison to object detection, tracking holds advantages when it comes to handling occlusions. Using the detected micro-agent positions $(\mathbf{p}_k)$ from the set $\mathbf{O}$ along with initial width (w) and height (h), tracking frames $(\mathbf{F}_k)$ with $k = 1, \ldots, r$ are initialized. The tracking algorithm which we utilized uses discriminative correlation filter with channel and spatial reliability (DCF-CSR) [29]. DCF-CSR tracker uses spatial reliability map for adjusting the filter support to the part of the selected region from the frame for tracking. Herein, it enables scaling and localization of the selected region and robust tracking of the non-rectangular objects. The algorithm continuously updates the position $(\mathbf{p}_k)$, width w_k, height h_k, and image matrix $(\mathbf{I}_k)$ of the tracking frame $(\mathbf{F}_k)$. An error detection algorithm is developed to run in parallel to the tracking algorithm so that it can re-initialize trackers in case of overlapping tracker frames:

$$\forall k = 1, \ldots, r \wedge \forall l = 1, \ldots, r | \ \mathbf{F}_k \cap \mathbf{F}_l \neq \emptyset \wedge k \neq l, \quad (1)$$

tracking frame position outside of vessel contour area $(A_\mathbf{C})$:

$$\forall k = 1, \ldots, r | \ \mathbf{F}_k \cap A_\mathbf{C} = \emptyset, \quad (2)$$

or if current $(curr)$ mean intensity (μ) of the tracking frame image matrix has changed significantly from the initial one $(init)$:

$$\mu(\mathbf{I}_{k,init}) \gg \mu(\mathbf{I}_{k,curr}) \vee \mu(\mathbf{I}_{k,curr}) \gg \mu(\mathbf{I}_{k,init}). \quad (3)$$

During initial detection of the micro-agents, those which are not in conflict with (1) and (2) are automatically tracked.

Finally, any pixel position in the US image is transformed to 3D coordinates in the frame $(\{\mathcal{B}\})$. Position of the tracked micro-agent in frame $(\{\mathcal{B}\})$ is calculated using the coordinates of $\mathbf{p}_k$ in the US image matrix (a-th column and b-th row):

$$\begin{bmatrix} {}^{\mathcal{B}}\mathbf{P}_k \\ 1 \end{bmatrix} = {}^{\mathcal{F}}_{\mathcal{B}}\mathbf{T} {}^{\mathcal{U}}_{\mathcal{F}}\mathbf{T} \begin{bmatrix} af - \dfrac{g}{2} & 0 & be & 1 \end{bmatrix}^T, \quad (4)$$

where g is the width of the probe and e and f are the height and width of any pixel, respectively. Homogeneous transformation between any two frames $(\{\mathcal{Y}\}$ and $\{\mathcal{X}\})$ is denoted as ${}^{\mathcal{X}}_{\mathcal{Y}}\mathbf{T} \in \mathrm{SE}(3)$. Robot flange pose is denoted as $({}^{\mathcal{F}}_{\mathcal{B}}\mathbf{T})$ and flange to US probe tip transformation $({}^{\mathcal{U}}_{\mathcal{F}}\mathbf{T})$ is derived by means of calibration.

B. US Probe Visual Servoing on a Body Surface

The prerequisite for achieving visual servoing on the body surface is a set of end-effector poses in frame $(\{\mathcal{B}\})$ from which the vessel containing the micro-agents can be imaged (Fig. 3). These end-effector poses are generated based on the surface geometry captured using the 3D camera and the positions of the vessel centroids derived from intraoperative US scan by using the vessel segmentation algorithm described in section II.A. Surface scanning and the impedance control used for realizing the robot motion are both described in our prior work [35]. Path planning algorithm developed for this application is designed to calculate a surface path for acquiring US images containing consecutive centroids in sagittal cross-section of the vessel. During US scanning impedance control is used to maintain the probe-to-surface contact while acquiring US images.

The input variables for the path planning algorithm registered to frame $(\{\mathcal{B}\})$ are the 3D point cloud of the body surface (κ) and consecutive centroids of the vessel defined as a set of points (λ), with each point $\lambda_q \in \mathbb{R}^3$. The algorithm starts by finding the nearest neighbour point in (κ) to the first point in (λ). The initial surface point $(\omega_{q=1} \in \mathbb{R}^3)$ for imaging of the centroid point $(\lambda_{q=1})$ is chosen as the one with minimal lenght $||\overline{\lambda_q \omega_q}||$ with angle (α_q) (defined between the line $\overline{\lambda_q \omega_q}$ and normal (n_q) to local surface plane) being smaller than α_{min}:

$$\omega_q = \underset{\mathbf{p} \in \{\kappa\}}{argmin}(||\overline{\lambda_q \mathbf{p}}||) \wedge \alpha_q < \alpha_{min}. \quad (5)$$

Normal to local surface plane unit vector (n_q) for point (ω_q) is calculated using the local neighborhood of points in a radius of size (g) around it. Radius (g) should be equal to the width of the probe. Using the function that solves the eigenvalues and eigenvectors of a given covariance matrix we estimate the least-squares plane normal (n_q). Optimal probe contact while scanning is achieved if the angle (α_q) between the US probe z-axis and the local surface normal is equal to zero. For each subsequent ω_q the initial search is started at the previous surface point (ω_{q-1}) and selected as:

$$\omega_q = \underset{\mathbf{p} \in \{\kappa\}}{argmin}(||\overline{\omega_{q-1} \mathbf{p}}|| + ||\overline{\lambda_q \mathbf{p}}||) \wedge \alpha_q < \alpha_{min}. \quad (6)$$

ensuring that when the vessel is being scanned, the US probe is moved only over short distances between any two consecutive vessel points and shallow imaging of the vessel centroid is secured.

US probe orientation $(\vartheta_q \in \mathrm{SO}(3))$ is defined in a way that the vessel centroid (λ_q) is horizontally centered in the US image that is also aligned with the line fitted between all the vessel centroid points that can fit in the US image Fig. 3.(1). In this manner we ensure that the US image plane is intersecting with the estimated middle cross-section of the vessel for imaging the micro-agents inside. Path point position (ω_q) and orientation (ϑ_q) together define a transformation $({}^{\mathcal{A}_q}_{\mathcal{B}}\mathbf{T})$. The final output of the path planning algorithm is a complete set of robot end-effector path poses $({}^{\mathcal{A}_1}_{\mathcal{B}}\mathbf{T}, \ldots, {}^{\mathcal{A}_s}_{\mathcal{B}}\mathbf{T})$.

The main objective of visual servoing is to move the probe along the planned path while keeping the tracked micro-agent in the US image plane. Once the distance of the micro-agent to the
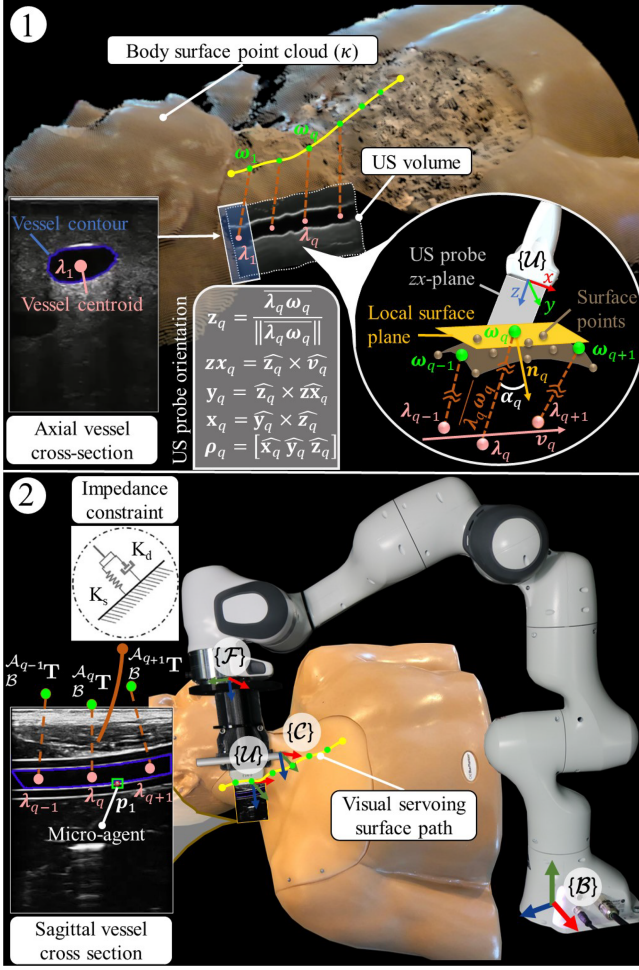
Fig. 3. (1) Path planning algorithm: Each surface path point ($\boldsymbol{\omega}_q$) (denoted as green points), where $q = 1, \ldots, s$, is derived from a point cloud ($\boldsymbol{\kappa}$) and a vessel centroid point ($\boldsymbol{\lambda}_q$) denoted as a pink point, where s is equal to the number of vessel centroid points. Position of each ($\boldsymbol{\omega}_q$) is selected in a way to minimize $||\overline{\boldsymbol{\lambda}_q \boldsymbol{\omega}_q}||$ and $||\overline{\boldsymbol{\omega}_q \boldsymbol{\omega}_{q-1}}||$ while keeping minimal chosen perpendicularity to the local surface $\alpha_q < \alpha_{min}$. The orientation ($\boldsymbol{\vartheta}_q$) of the ultrasound (US) probe in a path point ($\boldsymbol{\omega}_r$) is defined using the line $\overline{\boldsymbol{\lambda}_q \boldsymbol{\omega}_q}$ as z-axis and a vector ($\mathbf{v}_q$) that lies in the zx-plane of the probe frame ($\{\mathcal{U}\}$). Alignment of the probe with the sagittal vessel cross section is secured by calculating $\mathbf{v}_q$ using linear regression based on a set of points $\{\boldsymbol{\lambda}_{q-1}, \boldsymbol{\lambda}_q, \boldsymbol{\lambda}_{q+1}\}$. The planned path in the robot base frame ($\{\mathcal{B}\}$) is a set of poses $\left({}^{\mathcal{A}_1}_{\mathcal{B}}\mathbf{T}, \ldots, {}^{\mathcal{A}_s}_{\mathcal{B}}\mathbf{T}\right)$, where ${}^{\mathcal{A}_q}_{\mathcal{B}}\mathbf{T}$ is based on position ($\boldsymbol{\omega}_q$) and orientation ($\boldsymbol{\vartheta}_q$). (2) The robot base ($\{\mathcal{B}\}$), robot flange ($\{\mathcal{F}\}$), camera ($\{\mathcal{C}\}$), and ultrasound (US) probe ($\{\mathcal{U}\}$). Once tracking of the micro-agents has been initiated, as a response to their movement the algorithm continuously re-evaluates distance to the nearest vessel centroid ($\boldsymbol{\lambda}_q$) and moves the US probe frame ($\{\mathcal{U}\}$) to ${}^{\mathcal{A}_q}_{\mathcal{B}}\mathbf{T}$ with impedance control equilibrium point defined on the line $\overline{\boldsymbol{\lambda}_q \boldsymbol{\omega}_q}$. Stiffness ($\mathbf{K_s} \in \mathbb{R}^{6\times6}_{\geq0}$) and damping ($\mathbf{K_d} \in \mathbb{R}^{6\times6}_{\geq0}$) matrices are used to define the force-position relation between the US probe and the contact surface.

vessel centroid $||_{\mathcal{B}}\mathbf{p}_k - \boldsymbol{\lambda}_q|| < ||_{\mathcal{B}}\mathbf{p}_k - \boldsymbol{\lambda}_{q-1}||$, the impedance controler is used to move to the US probe to the the path pose ${}^{\mathcal{A}_q}_{\mathcal{B}}\mathbf{T}$ with impedance control equilibrium point positioned on the line $\overline{\boldsymbol{\lambda}_q \boldsymbol{\omega}_q}$. The impedance controler is used to ensure probe contact with the body surface during robot movement. Stiffness

$\mathbf{K_s} \in \mathbb{R}^{6\times6}_{\geq0}$ and damping $\mathbf{K_d} \in \mathbb{R}^{6\times6}_{\geq0}$ matrices define the force-position relation. For example, resulting contact force in the US probe z-direction can be modelled as ${}^z_{\mathcal{U}}F = K_z(\delta z)$, with K_z being z-axis stiffness and δz being $\mathcal{U}$ frame z-axis distance between the US probe location and equilibrium point position. By changing the parameters in $\mathbf{K_s}$ and $\mathbf{K_d}$ and the depth of the equilibrium point along the line $\overline{\boldsymbol{\lambda}_q \boldsymbol{\omega}_q}$ we achieve constant contact with deformable and moving surfaces.

III. EXPERIMENTS

The experiments are designed to evaluate vessel segmentation algorithm (Section III.A), the ability to compensate for breathing motion while scanning the surface (Section III.B), the tracking of the micro-agents (Section III.C) and visual servoing along the planned path (Section III.D).

Hardware components of the system used for the experiments are: a 7-DoF robot (Panda, Franka Emika GmbH) with integrated torque sensors and force sensing capabilities, a medical US machine (SonixTouch Q+, BK Medical) with a linear US L14-5/38 mm transducer which generates US images at 20 Hz, and a 3D camera (Intel Realsense SR305). The end-effector of the robot is mounted with the 3D camera and the US probe. Robot control and machine vision algorithms 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 (RT-PREEMPT). Point Cloud Library (PCL v1.9.1) is used for handling of 3D data and path planning, while Open Source Computer Vision Library (OpenCV v4.4.0) is used for image processing and tracking.

Experiments are performed using the gelatin-agar phantom (80 g/L gelatin and 10 g/L agar), anthropomorphic phantom (Gen II Ultrasound Central Line Training Model, CAE Healthcare, USA, a human body replica used for training medical personnel in diagnostic and interventional US procedures that is filled with simulated blood solution) and biological tissue (porcine hind limb obtained from a local slaughterhouse) referred to as gel#, ant# and bio#, respectively. Physiological motion is simulated by moving the phantom with motion amplitude of 10 mm at $\approx$0.27 Hz as reported in [37] for an average human chest movement during breathing. In order to establish artificial blood flow, a silicone tubes (6–10 mm inner diameter) with similar US characteristics of a human vessel [38] are inserted into gel# and bio#. The blood flow is simulated through the artificial vessel using a peristaltic pump (ISM 404, Ismatec, Wertheim, Germany) with water flow of 12-640 mL/min as a reference to the average adult human blood flow [39]. For each experiment, one or more micro-agents are manually inserted into the open end of the artificial vessel and their position inside the vessel is changed with a hand-held motion of an external permanent magnet. Micro-agents used in the experiments are 250 μm, 500 μm, and 1000 μm diameter magnetic beads (Nd-FeB) and $30 \times 500 \times 500 \mu m$ soft magnetic star-shaped agents. The star-shaped agents are 3-D printed on an IP-Dip photoresin using Direct Laser Writing technique (Nanoscribe GmbH, Germany) and later coated with metallic nanofilms to enable their
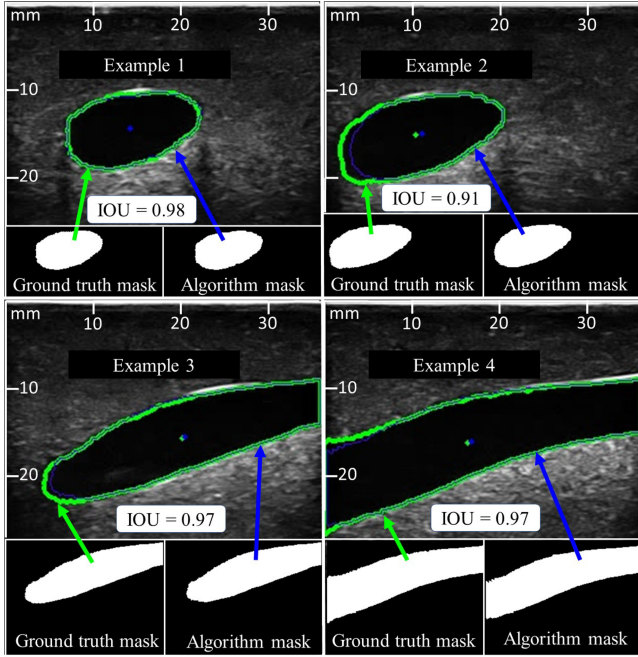
Fig. 4. Evaluation of vessel segmentation using Intersection over Union (IOU): During robotic US scanning, with US transducer frequency ($f = 10$ Mhz), dynamic range ($DR = 45$ dB) and US image gain ($K_u = 0.6$), the US probe is rotated so that it acquires a total of (275) ultrasound images (256×300 pixels, 38×45 mm) of the vessel from different angles. Edge preserving filter (EPF) threshold coefficient is set to 0.40, and size of the neighbourhood is 15px. Adaptive threshold is employed with with mean subtracting constant of 13, in pixel neighborhood of size 300px. Contour filter is set to remove all the contours that do not fit the parameters, $265 < L_C < 1500$ and $A_C/L_C < 7$. **Please refer to accompanying video for a demonstration of this experiment.**

magnetic actuation [40]. These layers made of NiFe (500 nm) and Au (20 nm) are deposited using thin film sputtering process.

A. Vessel Segmentation Experiment

The US imaging plane can intersect the vessel from various angles and therefore display arbitrary cross-sectional contours which need to be robustly detected and segmented using the proposed algorithm. The vessel segmentation algorithm is tested using ant# because visualization of the vessels in US images is the most similar to human blood vessels shown in the study [41]. The performance of the segmentation algorithm is evaluated based on two metrics, the ability to detect the vessel and the segmentation accuracy. Detection performance is measured as a binary value (true or false). Truthful detection means that the algorithm has correctly identified the presence (or lack of) of the vessel in the US image. In our experiment, the vessel contour is correctly detected in 100% of the cases i.e., in 275 out of 275 of the US images (Fig. 4). The Jaccard similarity coefficient, i.e., the Intersection over Union (IOU) is a standard metric used for evaluating the similarity and diversity of the binary segmentation masks in 2D images. IOU accuracy of the segmentation algorithm mask compared to the manually defined
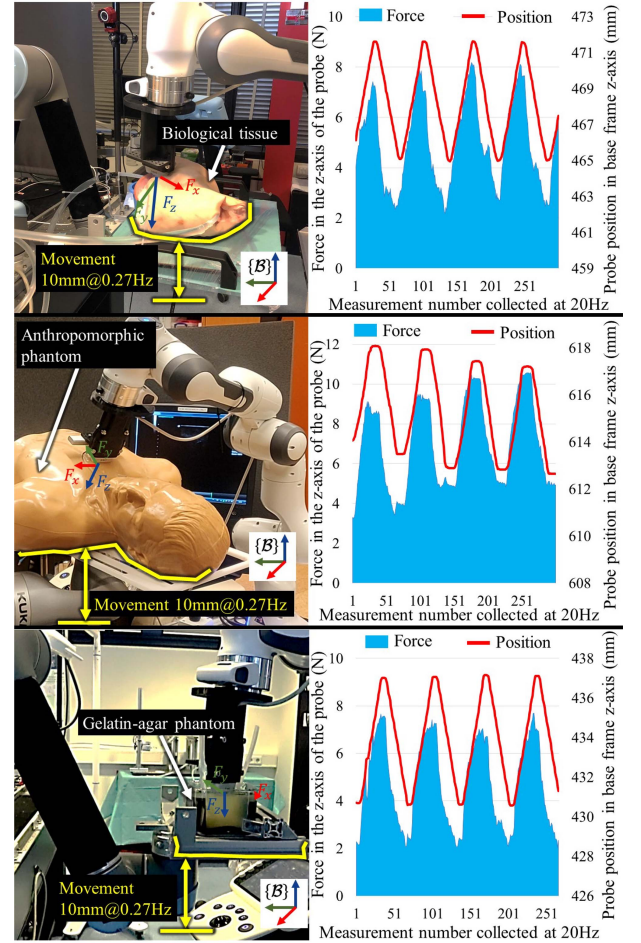
Fig. 5. Force and position feedback measurements from the robot are recorded at 20 Hz when simulating breathing motion with the bio#, ant# and gel#. Impedance control compliance matrix is set to $\mathbf{K_s} = diag\,([3000\ \ 3000\ \ 100\ \ 300\ \ 300\ \ 300])$ and $\mathbf{K_d} = 2\,\mathbf{K_s}^{\circ(1/2)}$. Each equilibrium point is translated along the z-axis of the US probe for $\delta z = 80$ mm from the starting surface point. **Please refer to accompanying video for a demonstration of these experiments.**

ground truth mask from twenty randomly selected images is 0.93 ± 0.05.

B. Breathing Motion Compensation Experiment

Robotic US scanning needs to be performed with constant contact between the US probe and the scanned surface. We utilize the an impedance controler to compensate for the breathing motion and surface deformations while manipulating the US probe on the body surface. The results of the experiment with moving surfaces of bio#, ant# and gel# are presented in Fig. 5. Each sample is moved using the secondary robot to simulate the breathing motion, while the US probe position in frame $(\mathcal{B})$ and the force in frame $(\mathcal{U})$ are measured. Force measured in the z-direction of frame $(\mathcal{U})$ on the bio#, ant# and gel# is 4.98 ± 1.70N, 7.04 ± 2.18N and 4.73 ± 1.82N, with maximum position amplitudes in $(\mathcal{B})$ frame z-axis of 6.67 mm, 5.43 mm and 6.60 mm, respectively. With lowest measured force of 2.02 N, this experiment demonstrates that the developed impedance

controler can successfully keep constant probe-to-surface contact with bodies that exhibit breathing motions, have different stiffness and shape.

C. Agent Tracking Experiment

The developed machine vision application, which integrates vessel segmentation, micro-agent detection and tracking is tested in series of different conditions which simulate clinical conditions (Fig. 6). Experiments measure the tracking and error recovery ability along with speed of processing in multiple testing scenarios. In this study, the tracking success rate is defined as a percentage of time in which one or more micro-agents (based on the exact known number of introduced agents) are tracked within the vessel. In the experiments, 250–1000 μm agents are observed with US while being subjected to breathing motion body movement, fluid flow through vessels and externally induced motion in bio#, ant# and gel#. Externally induced micro-agent motion is deliberately induced using a hand-held permanent magnet to re-position the agents out of US imaging plane and to test the error recovery ability. The vessels are located at depths in the range of 10–40 mm from the scanned surface. Bead micro-agents are inserted into the vessel using non-magnetic pincers, while star-shaped micro-agents are introduced by means of pipetting. Typically, 4-5 star-shaped agents are pipetted with 50uL aqueous solution into the vessels in a single experiment. After each experiment the micro-agents are pulled to the open end of the vessel using a permanent magnet and extracted.

While specific situations are best demonstrated in accompanying tracking videos, the developed machine vision application is able to detect and track single agents up to 99.8% of total experiment time at 28–36 frames per second (FPS) (images of size 256×300 pixels). Tracking of multiple (four) micro-agents is successfully achieved at 13FPS despite multiple occlusions and micro-agent overlapping. Successful tracking error recovery caused by micro-agent occlusions is demonstrated in experiments Bio_4, Bio_5, Bio_7, Bio_8, Ant_2, Ant_3, Ant_4, Gel_1 and Gel_4. The tracking frame drift is identified as powerful factor that influences tracking success rate in our experiments. Frame drift occurs in situations when the significant portion of the tracking object is no longer visible, and in our experiments it is induced by the artefacts on the vessel walls and micro-agent occlusions e.g., Ant_2 and Gel_4. This issue could be mitigated if the motion of the micro-agent is controlled and the agent is continuously re-positioned to the center of the vessel, thus avoiding occlusions.

Furthermore, an argument could be made that the echogenicity of blood affiliated particles, in comparison to water which is used in our experiments with fluid flow, would be an aggravating factor for the detection of the micro-agents. This subject should be investigated in future studies, even though the US frequencies in clinical use are essential for the size of the particles or anatomical structures that are being observed [42]. Hence, we expect that impaired visualization or tracking would not come to pass as long as the micro-agents and the particles found in blood are of different sizes, and adequate US frequency is used.

LEGEND — Flow (mL/min); Breathing motion body movement; Micro-agent movement; Star micro-agent (µm); Bead micro-agent (µm); Number of Micro-agents; Micro-agent occlusion

BIOLOGICAL TISSUE RESULTS (Original image; Vessel segmentation; Micro-agent detection; Micro-agent tracking)

Bead micro-agent (µm)	Star micro-agent (µm)	Breathing motion body movement	Micro-agent movement	Flow (mL/min)	Number of Micro-agents	Micro-agent occlusion	Segmentation speed (FPS)	Detection speed (FPS)	Tracking speed (FPS)	Tracking success rate (% of total time)	Micro-agent not visible (% of total time)	Successful error recoveries	Video name
250	-	-	✓	-	1	-	61	3424	79	99.8	-	0	Bio_1
250	-	✓	✓	-	1	-	58	3264	78	99.8	-	0	Bio_2
500	-	✓	✓	-	2**	-	68	2506	79	99.8	-	0	Bio_3
500	-	✓	✓	12	1	-	51	2350	82	99.5	-	1	Bio_4
500	-	✓	✓	12	2**	✓	47	2097	76	99.0	-	1	Bio_5
500	-	✓	-	500	1	-	45	2586	76	99.8	-	0	Bio_6
1000	-	✓	-	640	1	✓	61	2668	93	88.3	~10*	2	Bio_7
-	30 × 500	✓	-	-	4	✓	54	1388	18	59.0	~30*	7	Bio_8

ANTHROPOMORPHIC PHANTOM RESULTS (Original image; Vessel segmentation; Micro-agent detection; Micro-agent tracking)

Bead micro-agent (µm)	Star micro-agent (µm)	Breathing motion body movement	Micro-agent movement	Flow (mL/min)	Number of Micro-agents	Micro-agent occlusion	Segmentation speed (FPS)	Detection speed (FPS)	Tracking speed (FPS)	Tracking success rate (% of total time)	Micro-agent not visible (% of total time)	Successful error recoveries	Video name
500	-	-	✓	-	1	-	58	721	78	98.0	-	0	Ant_1
1000	-	-	✓	-	1	✓	51	1141	106	77***	~5*	7	Ant_2
1000	-	✓	✓	-	1	✓	58	2352	88	81.8	~3*	4	Ant_3
1000	-	-	✓	-	1	✓	58	1904	80	55.0	~20*	5	Ant_4

GELATIN-AGAR PHANTOM RESULTS (Original image; Vessel segmentation; Micro-agent detection; Micro-agent tracking)

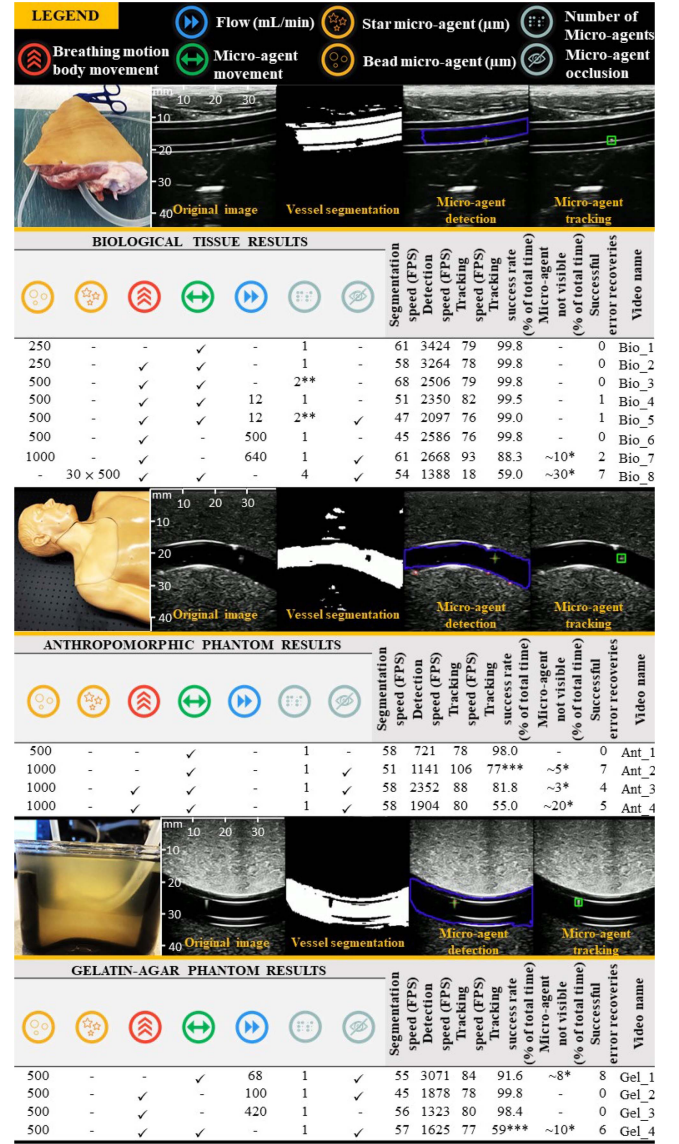| Bead micro-agent (µm)	Star micro-agent (µm)	Breathing motion body movement	Micro-agent movement	Flow (mL/min)	Number of Micro-agents	Micro-agent occlusion	Segmentation speed (FPS)	Detection speed (FPS)	Tracking speed (FPS)	Tracking success rate (% of total time)	Micro-agent not visible (% of total time)	Successful error recoveries	Video name
500	-	-	✓	68	1	✓	55	3071	84	91.6	~8*	8	Gel_1
500	-	✓	-	100	1	✓	45	1878	78	99.8	-	0	Gel_2
500	-	✓	✓	420	1	-	56	1323	80	98.4	-	0	Gel_3
500	-	✓	✓	-	1	✓	57	1625	77	59***	~10*	6	Gel_4

Fig. 6. Evaluation of the integrated machine vision application (vessel segmentation, detection and tracking) in simulated clinical conditions. Ultrasound (US) transducer frequency is set to f=10Mhz, while dynamic range (DR) and gain (K_u) are manually adjusted for each case. Vessel segmentation and microagent detection parameters are set based on prior information on agent and vessel size and shape. *Percent of the total video time in which the micro-agent is not visible is manually timed and approximated. **Multiple micro-agents are attracted and joined due to magnetic forces and form a single tracking object. ***Tracking drift problem is present in the video (false positive tracking results). **Please refer to accompanying video for a demonstration of these experiments.**

D. Visual Servoing Experiment

The final experiment integrates all the algorithms demonstrates visual servoing to the position of an actuated micro-agent. Vessel centroids are extracted using the vessel segmentation algorithm, during robotic US scanning with transducer frequency ($f = 10$ Mhz), dynamic range ($DR = 45$ dB) and US image gain ($K_u = 0.6$), on images of size 256×300 pixels (38×45 mm).
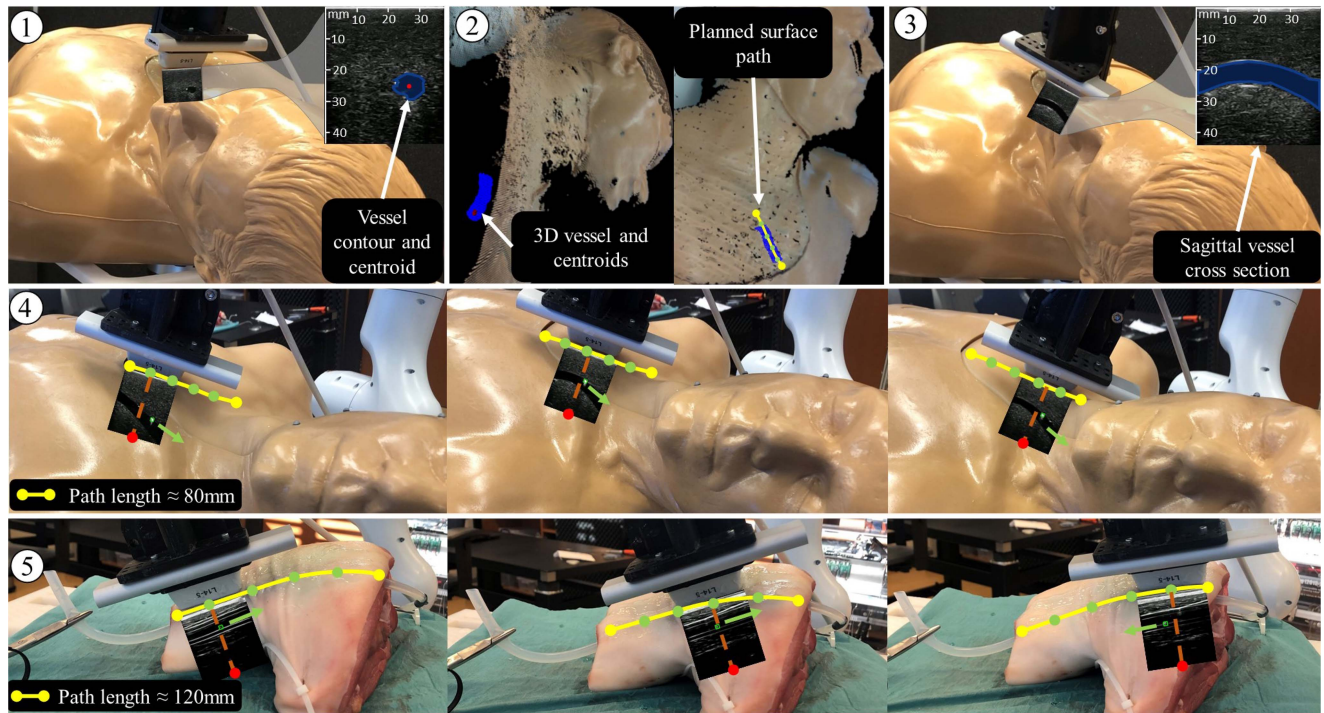
Fig. 7. Visual servoing experiment: (1) Initial ultrasound surface scanning for detection of vessel centroids; (2) 3D Vessel reconstruction, registration to body surface point cloud and path planning with parameter $\alpha_{min} = 10°$; (3) Planned path validation based on visualization of the sagittal cross-section of the vessel; Manual insertion and actuation of the micro-agent with robotic visual servoing to the tracked micro-agent positions in (4) ant# and (5) bio#. **Please refer to accompanying video for a demonstration of these experiments.**

The parameters of the filtering algorithm (EPF), adaptive threshold algorithm and the vessel contour are manually adjusted in the first US image. The vessel centroids and the body surface point cloud are registered to frame $(\mathcal{B})$ and used to generate a path for visual servoing using the algorithm described in Fig. 3. Image analysis of the acquired US images shows the algorithm was able to detect the vessel contour in the ant# as shown in Fig. 7.(1–3). Visual servoing is performed using the ant# and bio# with 1000 μm bead micro-agent. Successful visual servoing is performed in stationary and moving surface cases (breathing motion) along approximately 80 mm and 120 mm long paths as shown in Fig. 7.(4–5) and in the accompanying video. Methods and algorithms employed in this experiment can also be used along longer paths for clinical scenarios in which the US imaging line of sight is not obstructed.

Visual servoing is demonstrated with straight and moderately tortuous vessels, nevertheless, a limitation of the current work is a lack of experiments with complex vessel branches. This is due to the fact that manual actuation (based on a rotating magnetic field) limits the motion of the micro-agents to rolling in deeper seated vessels and hence prevents delivery and retrieval of the micro-agents from these locations.

IV. CONCLUSIONS & FUTURE WORK

In this study, we present methods for using a robotic US system to track intravascular micro-agents and perform visual servoing on a body surface in simulated clinical conditions.

The machine vision application performs vessel segmentation, detection and tracking of the intravascular micro-agents. Vessel segmentation experiment demonstrated 100% successful detection rate with 0.93 ± 0.05 IOU score in the ant#. Furthermore, detection and tracking algorithms are tested under various physiological conditions such as respiratory body movement, fluid flow thorough vessels, micro-agent movement and occlusions in bio#, ant# and gel#. Successful singe-agent tracking is demonstrated up to 99.8% of the time at 28–36FPS for agents of size 250–1000 μm. Multi-agent tracking under occlusions and overlapping is successfully performed using four star-shaped micro-agents at 13FPS. Breathing motion compensation is achieved using the impedance controler, with constant contact between the US probe and the surfaces of all the phantoms is measured in a range of 2.02–10.6 N. Finally, visual servoing was successfully performed while tracking the micro-agent using the planned paths on the static and moving surfaces of the ant# ($\approx$80 mm long path) and bio# ($\approx$120 mm long path).

Besides, the efficacy of segmentation and subsequent tracking performance could be improved with more controlled actuation of the micro-agent, whereby it can be reliably positioned in the centre of the vessel. Here, remote actuation such as mobile magnetic coils [43], or catheters with miniaturized magnetic coils [44] may overcome limitations of bulky and inaccessible hardware, typically associated with magnetic actuation. Future studies in this direction could accommodate such collaborative methods of tracking complemented by controlled actuation strategies in demanding scenarios such as arterial branches.

REFERENCES

[1] B. J. Nelson et al., "Microrobots for minimally invasive medicine," *Annu. Rev. Biomed. Eng.*, vol. 12, pp. 55–85, 2010.

[2] D. Jang et al., "Targeted drug delivery technology using untethered microrobots: A review," *J. Micromechanics Microengineering*, vol. 29, no. 5, 2019, Art. no. 053002.

[3] A. V. Singh et al., "Multifunctional magnetic hairbot for untethered osteogenesis, ultrasound contrast imaging and drug delivery," *Biomaterials*, vol. 219, 2019, Art. no. 119394.

[4] S. Jeon et al., "Magnetically actuated microrobots as a platform for stem cell transplantation," *Sci. Robot.*, vol. 4, no. 30, 2019, Art. no. eaav4317.

[5] Q. Wang et al., "Reconfigurable magnetic microswarm for thrombolysis under ultrasound imaging," in *Proc. Int. Conf. Robot. Automat.*, 2020, pp. 10285–10291.

[6] M. A. Ghanem et al., "Noninvasive acoustic manipulation of objects in a living body," *Nat. Acad. Sci.*, vol. 117, no. 29, pp. 16848–16855, 2020.

[7] F. Ongaro et al., "A contactless and biocompatible approach for 3D active microrobotic targeted drug delivery," *Micromachines*, vol. 10, no. 8, 2019, Art. no. 504.

[8] S. Mohanty et al., "Contactless acoustic micro/nano manipulation: A paradigm for next generation applications in life sciences," *Proc. Roy. Soc. A: Math., Phys. Eng. Sci.*, vol. 476, no. 2243, 2020, Art. no. 20200621.

[9] S. Zhang and A. R. Wheeler, "Light-driven microrobotics and their applications for single-cell manipulation," in *Proc. Int. Conf. Opt. Trapping Opt. Micromanipulation*, 2020, pp. 73–78.

[10] K. Villa et al., "Chemical microrobots as self-propelled microbrushes against dental biofilm," *Cell Rep. Phys. Sci.*, vol. 1, no. 9, 2020, Art. no. 100181.

[11] A. Aziz et al., "Medical imaging of microrobots: Toward *In Vivo* applications," *ACS Nano*, vol. 14, no. 9, pp. 10865–10893, 2020.

[12] X. Yan et al., "Multifunctional biohybrid magnetite microrobots for imaging-guided therapy," *Sci. Robot.*, vol. 2, no. 12, 2017, Art. no. eaaq1155.

[13] Q. Wang and L. Zhang, "Ultrasound imaging and tracking of micro/nanorobots: From individual to collectives," *IEEE Open J. Nanotechnol.*, vol. 1, pp. 6–17, 2020.

[14] A. Servant et al., "Controlled *In Vivo* swimming of a swarm of bacteria-like microrobotic flagella," *Adv. Mater.*, vol. 27, no. 19, pp. 2981–2988, 2015.

[15] S. Scheggi et al., "Magnetic motion control and planning of untethered soft grippers using ultrasound image feedback," in *Proc. Int. Conf. Robot. Automat.*, 2017, pp. 6156–6161.

[16] S. Scheggi et al., "A GPU-accelerated model-based tracker for untethered submillimeter grippers," *Robot. Auton. Syst.*, vol. 103, pp. 111–121, 2018.

[17] A. Sanchez et al., "Magnetic control of self-propelled microjets under ultrasound image guidance," in *Proc. 5th RAS/EMBS Int. Conf. Biomed. Robot. Biomechatronics*, 2014, pp. 169–174.

[18] Q. Chen et al., "Ultrasound tracking of the acoustically actuated microswimmer," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 3231–3237, Nov. 2019.

[19] S. Pane et al., "Real-time imaging and tracking of microrobots in tissues using ultrasound phase analysis," *Appl. Phys. Lett.*, vol. 118, no. 1, 2021, Art. no. 014102.

[20] E. E. Niedert et al., "A tumbling magnetic microrobot system for biomedical applications," *Micromachines*, vol. 11, no. 9, 2020, Art. no. 861.

[21] C. M. Heunis et al., "Real-time multi-modal sensing and feedback for catheterization in porcine tissue," *Sensors*, vol. 21, no. 1, 2021, Art. no. 273.

[22] M. Fiaz, A. Mahmood, and S. K. Jung, "Tracking noisy targets: A review of recent object tracking approaches," 2018, *arXiv:1802.03098*.

[23] H. Grabner et al., "Real-time tracking via on-line boosting," in *Proc. Brit. Mach. Vis. Conf. Brit. Mach. Vis. Assoc.*, 2006, pp. 1–6.

[24] B. Babenko, M. Yang, and S. Belongie, "Visual tracking with online multiple instance learning," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 983–990.

[25] J. F. Henriques et al., "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.

[26] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1409–1422, Jul. 2012.

[27] Z. Kalal, K. Mikolajczyk, and J. Matas, "Forward-backward error: Automatic detection of tracking failures," in *Proc. Int. Conf. Pattern Recognit.*, 2010, pp. 2756–2759.

[28] D. S. Bolme et al., "Visual object tracking using adaptive correlation filters," in *Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 2544–2550.

[29] A. Lukežič et al., "Discriminative correlation filter tracker with channel and spatial reliability," *Int. J. Comput. Vis.*, vol. 126, no. 7, pp. 671–688, 2018.

[30] J. Feddema and O. Mitchell, "Vision-guided servoing with feature-based trajectory generation (for robots)," *IEEE Trans. Robot. Automat.*, vol. 5, no. 5, pp. 691–700, Oct. 1989.

[31] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Automat.*, vol. 8, no. 3, pp. 313–326, Jun. 1992.

[32] R. Kelly et al., "Stable visual servoing of camera-in-hand robotic systems," *IEEE/ASME Trans. Mechatronics*, vol. 5, no. 1, pp. 39–48, Mar. 2000.

[33] M. Azizian et al., "Visual servoing in medical robotics: A survey. Part I: Endoscopic and direct vision imaging - techniques and applications," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 10, no. 3, pp. 263–274, 2014.

[34] O. Zettinig et al., "3D ultrasound registration-based visual servoing for neurosurgical navigation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 9, pp. 1607–1619, 2017.

[35] F. Šuligoj et al., "Robust-an autonomous robotic ultrasound system for medical imaging," *IEEE Access*, vol. 9, pp. 67456–67465, 2021.

[36] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," in *Proc. ACM SIGGRAPH*, 2011, vol. 30, no. 4, Art. no. 69.

[37] S. Quirk et al., "External respiratory motion analysis and statistics for patients and volunteers," *J. Appl. Clin. Med. Phys.*, vol. 14, no. 2, pp. 90–101, 2013.

[38] L. Maggi et al., "Development of silicon-based materials for ultrasound biological phantoms," in *Proc. IEEE Int. Ultrasonics Symp.*, 2009, pp. 1962–1965.

[39] C. P. Chung et al., "Flow volume in the jugular vein and related hemodynamics in the branches of the jugular vein," *Ultrasound Med. Biol.*, vol. 33, no. 4, pp. 500–505, 2007.

[40] S. Mohanty et al., "Acoustically-actuated bubble-powered rotational micro-propellers," *Sensors Actuators B: Chem.*, vol. 347, 2021, Art. no. 130589.

[41] P. Blanco, "Ultrasound-guided vascular cannulation in critical care patients: A practical review," *Medicina Intensiva (English Edition)*, vol. 40, no. 9, pp. 560–571, 2016.

[42] J. Battino, "Echogenicity of blood," *J. de Radiologie*, vol. 73, no. 12, pp. 705–708, 1992.

[43] J. Sikorski et al., "The ARMM system: An optimized mobile electromagnetic coil for non-linear actuation of flexible surgical instruments," *IEEE Trans. Magn.*, vol. 55, no. 9, Sep. 2019, Art. no. 5600109.

[44] J. Sikorski, S. Mohanty, and S. Misra, "MILiMAC: Flexible catheter with miniaturized electromagnets as a small-footprint system for microrobotic tasks," *IEEE Robot. Automat. Lett.*, vol. 5, no. 4, pp. 5260–5267, Oct. 2020.