Ultrasound Acoustic Phase Analysis Enables Robotic Visual-Servoing of Magnetic Microrobots

Stefano Pane, *Student Member, IEEE*, Giovanni Faoro, *Student Member, IEEE*, Edoardo Sinibaldi, *Member, IEEE*, Veronica Iacovacci, *Member, IEEE*, and Arianna Menciassi, *Senior Member, IEEE*

***Abstract*—Microrobots (MRs) have attracted growing interest for their potentialities in diagnosis and noninvasive intervention in hard-to-reach body areas. The safe operation of biomedical MRs requires fine control capabilities, which strongly depend on precise and robust feedback about their position over time. Ultrasound acoustic phase analysis (US-APA) may allow for a reliable feedback strategy for MR imaging and tracking in tissue. In this article, we combine task-specific magnetic actuation and related US-APA motion tracking to achieve closed-loop navigation of a magnetic MR, rolling on the boundary of a lumen in a tissue-mimicking phantom. A C-arm system attached to a robotic platform is used to precisely position the magnetic actuation source and US-APA detection unit within the workspace, thus enabling MR visual-servoing. In the first place, the proposed approach allows to perform supervised localization of the MR without any *a-priori* knowledge of its position. After localization, a robust real-time tracking enables closed-loop MR teleoperation in the phantom lumina over a travel distance of 80 mm (145 body lengths), both in static and counter flow, thus achieving an average position tracking error of 368 micron (0.67 body lengths). For the first time, our results validate US-APA as a reliable feedback strategy for visual-servoing control of MRs in simulated in-body environment.**

***Index Terms*—Acoustic phase analysis (APA), closed-loop control, magnetic actuation, medical microrobots (MRs), ultrasound (US) imaging, visual-servoing.**

I. Introduction

Microrobots (MRs) for biomedical applications hold the potential to revolutionize diagnosis and therapy, thanks to their ability to access and operate in hard-to-reach body districts [1]–[6]. Different strategies have been proposed to remotely navigate and maneuver microscale agents in confined and unstructured environments [7], among which the most popular employ ultrasound (US) [8], light [9], and magnetic fields [10], [11]. The last solution is arguably the most promising due to high controllability and deep penetration of magnetic fields without tissue attenuation, distortion, or harm to the patient. Several concepts of magnetic MRs have been proposed for medical applications [12]. Among these, surface microrollers look particularly promising due to their ability to navigate on the boundaries of body lumina, possibly against physiological flow [13], [14].

Most of the advancements witnessed in magnetic MRs control [15] have been allowed by optical microscopy, which provides precise real-time feedback on MRs position over time and on responses to triggering signals. Optical microscopy also served as enabling technology for implementing visual-servoing strategies in transparent media or tissue such as the eye, both *in-vitro* [16]–[20] and *ex-vivo* [21]. However, obtaining similar control performance in nontransparent deep tissue (where optical microscopy fails) is still an open challenge [22]. Biomedical imaging strategies should be used to obtain feedback on MRs states when moving inside the body. Different imaging techniques have been considered for this purpose, including both traditional techniques (e.g., magnetic resonance imaging [23] or single-photon emission computed tomography [24]) and innovative ones, such as photoacoustic [25] or magnetic particle imaging [26]. However, replicating optical microscopy contrast and spatial resolution while providing real-time imaging at high penetration depths is not straightforward [27]. In this scenario, US imaging stands as a mature technique, combining real-time imaging, low cost of the equipment, and no harm to the patient. Considering the frequency/power tradeoff characterizing traditional US imaging systems, the image spatial resolution can be improved by using high-frequency waves while reducing the imaging depth. Overall, commercial US probes provide acceptable spatial resolution ($100 - 500\ \mu\mathrm{m}$) at clinically relevant penetration depths ($2 - 30$ cm). Nevertheless, standard US imaging modalities, such as brightness (B)-mode, feature poor contrast resolution. Some attempts to perform visual-servoing of magnetic MRs with US B-mode images were reported in controlled experimental conditions (e.g., inhomogeneous and transparent media), specifically set to enhance MR contrast and

Manuscript received September 15, 2021; revised November 21, 2021; accepted January 6, 2022. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie under Grant 894425. This paper was recommended for publication by Associate Editor J. Abbott and Editor M. Yim upon evaluation of the reviewers'
Q1 comments. *(Corresponding author: Stefano Pane.)*

Stefano Pane, Giovanni Faoro, and Arianna Menciassi are with the
Q2 BioRobotics Institute of Scuola Superiore Sant'Anna, 56127 Pisa, Italy, and
Q3 also with the Department of Excellence in Robotics and AI of Scuola Superiore Sant'Anna, 56127 Pisa, Italy (e-mail: stefa.pane@gmail.com; gio-
Q4 vanni.faoro@santannapisa.it; arianna@sssup.it).

Edoardo Sinibaldi is with the Istituto Italiano di Tecnologia, 16163 Genova, Italy (e-mail: edoardo.sinibaldi@iit.it).

Veronica Iacovacci is with the BioRobotics Institute of Scuola Superiore Sant'Anna, 56127 Pisa, Italy, with the Department of Excellence in Robotics and AI of Scuola Superiore Sant'Anna, 56127 Pisa, Italy, and also with the Department of Mechanical and Automation Engineering, Chinese University of Hong Kong, Hong Kong (e-mail: veronica.iacovacci@santannapisa.it).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TRO.2022.3143072.

Digital Object Identifier 10.1109/TRO.2022.3143072

facilitate its detection and localization [28], [29]. On the other hand, MR visibility with B-mode imaging in biological tissue is limited, being tissue heterogeneous and thus highly reflective to US waves. The high-contrast imaging artifacts produced by acoustic impedance discontinuities (e.g., lumen boundaries, interfaces, bones, air sacs, etc.) introduce disturbances and instability in MR tracking, compromising the use of US B-mode as a feedback strategy for closed-loop control [30].

Recently, US Doppler techniques have been proposed to improve MRs visibility in biological media [31], [32]. US acoustic phase analysis (US-APA) has been proposed for MRs motion detection toward higher contrast resolution even in highly reflective and dynamic media [33], [34]. Although very promising, these techniques still require research efforts to be implemented in control scenarios, both to match real-time performance and to become robust enough against environmental disturbances.

In this article, we designed and implemented a control framework to exploit the US-APA detection technique in a stable closed-loop control system, and we developed a robotic platform for US-APA-enabled MR visual-servoing. More specifically, we contributed to the following advancements: 1) we designed a C-arm system for combining magnetic actuation with US-APA detection in a compact robotic platform; 2) we developed a real-time tracker based on the US-APA image processing technique, able to function both in open-loop and in closed-loop with the main robot controller; 3) we designed and implemented control features, such as the reduced search window (RSW) with adaptive size, which help minimize the US-APA's computational cost for smooth real-time operation; 4) we defined and characterized control signals for interfacing the robot controller and the US-APA tracker while ensuring feedback integrity and control stability; and 5) we designed and developed the main robot controller, able to function stably both in manual teleoperation of the robot arm, and in visual-servoing mode for closed-loop MR control. The developed platform allowed for controlled navigation of a magnetic MR in the lumina of a tissue-mimicking phantom. First, we evaluated the performances of the platform during supervised localization of the MR in the phantom, i.e., without any *a-priori* knowledge of the MR position. We then performed controlled MR teleoperation in the lumina, both in static flow and counter flow conditions.

The remainder of this article is organized as follows: Section II describes the analytical models that combine magnetic manipulation with US-APA detection for building a synergic actuation-feedback strategy. Section III provides an overview of the robotic platform enabling MR visual-servoing. Section IV is dedicated to the closed-loop control architecture. Section V presents the experimental setup used to evaluate the platform performances, whereas Section VI reports and discusses experimental results. Finally, Section VII remarks on the conclusions and the future perspectives stemming from this article.

II. Models for Magnetic Actuation and Acoustic Phase Modulation

This section presents the fundamental magnetic and acoustic principles that allow the development of a closed-loop control strategy for visual-servoing magnetic MRs in simulated in-body environments (i.e., closer to realistic bodily environments and farther from ideal lab conditions).

A. Magnetic Actuation

When subject to an external magnetic field B, a magnetic dipole with moment m experiences both torques and forces according to the following equations [35]:

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

For navigating the MR inside lumina, we exploited both the magnetic force $\boldsymbol{F}$ generated by the field gradient and the magnetic torque τ produced to align the MR magnetic moment $\boldsymbol{m}$ with $\boldsymbol{B}$. The magnetic force $\boldsymbol{F}$, pulling toward the magnet, systematically keeps the MR on the lumen boundary. The assumption of laminar flow (supported by the small lumen diameter and the relatively low flow rates) enables reduced drag force and adherence to the boundary surface to favor controlled locomotion. Once migrated at the border, the MR experiences a magnetic torque τ produced by a rotating/vibrating magnetic field $\boldsymbol{B}$. The employed MR (Section IV-B) can assume two states, based on the type of magnetic actuation mode: *idle state* when the MR experiences a vibrating magnetic field and *locomotion state* when a rotating magnetic field is produced (Fig. 1). We formulated the magnetic manipulation problem by considering the two components of the field B belonging to the US imaging plane, namely B_x and B_y. To produce the *idle state*, B_x and B_y should have the following time evolution:

$$\begin{cases} B_x \cong \bar{\theta} \, |B| \sin \left(2\pi f_{\mathrm{vib}} t \right) \\ B_y \cong |B| \, . \end{cases} \tag{2}$$

The field B oscillates in the xy plane at the frequency f_{vib} over a circular sector defined by a maximum angle $\bar{\theta}$ (assumed to be small). The resulting magnetic torque produces small harmonic rolling motions of the cylindrical MR. These motions do not produce locomotion but generate in-place vibrations, along the y axis, of the extreme points of the MR diameter parallel to the boundary surface [Fig. 1(a)]. The induced vibrations have amplitude $\bar{\theta}$ and frequency f_{vib}.

Otherwise, when in *locomotion state*, the MR is actuated with the following field sequence:

$$\begin{cases} B_x = |B| \; \sin \left(2\pi f_{\mathrm{rot}} t \right) \\ B_y = |B| \; \cos \left(2\pi f_{\mathrm{rot}} t \right). \end{cases} \tag{3}$$

The resulting field B rotates in the xy plane with frequency f_{rot}. The generated magnetic torque produces a continuous rotation of the MR, activating rolling locomotion on the boundary surface [Fig. 1(b)].

B. Motion-Induced Acoustic Phase Modulation

In pulse-echo US imaging, mechanical waves are emitted by a transducer, typically made of an array of piezoelectric elements. When a mechanical wave encounters an object (reflector), part of the wave is reflected to the transducer, in the form of an echo

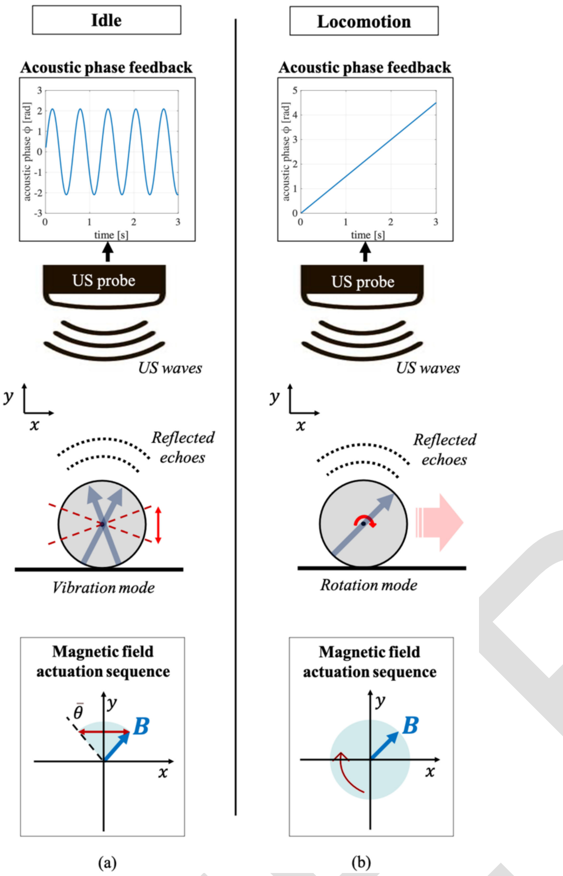


Fig. 1. Magnetic actuation sequences, MR states, and associated US acoustic phase feedbacks. (a) When in idle state, an oscillating magnetic field produces in-place vibrations of the microrobot and a consequent sinusoidal acoustic phase feedback. (b) When in locomotion state, a rotating magnetic field produces rolling locomotion and a linear modulation of the acoustic phase feedback.

signal $E(t)$:

$$E\ (t) = \ A\,(t)\,e^{j\varphi(t)}. \tag{4}$$

The instantaneous amplitude $A(t)$ of the echo signal is related to the acoustic impedance mismatch between the reflector and the surrounding medium [36]. In B-mode imaging, $A(t)$ is converted into grayscale levels to produce a contrast image. Biological tissues are highly reflective due to their heterogeneity. This implies that when considering one or more MRs in biological tissues, the echoes reflected by the tissues result higher than those reflected by the MRs (especially when the MR is close to an interface like the lumen boundary). This makes standard B-mode imaging unsuitable for automatized MR visualization and localization. Conversely, a reliable US-based technique for MR tracking should be robust to high contrast backgrounds. An emerging and promising strategy for enhancing MR contrast in reflective backgrounds is the US-APA motion detection. As a result of the Doppler effect, the echoes reflected by moving objects are shifted in phase with respect to the incident waves. If a wave propagates along the y direction with wavelength λ and encounters a moving reflector, the acoustic phase shift $\Delta\varphi$ in successively reflected echoes is proportional to the reflector's displacement along the direction of wave propagation Δu_y:

$$\Delta\varphi\ = \frac{4\pi}{\lambda}\,\Delta u_y. \tag{5}$$

In analogy, the acoustic frequency f, given by the time derivative of the acoustic phase, is shifted proportionally to the reflector's velocity v_y:

$$\frac{d\,(\varphi)}{dt} = \ f\ = \frac{4\pi}{\lambda}\,v_y. \tag{6}$$

According to (5) and (6), MR displacements (and respective velocity) are modulated on the acoustic phase (and respective frequency) of the received echo signals. By performing US-APA, this phenomenon can be exploited to detect MRs, even when the associated echoes feature low amplitude with respect to the background medium.

In this framework, producing magnetic field sequences has a double value: driving MR locomotion and producing acoustic phase feedback for localization and control (the feedback is meaningless when the MR is completely still). In particular, the vibration field sequence (2), associated with the MR idle state, induces an in-place vibrating motion pattern and a harmonic modulation of the acoustic phase signal [Fig. 1(a)], which can be detected through Fourier analysis [33]. Alternatively, the rotation field sequence (3), associated with the MR locomotion state, induces a rotating motion pattern and a linear modulation of the acoustic phase [Fig. 1(b)], resulting in a constant modulation of the acoustic frequency that can be detected through block-matching analysis [34].

Regarding the dynamics of the magnetic microroller, we identify two main contributions: 1) the dynamics of the rotational/vibrational MR motions in response to the magnetic torque generated by the driving magnetic field signals; and 2) the dynamics of the resulting translational rolling motion, thus the dynamics of the interaction with the environment (i.e., the fluid-filled lumen). In this article, we assume that the dynamics 1) can be neglected, considering the frequency range of the driving signals (1–5 Hz). The dynamics 2), which represent the major contributions, are characterized by many case-specific variables, including 1) the actual direction of the MR magnetic dipole moment, 2) the actual MR morphology, 3) the actual lumen morphology, 4) the precise static and dynamic friction coefficients between the MR and the lumen boundary surface, 5) the precise drag coefficients, and 6) the flow intensity and direction. In this article, we considered the deviations with respect to the ideal pure rolling behavior, due to the dynamics *B*, as disturbances to the control system. These assumptions on the MR dynamics are finally validated in the experimental section (Section VI).

Overall, the proposed combination of remote magnetic actuation with specific motion detection (through US-APA) provides a powerful strategy for US-guided visual-servoing of MRs inside the body.

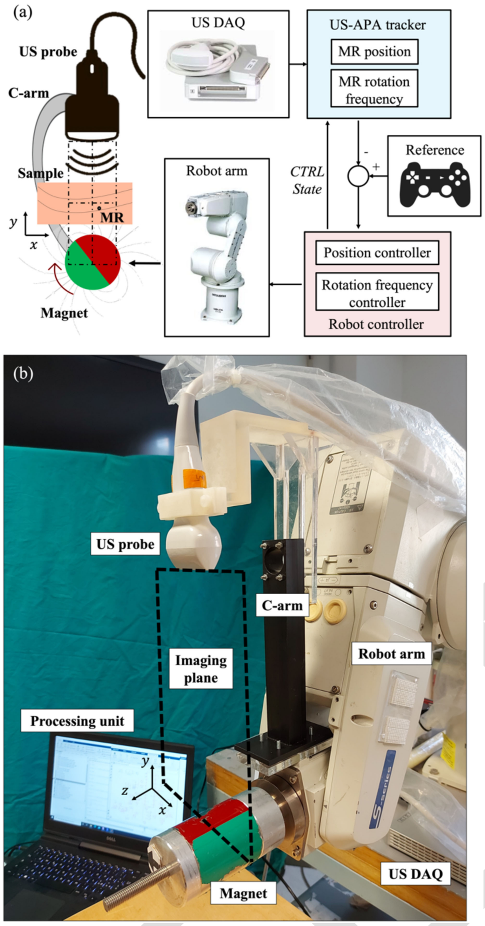


Fig. 2. Visual-servoing platform overview. (a) Schematic representation of the control scheme allowing MR teleoperation in closed feedback loop with the US-APA tracker. (b) Visual-servoing platform picture. The C-arm system prototype is integral with the robot arm and holds both the US probe (acoustic detection unit) and the magnet (magnetic actuation unit). A human–machine interface (joystick) allows to set the reference signals for teleoperation. The US DAQ and the processing unit enable data acquisition and processing, as well as closed-loop control implementation.

III. Robotic Visual-Servoing Platform Overview

For teleoperating the magnetic MR in lumina, we propose a robotic visual-servoing platform [Fig. 2(a)]. To combine magnetic manipulation with US-APA feedback, the platform integrates a magnetic actuation unit and an acoustic detection unit, arranged at the ends of a C-arm system integrated into a six-DOF robotic arm [Fig. 2(b)]. The distance between the two units can be manually regulated ($10 - 30$ cm) according to the desired workspace. The magnetic actuation unit consists of a mobile cylindrical permanent magnet (6 cm in diameter, 7 cm in height, NdFeB, diametral magnetization, and grade N35). The acoustic detection unit consists of a linear US probe (L15-7H40, Telemed, Lithuania) connected to an open architecture digital acquisition board (DAQ) (ArtUS, Telemed, Lithuania), which provides access to the raw radio-frequency (RF) data for customized analysis.

Having defined the acoustic axis of an US transducer as the direction of waves propagation, the spatial alignment of all acoustic axes of the employed probe's piezoelectric elements defines the imaging plane. On the other hand, the optimal magnetic manipulation plane is defined as the plane perpendicular to the cylinder axis and passing through its center. By design, the C-arm arrangement makes the imaging plane and the optimal magnetic manipulation plane coincide [dashed plane in Fig. 2(b)]. In such plane, we define the optimal manipulation point as the imaging plane center (i.e., the US probe focus point). The C-arm is attached to a six-DOF anthropomorphic robot arm with a spherical wrist (Melfa RV-3SB, Mitsubishi, Japan). The robot arm allows to precisely position the optimal manipulation point in space to systematically coincide with the estimated MR position and to rotate the permanent magnet for magnetic MR actuation. A closed-loop control architecture enables robotic MR visual-servoing within the C-arm workspace. The system main control parameters and building blocks are described in detail in the following sections.

IV. Control Architecture

In the proposed control architecture, the robot controller remotely actuates the magnetic MR and is connected in a closed feedback loop to the US-APA tracker [Fig. 2(a)]. The tracker sends information about the estimated MR state (i.e., position and rotation frequency) to the robot controller through a bilateral used datagram protocol communication. In turn, the controller sends some control state variables to the tracker for safe closed-loop implementation (Section IV-C). A human–machine interface (joystick) allows the user to teleoperate the robot arm (and consequently to move the imaging plane for MR localization) when in open-loop mode, and to provide reference MR rotation frequency values when in closed-loop mode. This grants flexible control and adaptivity to different task requirements. The implementation of the robot controller and the US-APA tracker are described in detail in Section V.

A. Robot Controller

The robot controller runs at a frequency of 100 Hz and is designed to function both in open-loop and closed-loop modes. When operating in open-loop mode, the controller does not receive inputs from the US-APA tracker and allows the user to teleoperate the robot arm for positioning the end-effector in the Cartesian space. When operating in closed-loop mode, the controller acquires feedback on MR states from the US-APA tracker and provides the control action for simultaneously actuating the MR and positioning the C-arm system. This control action is elaborated by two independent proportional subcontrollers (each associated with a control loop), namely the C-arm position controller and the magnet rotation frequency controller [Fig. 3(a)]. The C-arm position controller solves the problem of positioning the C-arm optimal manipulation point to minimize the mismatch with the estimated MR position, without accounting for the rotation of the magnet. In this way, the MR is systematically positioned in the imaging plane center for both optimal manipulation and detection. On the other hand, the

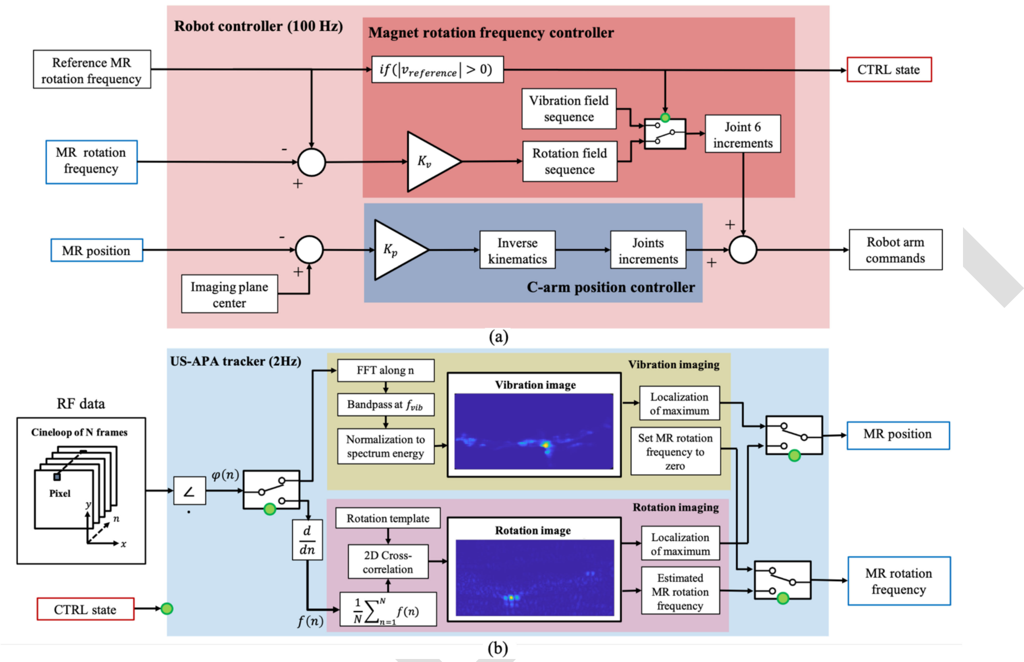


Fig. 3. Schematic representation of the robot controller and US-APA tracker developed for closed-loop MR visual-servoing. (a) Robot controller acquires MR position and rotation frequency data from the tracker and a reference MR rotation frequency from the teleoperation device (running control frequency is 100 Hz). In the upper loop (magnet rotation frequency controller), if the reference frequency is different from 0, rotation field sequence is implemented to rotate the MR and track the reference value. CTRL state is set to 1. Otherwise, vibration field sequence is implemented, and CTRL state is 0. In the lower loop (C-arm position controller), the C-arm positioning is adjusted to match the acquired MR position with the imaging plane center (optimal manipulation point). At every control iteration, the incremental joint commands, independently elaborated in the two loops, are combined and sent to the robot arm for motion control. (b) US-APA tracker acquires RF data from the DAQ and the value of CTRL state from the robot controller (running tracking frequency is 2 Hz). Based on CTRL state, the vibration imaging or rotation imaging tracking algorithm is executed. MR position and rotation frequency are estimated and provided as output. If vibration imaging is being performed, the estimated MR rotation frequency is set to zero.

magnet rotation frequency controller is designed to elaborate the trajectories of the sixth joint for rotating the permanent magnet and actuating the MR based on the driving field sequence requested by the user. This subcontroller also informs the tracker about the implemented field sequence (i.e., the MR state) through a binary control variable (CTRL state). In case the user-provided MR reference rotation frequency is null (i.e., the MR is required to be in idle state), the controller provides a harmonic rotation trajectory by processing predefined joint increments defined according to (2) (vibration field sequence). Otherwise, if the reference frequency is different from 0 (i.e., the MR is required to be in locomotion state), the controller provides a continuous rotation trajectory defined according to (3) (rotation field sequence). The rotating field frequency f_{rot} is regulated to minimize the mismatch between the estimated MR rotation frequency and the user-provided reference value (in the range $0 - 1.5$ Hz, bounded by the robot arm capabilities). In locomotion state, CTRL state is set to 1, otherwise it is set to 0.

This implementation allows simultaneous MR actuation and C-arm positioning during MR locomotion by solving the two problems independently and summing the solutions in terms of joints increments at every control iteration. This control strategy is robust since it allows MR position control and tracking during all states of a navigation task by relying on the continuous US-APA feedback. Furthermore, it is flexible because it allows the user to teleoperate the MR rotation frequency according to the task requirements.

B. US-APA Tracker

The US-APA tracker acquires RF data from the DAQ board and provides real-time feedback on MR position and rotation frequency to the robot controller [Fig. 3(b)].

Depending on the imposed MR state (either idle or locomotion, depending on CTRL state), the tracker implements two different tracking algorithms: *vibration imaging* and *rotation imaging*. The two algorithms are alternatively enabled by a tracking mode selector associated with the CRTL state variable [green bullet in Fig. 3(b)]: when MR state is *idle* (CRTL state = 0), the selector is in *vibration imaging*. In this modality, the MR position is identified by detecting magnetically induced microvibrations that produce a harmonic modulation of the acoustic phase (5) [33]. An ensemble of N RF data frames, namely cineloop, is acquired from the DAQ board and the acoustic phase signal is extracted. For each pixel in the imaging plane, the Fourier transform of the acoustic phase is computed, and a bandpass filter is applied to isolate the spectrum component relative to the actuation frequency f_{vib}. The value of such component is normalized to the spectrum energy to reduce noise effects and is mapped into the pixel intensity values. In this way, a vibration image is composed, representing the intensity of vibrations at the magnetic actuation frequency f_{vib}. Assuming that the MR is the only element vibrating at f_{vib}, this processing enables a significant MR contrast enhancement with respect to the background. Localizing the maximum in the vibration image provides the estimated MR position. In this modality, the output MR rotation frequency is set to zero since the MR is idle in place. Alternatively, when the MR is in *locomotion state* (CRTL state $= 1$), and the selector is in *rotation imaging*, the MR position and rotation frequency are estimated by analyzing the echoes frequency signal, which is modulated by the MR rotations (6) [34]. Here, the acoustic phase is time-derived to obtain the acoustic frequency. For each pixel, the acoustic frequency is averaged to obtain an image representation of the mean velocity distribution in the imaging plane. This image is cross-correlated with a template representing the velocity distribution of the rotating MR (known *a-priori*) to produce a rotation image. The maximum in the rotation image provides the estimated MR position, while the MR rotation frequency is estimated, in this case, from the measured mean acoustic frequency. This dual-mode US-APA tracker allows for continuous real-time feedback on MR position and rotation frequency throughout the different states assumed by the MR during navigation.

C. Closed-Loop Operation of the US-APA Tracker

To avoid instability due to tracking failure, the proposed tracker operates in two macrostates [Fig. 4(a)]: *supervised search* (default when the application is started) and *tracking*. In *supervised search* (open-loop), MR localization is performed over a relatively large search window since no *a-priori* knowledge on MR position is available to the operator. The tracker does not send data to the robot controller in this modality, and the control loop is open. The estimated MR position is screen printed and overlapped with B-mode images on a custom-developed GUI (Visual Studio 2019). The operator can move the imaging plane in space by teleoperating the robot arm while monitoring the GUI and compare morphological information provided by B-mode images (e.g., high contrast lumen boundaries) with the estimated MR centroid position provided by the US-APA tracker. The GUI also allows configuring the imaging and tracking parameters in real time (e.g., imaging depth, window size, and tracking frame rate). Once the MR has been localized in *supervised search* (e.g., by assessing stable estimation of MR centroid position) and the C-arm has been positioned accordingly by robot arm teleoperation, the operator can manually switch to *tracking* mode (closed-loop) through the GUI controls. In this modality, MR tracking is performed on an RSW around the optimal manipulation point, and the estimated MR position and rotation frequency are sent to the robot controller for visual-servoing. The operator can still monitor the US images through the GUI and manually switch back to *supervised search* at any time.

For the benefit of time efficiency, which is a strict requirement for feedback control applications, the tracker was designed to perform data acquisition and processing in parallel (by exploiting multithreading). To this purpose, the $i\mathrm{th}$ cineloop is processed while the $(i+1)\mathrm{th}$ cineloop is being acquired [Fig. 4(b)].

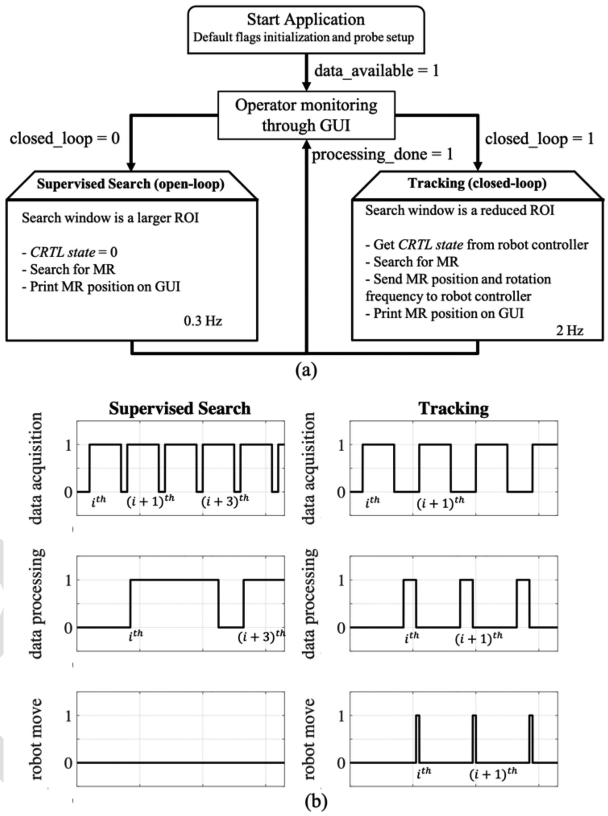


Fig. 4. Closed-loop operation of the US-APA tracker. (a) Schematic finite state machine representation of the tracker operation modes. When in the supervised search state, MR localization is performed over a larger search window and the tracker operates in open loop. When in tracking state, the search window is reduced, and the frame rate is increased for closed-loop operation. (b) Qualitative time evolution of the Boolean state variables "data acquisition," "data processing," and "robot move" in the two operating modes. In supervised search mode (left panel), "data processing" relative to the $i\mathrm{th}$ cineloop is performed when "data acquisition" relative to the $(i+1)\mathrm{th}$ cineloop is started and "robot move" is always set to zero. In tracking mode (right panel), to avoid motion-induced RF data distortion, "data acquisition" of the $(i+1)\mathrm{th}$ cineloop is started on the falling edge of the "robot move" relative to the $\mathbf{\mathit{i}th}$ cineloop.

In *supervised search* mode, the number of pixels on which to perform Fourier analysis is relatively high. Data processing takes longer than data acquisition [left panel of Fig. 4(b)], thus causing data losses and a low localization frame rate (0.3 Hz). However, in this case, the real-time requirements can be relaxed due to the open-loop operation mode.

In *tracking* mode, data processing is faster than data acquisition due to the RSW, resulting in no data loss and a higher localization frame rate [right panel of Fig. 4(b)]. The RSW size was selected based on two optimum criteria: 1) the RSW size must be minimized to reduce the computational cost of the US-APA processing; and 2) the RSW size must be large enough to prevent the MR from escaping the search area due to sudden motions during the localization time. To fulfill both these criteria, the RSW size should be selected based on the expected MR velocity, thus on the tracker frame rate f_{trk}, the MR radius r_{MR}, and its rotation frequency f_{rot}. For simplicity, we refer to a squared RSW, in which side L_{RSW} is defined as

$$L_{\mathrm{RSW}} = 2k\frac{1}{f_{\mathrm{trk}}}2\pi r_{\mathrm{MR}} f_{\mathrm{rot}}. \tag{7}$$

The quantity k is an arbitrary gain factor that can be chosen for ensuring stability. In fact, the stability of the control system is directly related to the RSW size. More specifically, if a disturbance is large enough to push the MR out of the search window (e.g., in presence of large counter flow), the tracking system could fail, leading to control loop instability. To ensure stability, potential case-specific disturbances must be investigated, and an appropriate search window size must be selected accordingly by tuning the value of k. In this article, we considered $k = 1.2$, a tracking framerate of 2 Hz, an MR diameter of 550 μm, and a rotation frequency of 1.5 Hz. With these conditions, the RSW side is 6 mm.

In *tracking* mode, since the robot moves the imaging plane in a closed-loop fashion, the tracker was implemented to start new data acquisition only after the robot arm has communicated motion completion. This workflow avoids distortion caused by simultaneous acquisition of RF data and motion of the imaging plane. In *tracking* mode, the frame rate primarily depends on the cineloop dimension *N*. Specifically, the tracking frame rate increases when *N* is decreased. To guarantee good quality of the Fourier analysis, it must hold that $N \geq 2\frac{\mathrm{fps}_{\mathrm{US}}}{f_{\mathrm{vib}}}$, where $\mathrm{fps}_{\mathrm{US}}$ is the frame rate of the US DAQ board (120 fps). Given a fixed f_{vib}, the tracking frame rate is defined as

$$f_{\mathrm{TR}} = \frac{\mathrm{fps}_{\mathrm{US}}}{N} = \frac{f_{\mathrm{vib}}}{2}. \tag{8}$$

In this article, the achievable f_{vib} was bounded to a maximum of 4 Hz by the robot arm capabilities, resulting in an f_{TR} of 2 Hz. However, higher MR vibration frequencies could enable higher tracking frame rates.

V. Experimental Setup

This section provides an overview of the experimental setup employed to evaluate the performances of the proposed visual-servoing system.

A. Experimental Validation Platform

To validate the system in a simulated in-body environment, we built an experimental platform to teleoperate the magnetic MR through the lumen of a tissue-mimicking phantom filled with a blood-mimicking fluid. The phantom was positioned in a water tank with an acoustic absorber on its bottom to ensure optimal acoustic coupling (Fig. 5). The tank was placed on a fixed support, and the robot was positioned so that the C-arm optimal manipulation point fell within the phantom, with the US detection unit on top and the magnetic actuation unit under the support. The C-arm width was regulated so that the magnet was positioned approximately $10\ \mathrm{cm}$ far away from the optimal manipulation point, thus testing working distances compatible with a future clinical translation. A micropump (M100S, TCS micropumps, U.K.) was used to pump the blood-mimicking fluid through the lumen. The induced flow allowed to demonstrate the robustness of the proposed visual-servoing approach in dynamic media. In this article, the flow rate was limited to a value of 3 mL/s with the aim of implementing magnetic MR actuation with relatively weak magnetic fields (in the order of 10 mT).

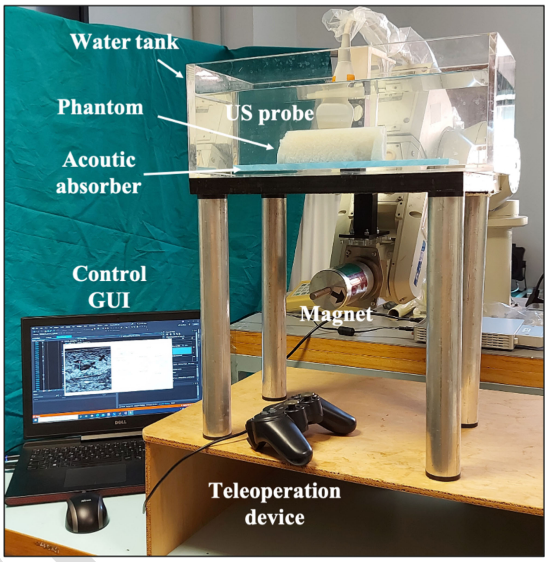


Fig. 5. Experimental platform for system performance evaluation. The MR is placed in the phantom lumen, immersed in a water tank with acoustic absorber. A teleoperation device allows to move the robot arm in open-loop mode and to provide reference MR angular velocity for visual-servoing in closed-loop mode. A control GUI allows for real-time monitoring, adjusting tracking parameters and collecting experimental data.

B. MR Fabrication

We required a magnetic surface microroller to perform controlled rolling along the internal lumen surface through external magnetic fields. For this purpose, we fabricated a submillimeter cylindrical MR with remanent magnetization along the radial direction. We employed extrusion-based printing of a UV curable magnetic ink. To achieve uniform radial magnetization, the printed cylindrical string was magnetized radially by an impulse magnetizer with a peak field intensity of $1.8\ T$ (T-Series, Magnet-Physik Dr. Steingroever GmbH, Germany). The final size of the cylindrical MR was defined by cutting the magnetized string into smaller segments, having length $990\ \mu\mathrm{m}$ and diameter $550\ \mu\mathrm{m}$.

C. Tissue-Mimicking Phantom and Fluid Preparation

The phantom was designed to mimic a tract of a medium artery ($3 - 4\ \mathrm{mm}$ in diameter) with the surrounding soft tissue, in terms of size and acoustic properties. The tissue-mimicking phantom was devised to simulate possible heterogeneity and high contrast regions in actual human soft tissues. To this aim, agarose was doped with soy milk used as a scatter-enhancing agent. Agarose powder (Sigma-Aldrich) was dissolved in a deionized and degassed water (dd-H_2O)—soy milk (5% v/v) solution and kept at $90^{\circ}\mathrm{C}$ for $1\ \mathrm{h}$ under continuous stirring.

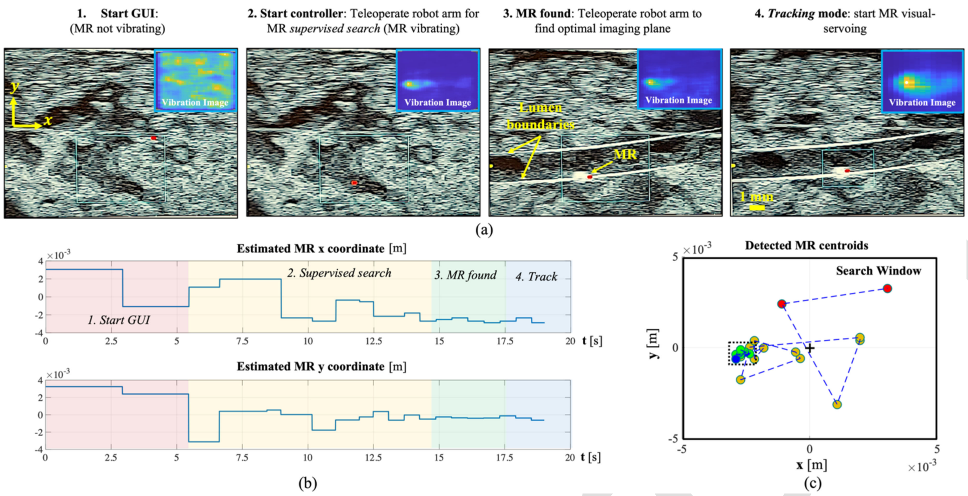


Fig. 6. Supervised MR search experiments. (a) Image acquired from the tracker GUI during the supervised search experiments. The light blue box represents the search window, the red dot represents estimated MR centroid, the yellow dot on the left image border represents US focus vertical position, and image insets show the vibration images. (1) GUI was started, the vibration image appeared noisy as no vibrations were detected. (2) Once activated the idle state control, the MR started vibrating and the imaging plane was swiped around the phantom while monitoring the plotted estimated centroid. When the vibrations appeared in the imaging plane, the MR featured high contrast in the vibration image and the estimated centroid (red dot) was stably plotted in the same position as an index of successful localization. (3) Robot arm was teleoperated to identify the best imaging plane. Lumen boundaries could be identified as continuous parallel white lines. (4) Operation modality was switched to tracking for MR visual-servoing and the search window size was reduced. (b) Time evolution of estimated MR centroid x and y coordinates with respect to the search window center. The values fluctuated during supervised search and became stable after successful MR localization. (c) Collection of all detected MR centroids within the search window during supervised search (20 s). The dot colors are associated to the respective phases 1, 2, 3, and 4 of panel (b). The distribution concentrates in the same region (black dashed box) after successful localization. During these experiments, the tracker operated in open loop and the robot was not moved automatically to match detected MR position with the search window center.

The proper agarose concentration (2% v/v) produces mechanical and acoustic properties that mimic human tissues [37]. A preshaped $3\,\mathrm{mm}$ diameter rubber tube was embedded in the phantom before reticulation. Physical reticulation occurred at room temperature in the target mold ($4.5 \times 4.5 \times 20\,\mathrm{cm}^3$). After reticulation, the tube was removed to generate the desired lumen in the phantom.

A fluid that mimics the blood in terms of viscosity and acoustic properties was obtained from an aqueous glycerol solution (60% v/v) [38].

VI. Experimental Validation

We conducted experimental validation to assess the system performance in terms of MR localization efficiency, tracking accuracy, and robustness to environmental disturbances (e.g., the presence of obstacles and counter fluid flow). In this section, the results of the experiments are presented and discussed.

A. Supervised MR Search in Echogenic Phantom

As a first step, we conducted supervised MR search experiments to validate the ability of the system to help the operator in localizing the MR within the phantom without any *a-priori* knowledge about its position. The experiment started by running the US-APA tracker in *supervised search* mode (default) [Fig. 6(a, 1)]. The MR was then put into vibrations by starting the robot controller to enable detection. In this phase, the robot arm was teleoperated for swiping the imaging plane along the phantom to search for the MR by comparing morphological information of B-mode images with MR specific information of the US-APA tracker. More specifically, by assessing the stable estimation of MR centroid position [red dot in Fig. 6(a)], the operator could verify MR localization within the lumen (Supplementary Video, part 1). In this open-loop tracking modality, the robot is not moved automatically to match the detected MR position with the optimal manipulation point (search window center).

The US-APA tracker helped find the MR even in suboptimal imaging conditions, e.g., when the imaging plane did not cross the lumen section completely [Fig. 6(a, 2)]. After finding the MR, the operator could teleoperate the robot arm to identify the optimal imaging plane crossing the lumen section [Fig. 6(a, 3)] by visual inspection of B-mode image features (e.g., the presence of reflective lumen boundaries). Once the optimal imaging conditions were identified, the *tracking* mode was activated [Fig. 6(a, 4)] for MR visual-servoing (closed loop). The time evolution of estimated MR x and y coordinates with respect to the center of the search window (optimal manipulation point) is reported in Fig. 6(b). The higher fluctuations were identified in phase 1, when the MR was not vibrating, and in the first part of phase 2, during supervised MR search. In the second part of phase 2, the estimated coordinates' values became more stable, providing a successful localization, and kept stable for all phases 3 and 4. This trend can be further confirmed by analyzing MR centroid positions estimation in the search window reference frame [Fig. 6(c)]. After the first phase of the supervised search, the distribution of the detected centroid points concentrates in a confined region (black dashed circle), thus indicating localization success. These experiments also confirmed that, for the chosen frequency range of the magnetic field sequences, the

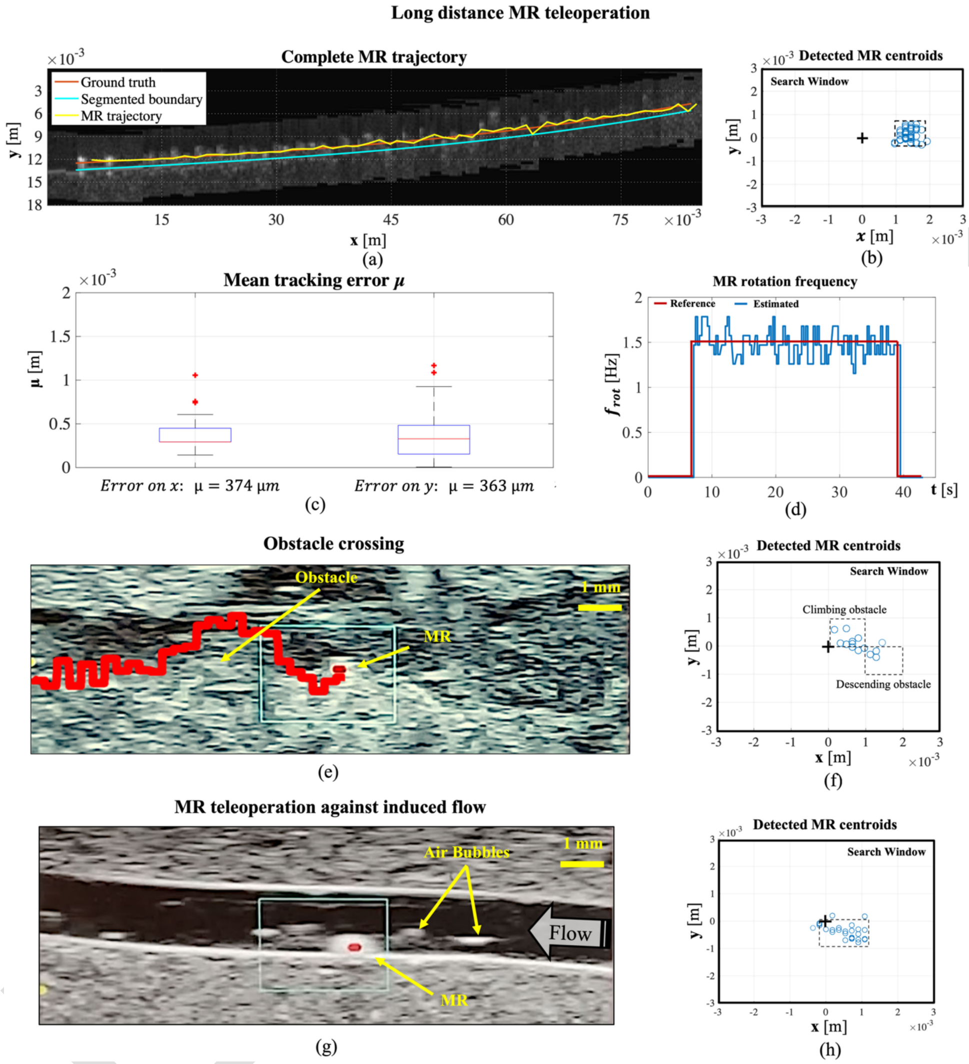


Fig. 7. MR visual-servoing experiments. (a) Complete trajectory traveled by the MR during long-distance teleoperation experiments and lumen morphology reconstructed through segmentation of B-mode images. (b) Collection of detected MR centroids with respect to the search window center for every control iteration. (c) Mean position tracking error for the lateral (x) and axial (y) dimensions. (d) Reference and estimated MR angular velocity. (e) B-mode images and overlapped tracked trajectory (red curve) of the MR when teleoperated across an obstacle in the lumen. (f) Detected MR centroids distribute in two clusters, namely climbing obstacle and descending obstacle, representing different linear motion conditions. (g) MR teleoperation against opposite laminar flow. White particles in B-mode images represent air bubbles generated by the pump. (h) Detected MR centroids are distributed in a larger cluster, indicating robust teleoperation, with some linear speed fluctuations due to flow disturbances. Subfigures (b), (f), and (h) report the detected MR centroids for every control iteration.

dynamics of the vibrational/rotational MR motions in response to the generated magnetic torque are negligible. In fact, the measured MR motion signals (vibrations/rotations) perfectly matched the magnetic driving signals, validating that potential nonideal MR responses (e.g., step-out behaviors) do not introduce any significant frequency or phase lag with respect to the driving signals.

B. Teleoperated MR Visual-Servoing in Phantom Lumen

Once localized the MR through supervised search, we conducted closed-loop control experiments to evaluate the system performances when navigating the MR through the phantom lumen. First, we conducted visual-servoing experiments to assess control stability and performances in terms of tracking errors. In these experiments, the MR was teleoperated along the phantom lumen for about 32 s (Supplementary Video, part 2), with a constant reference rotation frequency of 1.5 Hz. In closed-loop mode, at every control iteration, the robot automatically moved the search window center (optimal manipulation point) to match the estimated MR position. The complete trajectory traveled by the MR was derived by the trajectory of the robot arm end-effector, which corresponds to that of the optimal manipulation point [Fig. 7(a), yellow curve]. At every control iteration, the detected MR centroid position relative to the search window center was collected and reported in a graph [Fig. 7(b)]. The MR traveled along the phantom lumen for a total distance of 80 mm (145 body lengths) uphill.

To assess the system performances in terms of position tracking errors, we defined a ground truth trajectory by assuming that the MR performs a pure rolling motion. To estimate the tracking error on the axial dimension y, we exploited the high contrast fiducial markers provided by the lumen boundaries. We concatenated all the B-mode images acquired by the tracker during the whole MR trajectory and we applied a convolutional filter for horizontal edge detection. In this way, we could segment the lumen boundaries to reconstruct its actual morphology, accounting for potential fabrication defects in the phantom [Fig. 7(a), light blue curve]. Considering the pure rolling motion assumption, we defined a ground truth trajectory as the curve with one MR radius distance from the segmented lumen boundary [Fig. 7(a), red curve]. The deviation of the tracked MR trajectory from the ground truth trajectory resulted in an average axial tracking error of 363 μm (0.66 body lengths) [Fig. 7(c), *Error on y*]. To estimate the tracking error on the lateral dimension x, we defined again a ground truth reference. Given the pure rolling assumption, which implies constant MR linear velocity, at every control iteration, the MR has performed a fixed lateral displacement before the visual-servoing system repositions the search window center. This displacement is known, depends on the MR radius, rotation frequency, and on the tracking frame rate, and can be thus referenced as a ground truth. We compared the detected MR centroids [Fig. 7(b)] to the ground truth displacement, estimating an average lateral tracking error of 374 μm (0.68 body lengths) [Fig. 7(c), *Error on x*].

These experiments showed that the MR performs a rolling motion combined with a small tumbling motion (Supplementary Video, part 2). This behavior, considered as a disturbance to the control system, is most likely generated by MR fabrication defects. For example, if the remanent magnetization of the cylindrical MR is not purely radial but has a small axial component, a minor tumbling motion may arise. Although these deviations from the ideal behavior may introduce a slight inaccuracy in the estimated position tracking errors, these experiments demonstrated that the disturbances are completely rejected by the system, which preserved control stability for the whole traveled trajectory.

Indeed, the detected MR centroid positions [Fig. 7(b)] are concentrated in the same region (black dashed box), supporting the stable linear rolling motion assumption. The average MR linear rolling speed could be estimated by multiplying the average detected MR centroid position by the tracking frame rate (2 Hz), resulting in a speed of 2.5 mm/s (about 4.5 body lengths/s). This value is close to the theoretical linear speed of 2.59 mm/s corresponding to a cylinder with a diameter of 550 μm, performing pure rolling with a rotation frequency of 1.5 Hz. These results confirmed the assumption that the observed tumbling effect can be neglected. The estimated MR rotation frequency [Fig. 7(d), blue plot] followed the imposed reference value [Fig. 7(d), red plot] along the whole path, with a maximum absolute tracking error of 0.28 Hz, corresponding to a relative error of 18%.

We also assessed the robustness of the visual-servoing system to environmental disturbances, both in uneven terrain (e.g., obstacles) and with induced laminar flow in the lumen. First, the MR was teleoperated across an obstacle, represented by an occlusion in the lumen due to material accumulation (Supplementary Video, part 3) [Fig. 7(e), red curve]. Tracking the MR centroid position enabled to verify successful obstacle crossing. In this experiment, the detected MR centroid positions with respect to the search window center were distributed in two main clusters [Fig. 7(f), black dashed boxes], each representing a different linear velocity regimen. The left cluster reflected the obstacle climbing condition, featured by a reduced x component indicating reduced lateral speed and an increased y component indicating increased axial speed. Alternatively, the right cluster reflected the obstacle descending condition, featured by an increased x component indicating increased lateral speed, and negative y component indicating descending axial speed. Overall, the tracker was able to stably track the MR during the whole experiment allowing for successful visual-servoing across the obstacle. Finally, we conducted additional experiments to assess the system robustness in a laminar fluid flow regimen. The flow introduced a double disturbance: 1) it produced a drag force pushing the MR and 2) it generated high contrast moving reflectors in the form of air bubbles [Fig. 7(g)]. The MR was teleoperated against a continuous counter flow. In this case, the collection of detected MR centroids with respect to the search window center was distributed over a larger cluster [Fig. 7(h), black dashed box], reflecting higher fluctuations in the MR linear rolling speed due to flow disturbances. The average linear rolling speed value was here estimated around 0.8 mm/s (around 1.4 body lengths/s); compared to the linear rolling speed in static fluid conditions, this value suggests a possible slip between the MR and the boundary surface. Nonetheless, even when rolling with slip, the tracker could stably track the MR. Although with a lower linear speed, the visual-servoing system enabled robust teleoperation against the counter flow (Supporting Video, part 4). Overall, these experiments confirmed that, in the considered experimental conditions, the disturbances introduced by the unmodeled dynamics (Section II-B) could be successfully rejected by the visual-servoing system.

VII. Conclusion

In this article, we proposed a robust US-enabled MR visual-servoing system for controlling magnetic MRs in simulated in-body environments, where environmental disturbances (e.g., reflective objects and fluid flow) would hamper traditional US imaging techniques, such as B-mode or Color Doppler.

First, we designed a C-arm robotic system for holding the mobile permanent magnet and the US probe allowing for coordinated magnetic actuation and US-APA tracking within the C-arm workspace. Then, we developed a real-time MR tracker based on the US-APA technique and a robot arm controller, which allowed for flexible magnetic actuation and precise positioning of the C-arm system in space according to the tracker's feedback. Finally, we developed a control framework and a set of control signals to ensure feedback integrity and control stability, allowing for teleoperated MR visual-servoing.

We validated the tracker ability to help the operator in localizing the MR (open loop) within the highly reflective phantom without any *a-priori* knowledge about its position. Visual-servoing experiments (closed loop) allowed to evaluate the system performance in terms of tracking accuracy and robustness to environmental disturbances. The MR was navigated for 80 mm uphill along the lumen (145 body lengths) with an average position tracking error of 368 μm (0.67 body lengths). The MR was also teleoperated across obstacles represented by material accumulations occluding the lumen, demonstrating system robustness, and stability in uneven terrain. The system proved stable and robust even to drag forces and highly reflective moving air bubbles generated by a counter flow. Overall, the results reported in this article pave the way for the employment of US-APA as a precise and robust feedback strategy for closed-loop control of MRs inside the body.

In general, the resolution of the US-APA tracking technique is related to the pixel size in the processed phase images. The axial pixel size of a phase image can be up to four times smaller than that of a B-mode image (around 37 μm for standard linear high-definition probes) [33], but the lateral pixel size is limited by the US probe design parameters (around 150 μm for standard linear high-definition probes). However, combining US-APA with recent high-resolution US techniques such as super-resolution US imaging [39], which provide images with miniaturized pixel size down to 10 μm $\times$ 10 μm, could enable detection, visualization, and tracking of MRs with characteristic size in the low microscale range (around 10 μm). The US-APA tracking frame rate, also limited by the computational burdens of Fourier analysis, could be improved as the future development of this article. Furthermore, future article could focus on automatizing MR blind search, thus avoiding the need for an active operator by employing, for instance, learning-based image analysis techniques. Another significant achievement for translating the proposed approach to the clinics could be three-dimensional (3-D) visual-servoing. This would involve feedback control over the orthogonal component to the imaging plane. This problem could be addressed in future works, for example, by elaborating compensation trajectories based on lumen morphology estimated from preoperative imaging or an auxiliary 3-D US probe [40]. Alternatively, the problem could be addressed by implementing deep-learning approaches [41]. Finally, the possibility of extracting US-APA feedback from a wider repertoire of MRs, including helical propellers [42], crawlers [43], and oscillating swimmers [44] could also be investigated in future studies.

Acknowledgment

The authors thank Mohammed Hasan Dad Ansari for his support in fabricating the magnetic microroller.

References

[1] 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.

[2] M. Medina-Sánchez, L. Schwarz, A. K. Meyer, F. Hebenstreit, and O. G. Schmidt, "Cellular cargo delivery: Toward assisted fertilization by sperm-carrying micromotors," *Nano Lett.*, vol. 16, no. 1, pp. 555–561, Jan. 2016.

[3] K. L. Kozielski *et al.*, "Nonresonant powering of injectable nanoelectrodes enables wireless deep brain stimulation in freely moving mice," *Sci. Adv.*, vol. 7, no. 3, pp. 1–14, 2021.

[4] A. C. Hortelao *et al.*, "Swarming behavior and in vivo monitoring of enzymatic nanomotors within the bladder," *Sci. Robot.*, vol. 6, no. 52, 2021, Art. no. eabd2823.

[5] Q. Wang, B. Wang, J. Yu, K. Schweizer, B. J. Nelson, and L. Zhang, "Reconfigurable magnetic microswarm for thrombolysis under ultrasound imaging," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 10285–10291.

[6] Z. Wu *et al.*, "A swarm of slippery micropropellers penetrates the vitreous body of the eye," *Sci. Adv.*, vol. 4, no. 11, pp. 1–11, 2018.

[7] X. Z. Chen *et al.*, "Small-scale machines driven by external power sources," *Adv. Mater.*, vol. 30, no. 15, pp. 1–22, 2018.

[8] T. Xu, L.-P. Xu, and X. Zhang, "Ultrasound propulsion of micro-/nanomotors," *Appl. Mater. Today*, vol. 9, pp. 493–503, 2017.

[9] S. Palagi, D. P. Singh, and P. Fischer, "Light-controlled micromotors and soft microrobots," *Adv. Opt. Mater.*, vol. 7, no. 16, Aug. 2019, Art. no. 1900370.

[10] M. Latulippe and S. Martel, "Dipole field navigation: Theory and proof of concept," *IEEE Trans. Robot.*, vol. 31, no. 6, pp. 1353–1363, Dec. 2015.

[11] N. Ebrahimi *et al.*, "Magnetic actuation methods in bio/soft robotics," *Adv. Funct. Mater.*, vol. 31, no. 11, 2020, Art. no. 2005137.

[12] J. Li, B. Esteban-Fernández de Ávila, W. Gao, L. Zhang, and J. Wang, "Micro/nanorobots for biomedicine: Delivery, surgery, sensing, and detoxification," *Sci. Robot.*, vol. 2, no. 4, Mar. 2017, Art. no. eaam6431.

[13] D. Ahmed, T. Baasch, N. Blondel, N. Läubli, J. Dual, and B. J. Nelson, "Neutrophil-inspired propulsion in a combined acoustic and magnetic field," *Nature Commun.*, vol. 8, no. 1, pp. 1–8, 2017.

[14] Y. Alapan, U. Bozuyuk, P. Erkoc, A. C. Karacakol, and M. Sitti, "Multifunctional surface microrollers for targeted cargo delivery in physiological blood flow," *Sci. Robot.*, vol. 5, no. 42, May 2020, Art. no. eaba5726.

[15] 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, Jan. 2021, Art. no. 2002047.

[16] C. Huang, T. Xu, J. Liu, L. Manamanchaiyaporn, and X. Wu, "Visual servoing of miniature magnetic film swimming robots for 3-D arbitrary path following," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 4185–4191, Oct. 2019.

[17] Z. Yang, L. Yang, and L. Zhang, "3D Visual servoing of helical swimmers using parallel mobile coils," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2020, pp. 9223–9229.

[18] T. Xu, Y. Guan, J. Liu, and X. Wu, "Image-based visual servoing of helical microswimmers for planar path following," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 325–333, Jan. 2020.

[19] L. Yang, J. Yu, and L. Zhang, "Statistics-based automated control for a swarm of paramagnetic nanoparticles in 2-D space," *IEEE Trans. Robot.*, vol. 36, no. 1, pp. 254–270, Feb. 2019.

[20] F. Ongaro, S. Pane, S. Scheggi, and S. Misra, "Design of an electromagnetic setup for independent three-dimensional control of pairs of identical and nonidentical microrobots," *IEEE Trans. Robot.*, vol. 35, no. 1, pp. 174–183, Feb. 2018.

[21] C. Bergeles, B. E. Kratochvil, and B. J. Nelson, "Visually servoing magnetic intraocular microdevices," *IEEE Trans. Robot.*, vol. 28, no. 4, pp. 798–809, Aug. 2012.

[22] M. Medina-Sánchez and O. G. Schmidt, "Medical microbots need better imaging and control," *Nature*, vol. 545, no. 7655, pp. 406–408, 2017.

[23] O. Erin, H. B. Gilbert, A. F. Tabak, and M. Sitti, "Elevation and azimuth rotational actuation of an untethered millirobot by MRI gradient coils," *IEEE Trans. Robot.*, vol. 35, no. 6, pp. 1323–1337, Dec. 2019.

[24] V. Iacovacci *et al.*, "High-resolution SPECT imaging of stimuli-responsive soft microrobots," *Small*, vol. 15, no. 34, 2019, Art. no. 1900709.

[25] D. Li, C. Liu, Y. Yang, L. Wang, and Y. Shen, "Micro-rocket robot with all-optic actuating and tracking in blood," *Light Sci. Appl.*, vol. 9, no. 1, p. 84, 2020. Q5

[26] T. Le, X. Zhang, A. K. Hoshiar, and J. Yoon, "An electromagnetic navigation system with real-time 2D magnetic particle imaging for targeted drug delivery," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1895–1900.

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

[28] F. Ongaro, D. Niehoff, S. Mohanty, and S. Misra, "A contactless and biocompatible approach for 3D active microrobotic targeted drug delivery," *Micromachines*, vol. 10, no. 8, p. 504, 2019.

[29] Q. Wang, L. Yang, J. Yu, P. W. 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.

[30] 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.

[31] Q. Wang *et al.*, "Ultrasound Doppler-guided real-time navigation of a magnetic microswarm for active endovascular delivery," *Sci. Adv.*, vol. 7, no. 9, Feb. 2021, Art. no. eabe5914.

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

[33] 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, no. 1, Jan. 2021, Art. no. 14102.

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

[35] S. L. Charreyron, Q. Boehler, B. Kim, C. Weibel, C. Chautems, and B. J. Nelson, "Modeling electromagnetic navigation systems," *IEEE Trans. Robot.*, vol. 37, no. 4, pp. 1009–1021, Aug. 2021.

[36] H. Azhari, *Basics of Biomedical Ultrasound for Engineers*. Hoboken, NJ, USA: Wiley, 2010.

[37] 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 Biomater.*, vol. 49, pp. 368–378, 2017.

[38] H. Samavat and J. A. Evans, "An ideal blood mimicking fluid for Doppler ultrasound phantoms," *J. Med. Phys.*, vol. 31, no. 4, pp. 275–278, Oct. 2006.

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

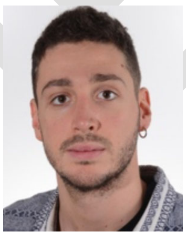
[40] Q. Huang and Z. Zeng, "A review on real-time 3D ultrasound imaging technology," *Biomed. Res. Int.*, vol. 2017, 2017, Art. no. 6027029.

[41] E. Cipolletta *et al.*, "Artificial intelligence for ultrasound informative image selection of metacarpal head cartilage. A pilot study," *Front. Med.*, vol. 8, pp. 1–8, 2021.

[42] D. Li, M. Jeong, E. Oren, T. Yu, and T. Qiu, "A helical microrobot with an optimized propeller-shape for propulsion in viscoelastic biological media," *Robotics*, vol. 8, no. 4, p. 87, 2019.

[43] W. Hu, G. Z. Lum, M. Mastrangeli, and M. Sitti, "Small-scale soft-bodied robot with multimodal locomotion," *Nature*, vol. 554, no. 7690, pp. 81–85, 2018.

[44] L. Yang, Q. Wang, C. I. Vong, and L. Zhang, "A miniature flexible-link magnetic swimming robot with two vibration modes: Design, modeling and characterization," *IEEE Robot. Autom. Lett.*, vol. 2, no. 4, pp. 2024–2031, Aug. 2017.



Stefano Pane (Student Member, IEEE) received the B.Sc. degree in electronic engineering from the University of Roma3, Rome, Italy, in 2015, and the M.Sc. degree (Hons.) in mechatronics engineering from the Politecnico di Torino, Turin, Italy, in 2017. He is currently working toward the Ph.D. degree with the Biorobotics Institute of Scuola Superiore Sant'Anna, Pisa, Italy.

His research interests include medical robotics, microrobotics, magnetic control, and medical ultrasound, as well as biomechatronics and implantable devices.

Giovanni Faoro (Student Member, IEEE) received the B.Sc. degree in biomedical engineering from Università degli Studi di Padova, Padova, Italy, in 2019, and the M.Sc. degree (Hons.) in bionics engineering jointly from the University of Pisa, Pisa, Italy and Scuola Superiore Sant'Anna, Pisa, Italy, in 2021. He is currently working toward the Ph.D. degree with The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy.

His research interests include medical and probabilistic robotics, image processing, and visual-servoing control of microrobots.

Edoardo Sinibaldi (Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in aerospace engineering from the University of Pisa, Pisa, Italy, in 2002, and the Ph.D. degree (Hons.) in mathematics for technology and industry from the Scuola Normale Superiore, Pisa, in 2006.

He was an intern with Rolls-Royce plc, Derby, U.K. After a Postdoc period with BioRobotics Institute, Scuola Superiore Sant'Anna, he moved to the Italian Institute of Technology, Genoa, Italy. His research interests include modeling (at large, and in particular for biomedical applications), modelbased design (in particular for flexible tools), and bioinspired soft robotics (at large, and in particular bioinspired actuation).

Dr. Sinibaldi is currently an Editorial Board Member (Associate/Academic Editor) for *Scientific Reports* (Nature Research), *PLOS One*, and *International Journal of Advanced Robotic Systems*.

Veronica Iacovacci (Member, IEEE) received the Ph.D. degree in biorobotics from Scuola Superiore Sant'Anna, Pisa, Italy, in 2017.

She has been a Postdoctoral Fellow with Scuola Superiore Sant'Anna and Swiss Federal Institute of Technology. She is currently a Marie Curie Global Fellow jointly with Scuola Superiore Sant'Anna and at the Chinese University of Hong Kong, Hong Kong. In the recent years, she has been working on microrobots imaging and retrieval, to bring these technologies closer to the clinical practice. Her research interests include microrobotics for medical applications with a focus on magnetic systems.

Arianna Menciassi (Senior Member, IEEE) is currently a Professor of Bioengineering and Biomedical Robotics with the Scuola Superiore Sant'Anna, Pisa, Italy, where she is a Team Leader of the "Surgical Robotics & Allied Technologies" area with The BioRobotics Institute. She has been a Coordinator of the Ph.D. in biorobotics, since 2018, and she was appointed in 2019 as a Vice-Rector with the Scuola Superiore Sant'Anna. Her research interests include surgical robotics, microrobotics for biomedical applications, biomechatronic artificial organs, and smart and soft solutions for biomedical devices. She pays a special attention to the combination between traditional robotics, targeted therapy, and wireless solution for therapy (e.g., ultrasound- and magnetic-based).

Prof. Menciassi was on the Editorial Board of IEEE-ASME Transactions on Mechatronics, was a Topic Editor of the *International Journal of Advanced Robotic Systems* (2013–2020). She is currently an Editor of IEEE Transactions on Robotics, *APL Bioengineering*, and IEEE Transactions on Medical Robotics and Bionics. She is the Co-Chair of the IEEE Technical Committee on Surgical Robotics.