Optical Flow and Acoustic Phase Analysis Comparison in Ultrasound-Based Microrobot Tracking

Giovanni Faoro, *Student Member, IEEE*, Veronica Iacovacci, *Member, IEEE*, and Arianna Menciassi, *Fellow, IEEE*

***Abstract*—Micrometer-sized devices hold the potential to revolutionize medicine by improving diagnostic abilities, increasing therapeutic efficiency and reducing adverse effects. To safely operate microrobots (MRs) inside the human body, accurate localization strategies based on medical imaging should be adopted. Ultrasound (US) imaging has recently gained attention for medical robotics and microrobotics thanks to its noninvasiveness and real-time performances. In this letter we compare US Acoustic Phase Analysis (US-APA), a state-of-the-art Phase tracking method, with US Optical Flow (US-OF), a computer vision Speckle tracking method. To this aim, different MRs dimensions, different locomotion strategies and working environments were considered. In particular, cylindrical magnetic MRs with diameters spanning from 1.20 down to 0.25 mm were fabricated. They were actuated in tissue-mimicking phantoms through a permanent magnet mounted on a robotic arm to reproduce intravascular rolling and in place vibration in tissue. US-OF proved to be comparable to US-APA in localizing vibrating MRs, achieving a tracking error comparable to one body length. It slightly improved localization during rolling, by consistently tracking the MR with errors lower than 0.6 body length. Additionally, US-OF outperformed US-APA in temporal performances, reaching an output rate forty times higher (about 40 Hz in US-OF and about 1 Hz in US-APA). We believe that these results demonstrate the possible effective use of Optical Flow for MRs tracking, with a relevant advantage in terms of portability and intelligibility.**

***Index Terms*—Acoustic phase analysis, medical microrobots, optical flow, ultrasound imaging.**

I. Introduction

THE final goal of medicine is to diagnose diseases as soon as possible and deliver therapy in a patient-specific way, when and where needed. Microrobots (MRs) have been recently introduced to move a step forward toward such goal, allowing microsurgical tasks and targeted drug release [1], [2], [3], [4], [5]. The safe operation of biomedical MRs is based on accurate and robust feedback about their position and locomotion status over time. Due to their small size (between few μm and a couple of mm), sensors integration for sensor-based tracking cannot be pursued and image-based tracking is typically adopted [4], [6], [7]. Several biomedical imaging strategies have been investigated for MRs localization ranging from optical imaging [8], computed tomography [9], positron emission tomography [10], single photon emission computed tomography [11], magnetic resonance or particle imaging [12], [13] to ultrasound (US) [14], [15], [16]. Each technique holds specific pros and cons when applied to MRs tracking in light of the stringent requirements this task sets in terms of spatial resolution, contrast, penetration depth and temporal performances. Nonetheless, US imaging appears particularly promising by combining good spatial resolution (100-500 μm), deep tissue real-time imaging (up to 20 cm), safe and noninvasive tissue interaction and relatively low cost of the equipment. Such a combination of optimal features makes US imaging one of the preferred option in medical robots and microrobots visualization and tracking [6], [7], [17], [18].

Several approaches have been proposed for US-based MRs localization, spanning from static to motion based methods. Static tracking strategies commonly rely on Brightness (B)-mode image analysis and can be distinguished into contrast-based [14], [19], template matching [20], [21] or deep learning [22]. Anyway, due to MRs small dimensions and generally low contrast, such approaches are often discarded in favor of motion based methods, as Doppler [15], [23] and Acoustic Phase Analysis (APA) [16], [24], which have demonstrated improved robustness and precision. Both these motion-based techniques inspect the acoustic phase signal coming from the radio frequency (RF) data of US, that contains motion information not related to contrast. For this reason, they can be referred to as Phase tracking methods. Despite relying both on acoustic phase data, they differ in analysis and performances. Doppler imaging is grounded on the computation of velocity maps by deriving the acoustic phase signal in time. Wang et al. successfully demonstrated Doppler-based MRs tracking in both parallel and counter flow experiments in vessels. Additionally, the possibility to track single helical MRs (2.15 mm in diameter and 7.30 mm in length) or nanoparticle swarms (0.6 to 2.4 mm as diameter of the swarm) was validated [15], [23]. On the other hand, US-APA is based on the application of selective motion filtering through

Manuscript received 4 August 2023; accepted 8 December 2023. Date of publication 4 January 2024; date of current version 18 January 2024. This letter was recommended for publication by Associate Editor Z. Zhang and Editor X. Liu upon evaluation of the reviewers' comments. This work was supported by the European Union's Horizon Europe Research and Innovation Program under Grant 101070066 (REGO). *(Corresponding author: Giovanni Faoro.)*

The authors are with the BioRobotics Institute of Scuola Superiore Sant'Anna, 56127 Pisa, Italy, and also with the Department of Excellence in Robotics and AI of Scuola Superiore Sant'Anna, 56127 Pisa, Italy (e-mail: giovanni.faoro@santannapisa.it; veronica.iacovacci@santannapisa.it; arianna.menciassi@santannapisa.it).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LRA.2024.3349815, provided by the authors.

Digital Object Identifier 10.1109/LRA.2024.3349815

Fourier analysis, allowing the rejection of physiological motions not related to microrobot movements [16], [24]. This approach proved robust with respect to different imaging conditions, such as low contrast environments, and was successfully integrated into a closed-loop visual servoing platform [25]. Recently, researchers demonstrated that US-APA can outperform Doppler in MRs localization thanks to its selectivity [26]. Anyway, the peculiar need for phase spectrum analysis limits the temporal performances of such precise localization method to only few Hz by requiring the acquisition and elaboration of multiple frames at once. In addition, RF data are commonly not available in clinical US machines, requiring specific research dedicated modules.

Motion analysis in US imaging has been investigated also through computer vision approaches based on B-mode images. This separate branch of US motion detection strategies can be referred to as Speckle tracking methods and accounts for echo particle image velocimetry (Echo PIV) and US optical flow (US-OF). Echo PIV is based on the detection of speckle patterns. A velocity map is computed by dividing the image through a grid and matching each grid cell with a portion of the subsequent image [27]. Conversely, OF is a mature computer vision technique that computes velocity maps by looking at brightness changes in the entire image [28]. It has been used in the context of medical robotics for, e.g., US-based motion compensation [29]. Echo PIV has been successfully applied for blood flow analysis through injection of scattering particles [30], [31], [32] while US-OF has been exploited for motion tracking, as in abdominal organs [33] or cardiovascular tissues [34], [35].

Despite their advantages in terms of data availability, intelligibility and time performances there are no evidence of their use for MRs localization. However, US-OF appears particularly suitable for MRs real-time tracking in complex environment as eligible to provide a compromise between robustness to physiological motion and high framerate for real-time tracking. Indeed, among the Speckle tracking methods, Echo PIV appears less suitable for MRs tracking since the quality of the velocity map depends on the resolution of the applied grid and on the distributed appearance of the agents. Thus, to achieve proper tracking performances for a single MR, a high resolution grid should be adopted hampering real-time application due to the increase in computational complexity. In this article, we investigate for the first time the possible application of US-OF-based MRs localization by comparing its spatial and temporal performances to the one obtained with US-APA. This will allow the validation of this computer vision approach in an unexplored yet scientifically relevant scenario as MRs tracking, in which OF can play a pivotal role thanks to the combination of robustness and portability, typical of motion and static based US tracking methods, respectively.

The comparison between US-OF and US-APA has been conducted considering different environments, different microrobots size and different locomotion modalities. More specifically, we considered experiments in 1) structured vessel-like environment to mimic fine microrobot control inside the vasculature for reaching a target position; 2) unstructured tissue-like environment to mimic tracking at the target site during therapy delivery. To evaluate algorithms robustness with respect to different appearance, we also evaluated different microrobots dimensions (from 1.20 down to 0.25 mm). Last, to assess the versatility with respect to different locomotion modalities, tracking performances during rolling locomotion and in place vibration (from 5 to 1 Hz) were verified. In place vibration enabled the evaluation of both tracking methods sensitivity to small net displacements covered by the microrobot with respect to a rest position.

The remainder of this article is organized as follows. Section II presents an overview of the experimental setup in terms of phantom and microrobot fabrication, magnetic actuation and US image acquisition units. Mathematical modeling of microrobot actuation is reported in Section III. Section IV specifies the investigated motion-based localization algorithms. Tracking performances of the two methods are reported in Section V whereas Section VI draws the conclusions of this letter.

II. Experimental Setup

In this section the details on the experimental setup are reported including the fabrication of US compatible phantoms and magnetic microrobots, as well as the robotic platform for proper magnetic control.

A. US Compatible Phantom

The two scenarios under study (locomotion in a vessel and vibration in a tissue) differ for the environment surrounding the MR.

To simulate intravascular locomotion, an agarose phantom was prepared. It was designed to mimic a tract of medium artery (3-4 mm in diameter) with the surrounding soft tissue in terms of size and acoustic properties. A 4 mm diameter rubber tube was placed in a rectangular mold ($5 \times 5 \times 11$ cm^3) in the desired anatomy-like position. Agarose powder (Sigma-Aldrich), at 2% v/v concentration to replicate acoustic properties of human tissues, was dissolved in a deionized and degassed water (dd-H2O)—soy milk (5% v/v) solution [36]. The solution was then poured into the rectangular mold and the rubber tube removed after reticulation to generate the desired lumen in the phantom.

To mimic an unstructured tissue-like environment, a second phantom, characterized by a homogeneous medium with several scattering elements and no landmarks, was manufactured. Agarose and Polyvinyl alcohol (PVA) powders (Sigma-Aldrich), at 2% v/v and 4% v/v respectively, were dissolved in a deionized and degassed water (dd-H2O)-soy milk (5% v/v) solution [37]. The solution was then poured into the rectangular mold and let cool down. This resulted into a sticky gel-like material compatible with ultrasound imaging.

An example of US images acquired on the two phantoms is reported in Fig. 1.

B. Microrobot Fabrication

Among the different actuation strategies for endoluminal devices at macro and micro scale, magnetism is one of the preferred approaches [38]. It is characterized by several

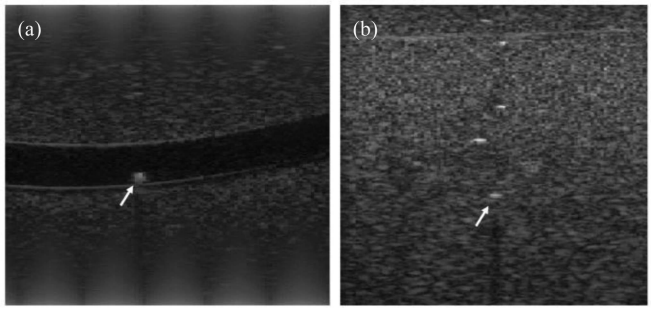

Fig. 1. Exemplifying US B-mode images of the microrobot in the two tissue mimicking phantoms. (a) Intravascular agar phantom with a visible vessel lumen. (b) US image of the tissue mimicking phantom. In both images the 0.84 mm MR is present (white arrow). While it is clearly recognizable in the intravascular environment, as characterized by different echogenicity with respect to the vessel lumen, it is hardly visible in the unstructured environment.

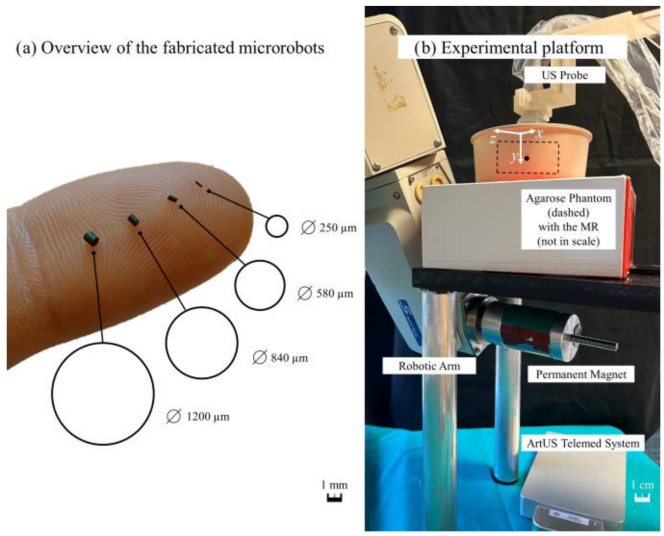

Fig. 2. Experimental setup. (a) An overview of the fabricated magnetic microrobots featuring variable diameter (from 1200 down to 250 μm), with 1 mm scale bar for reference. (b) The experimental platform, including a robotic arm, a permanent magnet and a US probe is depicted. The magnetic MR is placed inside the phantom (dashed in the figure) arranged in a small water tank to ensure optimal acoustic coupling. The reference frame associated to the US probe is reported in white.

advantages, such as tissue transparency to magnetic fields, real-time operation, precise and contactless control. Additionally, thanks to the adaptability of magnetic-based solutions in terms of device dimensions and fabrication methods, magnetism has been widely used for teleoperating catheters and MRs [39], [40]. For the above reasons magnetic MRs were selected in this study.

Both rolling and vibrating locomotion modalities can be covered by a surface microroller with remanent magnetization along the diametral direction. A submillimetric cylindrical structure was fabricated through extrusion-based printing of a UV curable magnetic ink [41]. To achieve uniform diametral magnetization, fundamental for proper rolling, the printed cylinder was magnetized diametrically by an impulse magnetizer with peak intensity field of 1.8 T (T-Series, Magnet-Physik Dr. Steingroever GmbH, Germany). The final height of the cylindrical MR was defined by cutting the magnetized string into smaller segments, having length of approximately 1 mm. Conversely, different MR diameters were achieved by changing printer nozzle dimensions. The ones considered in this study (Fig. 2(a)) were: 1.20, 0.84, 0.58, 0.25 mm.

C. Platform Hardware Components

The experiments were carried out by using a robotic platform including a six Degrees of Freedom (DoF) robotic arm (Melfa RV-3SB, Mitsubishi, Japan), a linear US probe (L15-7H40, Telemed, Lithuania) and a cylindrical permanent magnet (60 mm in diameter, 70 mm in height, NdFeB, diametral magnetization and grade N35).

The cylindrical permanent magnet is attached to the robotic arm through a custom holder (Fig. 2(b)). The use of a robotic arm allowed us to precisely control the spatial pose of the permanent magnet and thus to produce dynamic and versatile magnetic fields in the workspace. In particular, the robotic arm is controlled thanks to a custom Matlab Simulink Real-Time routine (version R2020a, The Mathworks Inc., Natick, MA).

The US probe was rigidly attached to the workbench to ensure constant optimal acoustic coupling with the phantom. Additionally, it was connected to an open architecture digital acquisition board (DAQ) (ArtUS, Telemed, Lithuania), which provides access to the raw RF data. These are directly used as input for US-APA analysis while for US-OF a preprocessing step, represented by logarithmic conversion to B-mode, is needed. All the imaging parameters, such as US frequency, gain and depth were kept constant throughout the experiments and were set to 10 MHz, 80% and 4 cm, respectively. Given these parameters, lateral resolution was 0.15 mm while the axial one was 0.06 mm for B-mode images and around 0.01 mm for RF data.

III. Magnetic Locomotion Models

Two different locomotion patterns were explored, in order to test algorithms robustness and precision in different physiological scenarios. Firstly, to mimic microrobot locomotion for reaching the target site of therapy release, constrained rolling inside a vessel is considered. Secondly, to mimic microrobot activation once at the target site, in place vibration inside a homogeneous and scattering medium is accounted.

By modeling the microrobot as a magnetic dipole ($\mathbf{m} \in \mathbb{R}^3$) it is possible to express the torques and forces generated by an external magnetic field $\mathbf{B} \in \mathbb{R}^3$ as:

$$\begin{cases} \tau = m \times B \\ F = \nabla (m \cdot B) \end{cases} \tag{1}$$

where $\times$ and $\cdot$ represent cross and dot product, respectively and ∇ indicate field gradient. By properly modeling the external field $\mathbf{B}$, the desired microrobot locomotion strategy can be achieved. To perform controlled rolling or vibration, we exploited the magnetic torque τ to align the magnetic MR with the external magnetic field $\mathbf{B}$. Among the three spatial components of $\mathbf{B}$, we should mainly focus on the two belonging to the US imaging plane (see Fig. 2(b), namely B_x and B_y since these are the ones playing a role in detectable microrobot motion.

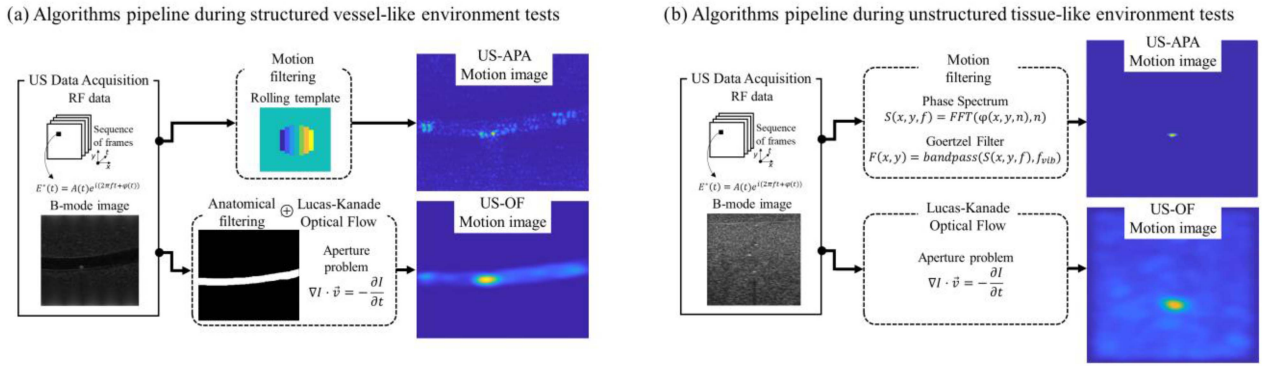

Fig. 3. Conceptual sketch of the motion based algorithms for microrobot localization under different testing conditions. In (a) the pipeline for the vessel-like environment is reported. In this case standard Lucas-Kanade Optical Flow is integrated with anatomical filtering considering vascular mask coming from US image segmentation. On the other hand, US-APA exploits the definition of a rolling motion template, searching for it through cross-correlation. In (b) the pipeline for the unstructured environment is shown. Here only Lucas-Kanade Optical Flow can be estimated as no relevant anatomies can be considered for spatial filtering.. Conversely, US-APA detects MR position based on the intensity of phase spectrum at the given vibration frequency. For each case ((a) and (b), US-OF and US-APA), an exemplifying motion image is reported, with the detected microrobot position as a yellowish region (i.e., a region of high motion).

To perform rolling locomotion over the lumen boundary, the microrobot should experience a rotating magnetic field in the *xy* plane. This can be formalized considering the following magnetic field time evolution:

$$\begin{cases} B_x = |\boldsymbol{B}| \sin(2\pi f_{rot} t) \\ B_y = |\boldsymbol{B}| \cos(2\pi f_{rot} t) \end{cases} \tag{2}$$

where f_{rot} represents the desired rotation frequency.

On the other hand, to let the microrobot vibrate in place the following time sequence for B_x and B_y was set:

$$\begin{cases} B_x = \bar{\theta}\,|\boldsymbol{B}| \sin(2\pi f_{vib} t) \\ B_y = |\boldsymbol{B}| \end{cases} \tag{3}$$

The field **B** oscillates in the *xy* plane over a circular sector defined by a maximum angle $\bar{\theta}$ at frequency f_{vib}. Thus, the induced microrobot vibrations have amplitude $\bar{\theta}$ and frequency f_{vib}.

The robotic platform allowed us to easily set the major magnetic control parameters such as f_{rot}, f_{vib} and $\bar{\theta}$ and to control the intensity **|B|** of the applied magnetic field. The distance between MR and permanent magnet was set to about 20 cm. Through dipole approximation, the maximum magnetic field intensity **|B|** generated by the permanent magnet was computed and resulted around 500 mT. The frequency of rotation f_{rot} and the vibration angle $\bar{\theta}$ were considered constant to 1.5 Hz and 35°, respectively. This choice was driven by considerations on MR control stability and optimal imaging conditions: the MR should not escape the field of view too quickly, but at the same time it should have detectable motions. Inversely, vibration frequency f_{vib} was varied from 5 to 1 Hz with decrements of 1 Hz to simulate conditions of net displacement with respect to the rest condition at decreasing velocities. This is remarkable to evaluate localization algorithms sensitivity to motion with different dynamics.

IV. Motion-Based Localization Algorithms

The US echo after the transmission of a sinusoidal pulse can be expressed as:

$$E^*(t) = A(t)\, e^{i(2\pi f t + \varphi(t))} \tag{4}$$

where, *A(t)* represents the amplitude of the signal (or pulse envelope) which is the physical quantity of interest for Speckle tracking methods. The frequency of the echo is represented by *f*, while *φ(t)* defines the echo phase signal, fundamental for Phase tracking methods. In the following subsections the working principle of US-APA and US-OF are briefly reported and a schematic of their pipeline for the two addressed experimental conditions is reported in Fig. 3.

A. US Acoustic Phase Analysis

US-APA is considered one of the best performing motion-based MRs tracking strategy due to its precision and robustness in localizing the MR. It is based on the analysis of the echo phase signal *φ(t)* [16], [24]. Indeed, shifts are impressed in this signal when a movement between the propagating US wave and the reflecting interface (i.e., the MR) occurs. Such shifts can be related to the projection *u* of object movements along the wave propagation direction and can be expressed as:

$$\Delta\varphi = \varphi_2 - \varphi_1 = \frac{u 4\pi}{\lambda} \tag{5}$$

where λ represents pulse wavelength.

Since the projection *u* depends on object movements, different microrobot locomotion strategies lead to different phase shifts. Therefore, a specific shift pattern can be identified and appropriately detected once a certain locomotion is imposed. The most relevant locomotion patterns investigated with such approach are in place vibration, rolling motion and corkscrew motion [16], [24], [26].

In place vibrations impress a sinusoidal variation of the phase signal *φ(t)* allowing MR localization through Goertzel filtering applied to the phase spectrum (at the MR vibration frequency

f_{vib}) [16]:

$$S(x,y,f) = FFT\ (\varphi(x,y,t),t) \tag{6}$$

$$F(x,y) = bandpass\,(S(x,y,f),f_{vib}) \tag{7}$$

On the other hand, rolling locomotion linearly modulates the phase signal by creating a two lobes pattern of positive and negative shifts. Microrobot location can be detected by searching for such rolling template in the phase image [24]. For a detailed description of US-APA mathematical modeling and implementation an interested reader can consult previous works by our research group [16], [24], [25], [26].

Despite the different phase shift identification routines, US-APA is always characterized by the need to analyze phase evolution in time. This means that a batch of multiple frames (i.e., a cineloop) is always needed for a single detection, thus producing a downscale of the output rate. Cineloop dimension is empirically set to 60 in the case of rolling locomotion. On the other hand, to guarantee the acquisition of at least two sinusoidal evolutions during in place vibration, cineloop dimension is set as follows:

$$cineloop\ =\ \frac{2*fps}{f_{vib}} \tag{8}$$

where *fps* indicates the frame per second rate of the US probe.

B. US Optical Flow

Optical flow is based on the analysis of pixel brightness changes among consecutive frames. Given a sequence of B-mode images *I(x,y,t)*, where *x* and *y* refers to the spatial coordinates of the pixel and *t* represents the time evolution, and pixel small displacements Δx and Δy in a short timeframe Δt, OF Aperture problem can be formulated:

$$\nabla I \cdot\ \vec{v} = -\frac{\partial I}{\partial t} \tag{9}$$

where $\vec{v}$ identifies the unknown velocity vector $(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t})$. This is obtained by considering Taylor's expansion (10) and brightness consistency constrain (11):

$$I\ (x+\Delta x,\ y+\Delta y,\ t+\Delta t) =\ I(x,y,t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t \tag{10}$$

$$I\ (x+\Delta x,\ y+\Delta y,\ t+\Delta t) =\ I(x,y,t) \tag{11}$$

The Aperture problem is an ill-posed problem since a single equation is available for the two unknown velocity components. For this reason, additional constrains are commonly considered distinguishing the adopted method (e.g., Horn-Schunk [42] or Lucas-Kanade [43]).

In this study we considered standard Lucas-Kanade US- OF with the integration of anatomical information coming from an artificial intelligence (AI) routine. When considering locomotion in a vessel-like environment, a VGG16-UNet architecture is exploited for vessel lumen identification from US images, so to limit US-OF computation to an image portion. Based on the assumption that the MR is rolling inside the vascular lumen, US-OF is computed only for the pixels corresponding to this anatomical region, where MR is expected to stay. This is indicated in Fig. 3(a) as "Anatomical filtering". Additionally, the information on the relative position of magnetic actuation unit (i.e., permanent magnet) with respect to the US probe can be exploited to further refine such region of interest by drawing assumptions on the vessel border on which the microrobot is rolling. The details on network training and performances are reported in [37]. Conversely, when no anatomical constrains can be set, as in the case of in place vibration inside the tissue, the whole image is analyzed. Last, raw US-OF output is filtered in space through Gaussian filtering, to smoothen the motion map, and in time through a moving median filter, in order to stabilize the microrobot tracked position. This comes from the assumption that the microrobot locomotion is always happening in a smooth way, without sudden and continuous jumps from one location to another. In particular, the Gaussian filter has dimensions 15x15 pixels with standard deviation 5 and the buffer size of the moving median filter is set to 20, which corresponds to a time window of 0.3 s given the US probe *fps* around 70 Hz. Such parameters were empirically set to guarantee US-OF output smoothness without hampering promptness.

V. Results

This section is devoted to the assessment of algorithms performances in terms of spatial tracking error and output rate under different environmental conditions and locomotion strategies. Experimental results are also shown in the ***Supporting Video***.

A. Tracking of a Rolling MR in a Vascular Phantom

Precise tracking in vascular anatomy is fundamental for accurate control during MR locomotion to reach pathological targets. Each magnetic MR (from 1.20 to 0.25 mm in diameter) was actuated inside the vasculature-mimicking phantom while acquiring the sequence of RF data (n = 3). To objectively evaluate the localization accuracies of US-APA and US-OF, a ground truth trajectory was defined by exploiting the information on vessel lumen features. Based on the relative position of the phantom with respect to the permanent magnet, the MR is expected to move along the bottom vessel boundary. From the AI generated vascular segmentation [37], such boundary is determined, and the expected MR trajectory *T* is computed as the one parallel to this border and elevated from it by a quantity equal to MR radius under testing. Once the expected MR trajectory is defined, an objective tracking error can be estimated by root square error minimization. Given a tracked MR position (D_x, D_y), to define a tracking error (e_x, e_y) the closest expected MR trajectory point (T_x, T_y) can be identified by considering (12):

$$T(i) = min_i\ \left\{\sqrt{(T_x(i)-D_x(t))^2+(T_y(i)-D_y(t))^2}\right\} \tag{12}$$

This process allows us to compute the optimal index *i* for the proper reference point (T_x, T_y) against which to compare algorithms results at each time step *t*. Once the optimal index *i* is defined, the *x* and *y* tracking error (e_x, e_y) can be computed

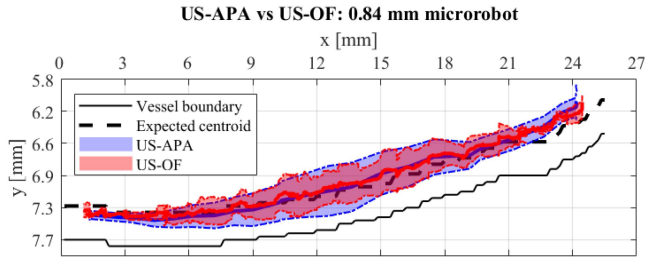

Fig. 4. Graphical comparison of the tracked trajectories for 0.84 mm diameter MR rolling inside a vascular phantom. The continuous black line represents the vessel inferior boundary extracted from image segmentation. The expected MR trajectory is reported with a dashed black line. 3 repetition mean trajectories obtained with the two algorithms are reported with a colored line while a colored shadow represents standard deviation (blue and red for US-APA and US-OF respectively).

as:

$$\begin{cases} e_x(t) = \frac{|T_x(i) - D_x(t)|}{d_{MR}} \\ e_y(t) = \frac{|T_y(i) - D_y(t)|}{d_{MR}} \end{cases} \tag{13}$$

where d_{MR} stays for the MR diameter. This allows us to compute directly unitless body length errors. The minimization process reported in (12) should be conducted separately for US-APA and US-OF due to their different output rate and thus the differences in tracked sequence length. As we indicated in Section IV A, US-APA needs the acquisition of a cineloop (60 frames) to provide MR position. On the other hand, US-OF provides an update on the MR tracked position for each new acquired image. This means that, in the timeframe required by US-APA to provide one tracked position, US-OF already computed 60 points. This makes the sequence D of detected points for US-OF around 60 times longer than the one corresponding to US-APA.

However, it should be remarked that tracking analysis was performed offline. Real-time analysis framerate cannot reach 60 Hz or more. To evaluate also the possible real-time implementation of US-OF, we monitored the time needed for each tracking update. The US-OF computing time was 23.8 ms, corresponding to an output rate of 42 Hz, showing a vast improvement with respect to US-APA (1-2 Hz).

A graphical comparison of the two tracked trajectories with respect to the expected one, for the case of 0.84 mm MR, is reported in Fig. 4. The graphical comparisons of tracking performances for the other MRs are shown in Part I of the *Supporting Video*. Numerical results, as mean and median error, standard deviation (σ) and interquartile range (IQ), for all different microrobots are reported in Table I.

US-OF reported consistently better tracking performances compared to US-APA with major improvement when reducing MR diameter. This deterioration in US-APA performances for smaller MR (i.e., 0.25 mm) can be due to an alteration in the locomotion pattern and thus in the impressed phase signal.

B. In Place Vibration Inside Tissue

To mimic unfavorable conditions of MRs in soft tissue (low contrast mismatch) without anatomical markers and in case of small net displacement, each magnetic MR was vibrated in place inside the tissue mimicking phantom. Also in this scenario, each trial was repeated 3 times to ensure robustness and repeatability. The absence of anatomical references forced us to change the definition of position landmark for error estimation. In this scenario US-APA and US-OF tracking error were computed through (13) by considering (T_x, T_y) as constant and defined manually by visual inspection of the US images sequence.

Also in this testing case a significant difference in algorithms output rate can be noted. When considering in place vibrations, there is a direct dependence of cineloop dimension on the vibration frequency, as stated in (8). In particular, lowering the vibration frequency imposes an increase in cineloop dimension and thus a reduction in US-APA promptness. As an example, when considering vibration frequencies < 2 Hz the cineloop acquisition requires more than 1 s (cineloop $>$ fps) and thus US-APA output rate further reduces below 1 Hz. On the other hand, US-OF output rate does not depend on MR locomotion features, thus remaining unaltered as compared to rolling.

A graphical comparison of US-APA and US-OF MR tracked centroid with respect to the manual landmark, for the case of 0.84 mm in diameter and 1 Hz vibration, is depicted in Fig. 5. An ellipse representation is considered, indicating the mean tracked position with a colored dot, while standard deviation is represented by the ellipse axes, thus showing data distribution. Additionally, a comparison among different MR diameters and vibration frequencies is shown in Fig. 6 indicating the evolution of median body length errors for the two tracking strategies

TABLE I
BODY LENGTH TRACKING ERRORS OF US-APA AND US-OF FOR DIFFERENT MICROROBOTS DIAMETERS

MR diameter	Axis	Mean		Median		σ		IQ	
		APA	OF	APA	OF	APA	OF	APA	OF
1200 µm	x	0.06	0.04	0	0	0.12	0.09	0.06	0
	y	0.15	0.11	0.14	0.08	0.12	0.07	0.17	0.09
840 µm	x	0.07	0.10	0	0	0.15	0.17	0	0.08
	y	0.12	0.07	0.08	0.08	0.07	0.05	0.08	0.08
580 µm	x	0.17	0.09	0	0	0.34	0.14	0.26	0.12
	y	0.22	0.10	0.17	0.09	0.26	0.07	0.12	0.09
250 µm	x	0.64	0.32	0	0	1.20	0.52	1.20	0.60
	y	0.76	0.40	0.60	0.36	0.76	0.24	0.28	0.20

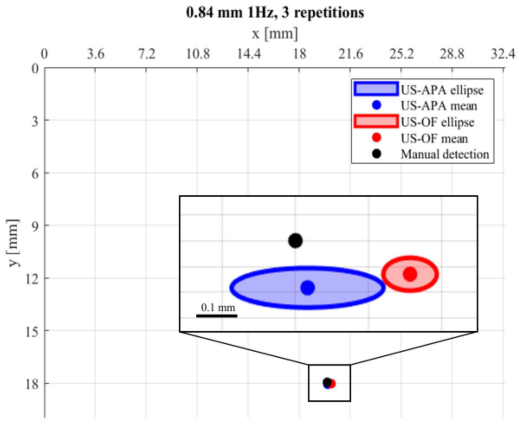

Fig. 5. Graphical comparison of the detected vibrating MR centroid for 0.84 mm diameter and 1 Hz frequency. The black dot represents the visually detected MR centroid. US-APA and US-OF mean positions are reported with blue and red dot respectively. Algorithms output standard deviation in x and y are exploited to define a tracking ellipse to visually evaluate for stability throughout the 3 repetitions. A 0.1 mm scalebar is indicated for reference.

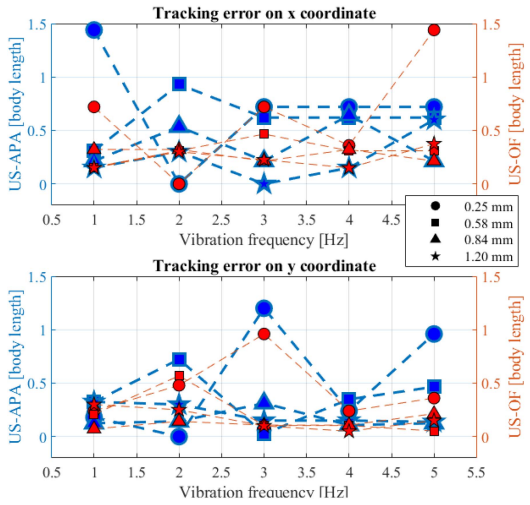

Fig. 6. Tracking error evolution of different MRs at different vibration frequencies. On the horizontal axis the vibration frequency is reported, while on the vertical one the body length error is shown. Blue color is used to indicate US-APA error, while red color is exploited for US-OF. Different MR diameters are illustrated by different markers as shown in the legend.

under consideration. A color-marker code, as indicated in the figure caption, is exploited to compare the mean tracking results, showing the absence of vibration dependency. Error boxplots for each experimental combination are reported in Part II of the *Supporting Video*.

In this scenario, no clear differences can be highlighted when comparing US-APA and US-OF localization performances. Additionally, we do not observe any trends indicating a dependence of tracking accuracy on vibration frequency or, differently from rolling, on MR diameter. This different behavior in US-APA when considering rolling or vibration at different scales can be due to the different detection strategies used for the two locomotion modalities: rolling is based on finding the best match with an expected motion template, whereas vibration is based merely on filtering.

VI. Discussion & Conclusion

In this article, we proposed the use of a mature computer vision method, namely Optical Flow, to track MRs in tissue through US images. The performances of this Speckle tracking method were compared with the state-of-the-art Phase tracking method, US-APA, both in terms of localization accuracy and update rate under different experimental conditions. In particular, different MR dimensions, different locomotion strategies and different surrounding environments were analyzed. US-OF consistently achieved submillimetric tracking accuracies in all tested conditions, reporting a body length error lower than 0.6 for the intravascular scenario and comparable to one for the *quasi-static* in place vibration in tissues. On the other hand, US-APA showed a deterioration in tracking performances when considering rolling locomotion of MRs as small as 0.25 mm (due to irregular locomotion pattern), while being comparable to US-OF in the other case studies.

It is not straightforward to compare such performances with other tracking methods proposed in literature. Indeed, [14] and [22] reported tracking accuracy as a binary metric, representing failure or success in microrobot detection. In [20] authors reported a deviation between detected position and expected vessel centerline in the range 0.5–1 body length when using template-based millirobot tracking. Similar performances were reported for Doppler-based millirobot localization [23]. Wang et al. reached microswarm (from 0.6 to 2.4 mm in diameter) localization error around 0.2 body length. Anyway, it is important to notice that these performances were obtained considering dynamic tasks in a vascular mimicking environment and thus they cannot be compared to the performances reported here for in place vibration. Overall, the achieved US-OF tracking performances are comparable with the one reported in the state-of-the-art for Phase tracking methods and, generally, overtakes the one of other B-mode-based strategies. Furthermore, US-OF outperformed US-APA in terms of temporal performances demonstrating an update rate up to forty times higher, in addition to not requiring raw RF data. Indeed, US-APA output rate is limited to 1-2 Hz both theoretically, by the need to run Fourier analysis, and practically, by the limited US probe acquisition rate and available computation power. On the other hand, US-OF can readily update the MR tracked location as fast as few tens of milliseconds. This maps to an output rate higher than 40 Hz, which is comparable to standard US probes frame rate. Additionally, these results were achieved without code optimization and using a standard PC, thus the computation time can furtherly be reduced. From the comparison drawn in this study, it is possible to formulate a rule of thumb for US tracking of MRs. If a precise model for target locomotion is available, US-APA is a robust approach, being a signal filtering tailored for the specific scenario. Anyway, it comes with increased computational cost and difficulties in arranging the experimental setup, as RF data are required. Conversely, US-OF stands as a general purpose alternative, more easily integrated in different applications thanks to the faster computation time, the possibility to work with every US equipment and similar accuracy at least in the two locomotion patterns under study, which are quite representative of multiple locomotion modalities.

In general, from the hardware viewpoint, most of the works dealing with US-based MR tracking use linear probes with relatively high US frequency, from 10 up to 16 MHz [14], [15], [16], [19], [20], [21], [22], [23], [24], [25], [26], [44]. Anyway, due to the frequency-depth trade off typical of US imaging, such kind of probes only allows to image targets at few centimeters from the skin surface. This might not be the case for deep vessels and tissues in the abdomen or brain. To extend MR imaging capabilities even to these scenarios, a transition to different US probes, such as convex or phased-array ones, with a lower imaging frequency and an extended field of view, is needed. Anyway, a decrease in imaging frequency also maps to a deterioration in image resolution which is a fundamental parameter when dealing with MRs due to their small dimensions. To tackle this problem, recent high-resolution US techniques such as super-resolution US imaging could be exploited, in order to improve US penetration depth while avoiding resolution

worsening [45]. In the near future, we intend to: 1) evaluate US-OF performances with different microrobots design and locomotion strategies; 2) assess US-OF robustness under different environmental conditions, as in *ex vivo* tissues; 3) integrate US-OF in a closed-loop control for a robotic visual servoing platform.

REFERENCES

[1] M. Sitti et al., "Biomedical applications of untethered mobile milli/microrobots," *Proc. IEEE*, vol. 103, no. 2, pp. 205–224, Feb. 2015.

[2] H. Ceylan, I. C. Yasa, U. Kilic, W. Hu, and M. Sitti, "Translational prospects of untethered medical microrobots," *Prog. Biomed. Eng.*, vol. 1, no. 1, 2019, Art. no. 1012002.

[3] C. K. Schmidt, M. Medina-Sánchez, R. J. Edmondson, and O. G. Schmidt, "Engineering microrobots for targeted cancer therapies from a medical perspective," *Nature Commun.*, vol. 11, no. 1, pp. 1–18, 2020.

[4] B. Wang, K. Kostarelos, B. J. Nelson, and L. Zhang, "Trends in micro-/nanorobotics: Materials development, actuation, localization, and system integration for biomedical applications," *Adv. Mater.*, vol. 33, no. 4, 2021, Art. no. 25201.

[5] T.-Y. Huang, H. Gu, and B. J. Nelson, "Increasingly intelligent micromachines," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 5, pp. 279–310, 2022.

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

[7] S. Pané et al., "Imaging technologies for biomedical micro- and nanoswimmers," *Adv. Mater. Technol.*, vol. 4, no. 4, 2019, Art. no. 1800575.

[8] S. L. Charreyron et al., "A magnetically navigated microcannula for subretinal injections," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 1, pp. 119–129, Jan. 2021.

[9] J. Hwang et al., "An electromagnetically controllable microrobotic interventional system for targeted, real-time cardiovascular intervention," *Adv. Healthcare Mater.*, vol. 11, 2022, Art. no. 2102529.

[10] A. C. Hortelao et al., "Swarming behavior and in vivo monitoring of enzymatic nanomotors within the bladder," *Sci. Robot.*, vol. 6, 2021, Art. no. 25100.

[11] V. Iacovacci et al., "High-resolution SPECT imaging of stimuli-responsive soft microrobots," *Small*, vol. 15, 2019, Art. no. 1900709.

[12] M. E. Tiryaki, S. O. Demir, and M. Sitti, "Deep learning-based 3D magnetic microrobot tracking using 2D MR images," *IEEE Robot. Automat. Lett.*, vol. 7, no. 3, pp. 6982–6989, Jul. 2022.

[13] N. Nothnagel, J. Rahmer, B. Gleich, A. Halkola, T. M. Buzug, and J. Borgert, "Steering of magnetic devices with a magnetic particle imaging system," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 11, pp. 2286–2293, Nov. 2016.

[14] F. Suligoj, C. M. Heunis, S. Mohanty, and S. Misra, "Intravascular tracking of micro-agents using medical ultrasound: Towards clinical applications," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 12, pp. 3739–3747, Dec. 2022.

[15] Q. Wang et al., "Ultrasound Doppler-guided real-time navigation of a magnetic microswarm for active endovascular delivery," *Sci. Adv.*, vol. 7, 2023, Art. no. eabe5914.

[16] S. Pane, V. Iacovacci, E. Sinibaldi, and A. Menciassi, "Real-time imaging and tracking of microrobots in tissues using ultrasound phase analysis," *Appl. Phys. Lett.*, vol. 118, 2021, Art. no. 014102.

[17] F. von Haxthausen, S. Böttger, D. Wulff, J. Hagenah, V. García-Vázquez, and S. Ipsen, "Medical robotics for ultrasound imaging: Current systems and future trends," *Curr. Robot. Rep.*, vol. 2, no. 1, pp. 55–71, Mar. 2021.

[18] S. E. Salcudean, H. Moradi, D. G. Black, and N. Navab, "Robot-assisted medical imaging: A review," *Proc. IEEE*, vol. 110, no. 7, pp. 951–967, Jul. 2022.

[19] Q. Wang, L. Yang, J. Yu, P. W. Y. Chiu, Y. P. Zheng, and L. Zhang, "Real-time magnetic navigation of a rotating colloidal microswarm under ultrasound guidance," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 12, pp. 3403–3412, Dec. 2020.

[20] X. Du et al., "Real-time navigation of an untethered miniature robot using mobile ultrasound imaging and magnetic actuation systems," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 7668–7675, Jul. 2022.

[21] Z. Yang, L. Yang, M. Zhang, N. Xia, and L. Zhang, "Ultrasound-guided wired magnetic microrobot with active steering and ejectable tip," *IEEE Trans. Ind. Electron.*, vol. 70, no. 1, pp. 614–623, Jan. 2023.

[22] K. Botros, M. Alkhatib, D. Folio, and A. Ferreira, "USMicroMagSet: Using deep learning analysis to benchmark the performance of microrobots in ultrasound images," *IEEE Robot. Autom. Lett.*, vol. 8, no. 6, pp. 3254–3261, Jun. 2023.

[23] Q. Wang, X. Du, D. Jin, and L. Zhang, "Real-time ultrasound Doppler tracking and autonomous navigation of a miniature helical robot for accelerating thrombolysis in dynamic blood flow," *ACS Nano*, vol. 16, no. 1, pp. 604–616, 2022.

[24] S. Pane, V. Iacovacci, M. H. D. Ansari, and A. Menciassi, "Dynamic tracking of a magnetic microroller using ultrasound phase analysis," *Sci. Rep.*, vol. 11, 2021, Art. no. 23239.

[25] S. Pane, G. Faoro, E. Sinibaldi, V. Iacovacci, and A. Menciassi, "Ultrasound acoustic phase analysis enables robotic visual-servoing of magnetic microrobots," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1571–1582, Jun. 2022.

[26] S. Pane, M. Zhang, V. Iacovacci, L. Zhang, and A. Menciassi, "Contrast-enhanced ultrasound tracking of helical propellers with acoustic phase analysis and comparison with color Doppler," *APL Bioeng.*, vol. 6, 2022, Art. no. 36102.

[27] C. Poelma, "Ultrasound imaging velocimetry: A review," *Exp. Fluids*, vol. 58, no. 1, 2017, Art. no. 396.

[28] S. T. H. Shah and X. Xuezhi, "Traditional and modern strategies for optical fow: An investigation," *SN Appl. Sci.*, vol. 3, 2021, Art. no. 289.

[29] G. Faoro, N. Pasini, A. Mariani, L. Morchi, S. Tognarelli, and A. Menciassi, "Robot-assisted motion compensation based on optical flow in ultrasound images," in *Proc. IEEE Int. Ultrasonics Symp.*, 2023, pp. 1–4.

[30] F. Zhang et al., "In vitro and preliminary in vivo validation of echo particle image velocimetry in carotid vascular imaging," *Ultrasound Med. Biol.*, vol. 37, no. 3, pp. 450–464, 2011.

[31] K. Sampath, T. T. Harfi, R. T. George, and J. Katz, "Optimized time-resolved echo particle image velocimetry – particle tracking velocimetry measurements elucidate blood flow in patients with left ventricular trombus," *J. Biomechanical Eng.*, vol. 140, 2018, Art. no. 236.

[32] M. Strachinaru et al., "Left ventricular high frame rate echo-particle image velocimetry: Clinical application and comparison with conventional imaging," *Cardiovasc. Ultrasound*, vol. 20, 2022, Art. no. 11.

[33] T. Williamson, W. Cheung, S. K. Roberts, and S. Chauhan, "Ultrasound-based liver tracking utilizing a hybrid template/optical flow approach," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, pp. 1605–1615, 2018.

[34] S. Golemati, A. Gastounioti, and K. S. Nikita, "Ultrasound-image-based cardiovascular tissue motion estimation," *IEEE Rev. Biomed. Eng.*, vol. 9, pp. 208–218, 2016.

[35] E. Karami, M. S. Shehata, and A. Smith, "Tracking of the internal jugular vein in ultrasound images using optical flow," in *Proc. IEEE 30th Can. Conf. Elect. Comput. Eng.*, 2017, pp. 1–4.

[36] A. Cafarelli, A. Verbeni, A. Poliziani, P. Dario, A. Menciassi, and L. Ricotti, "Tuning acoustic and mechanical properties of materials for ultrasound phantoms and smart substrates for cell cultures," *Acta Biomaterialia*, vol. 49, pp. 368–378, 2017.

[37] G. Faoro, S. Maglio, S. Pane, V. Iacovacci, and A. Menciassi, "An artificial intelligence-aided robotic platform for ultrasound-guided transcarotid revascularization," *IEEE Robot. Automat. Lett.*, vol. 8, no. 4, pp. 2349–2356, Apr. 2023.

[38] N. Ebrahimi et al., "Magnetic actuation methods in bio/soft robotics," *Adv. Funct. Mater.*, vol. 31, 2021, Art. no. 11.

[39] 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. Automat. Lett.*, vol. 5, no. 2, pp. 705–712, Apr. 2020.

[40] X. Du and J. Yu, "Image-integrated magnetic actuation systems for localization and remote actuation of medical miniature robots: A survey," *IEEE Trans. Robot.*, vol. 39, no. 4, pp. 2549–2568, Aug. 2023.

[41] M. H. D. Ansari et al., "3D printing of small-scale soft robots with programmable magnetization," *Adv. Funct. Mater.*, vol. 33, no. 15, 2023, Art. no. 2211918.

[42] B. Horn and B. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 7, pp. 185–203, 1981.

[43] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. DARPA Image Understanding Workshop*, 1981, pp. 674–679.

[44] Q. Chen, F. W. Liu, Z. Xiao, N. Sharma, S. K. Cho, and K. Kim, "Ultrasound tracking of the acoustically actuated microswimmer," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 3231–3237, Nov. 2019.

[45] K. Christensen-Jeffries et al., "Super-resolution ultrasound imaging," *Ultrasound Med. Biol.*, vol. 46, no. 4, pp. 865–891, 2020.