IEEE INSTRUMENTATION & MEASUREMENT SOCIETY

Received March 2025; revised June 2025; accepted June 2025.

Digital Object Identifier 10.1109/10.1109/OJIM.2025.3581634

Integrating Confidence Maps and Visual Servoing for Needle Tracking in Robotic US-Guided Percutaneous Nephrolithotomy

Hoorieh Mazdarani[1] (Graduate Student Member, IEEE), James Watterson[2], Rebecca Hibbert[3], AND Carlos Rossa[1] (Senior Member, IEEE)

[1]Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada
[2]Faculty of Medicine, Division of Urology, University of Ottawa, Ottawa, ON, Canada
[3]Department of Radiology, Mayo Clinic, Rochester, MN, USA

CORRESPONDING AUTHOR: Hoorieh Mazdarani (e-mail: hooriehmazdarani@cmail.carleton.ca).

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, [funding # ALLRP 580714-22].

ABSTRACT Ultrasound-guided percutaneous nephrolithotomy (PCNL) is a minimally invasive procedure to remove large kidney stones through an incision in the patient's back. PCNL requires a high level of dexterity to steer a surgical tool while visualizing it using ultrasound images. A robotic system that controls the ultrasound probe to automatically image the tool would alleviate the surgeon's cognitive workload and potentially lead to more accurate kidney access.

We propose a novel algorithm that combines visual servoing and confidence maps to track the position of a manually steered needle using a robotically actuated ultrasound probe. The algorithm automatically adjusts the position of the ultrasound probe so that the same longitudinal portion of the needle shaft is visible in the image, while simultaneously ensuring acoustic contact between the ultrasound probe and the tissue over uneven surfaces. Unlike previous methods, where confidence maps were used for probe positioning with static targets, this paper introduces the first unified algorithm that optimizes image quality while tracking a moving tool. It ensures continuous probe-tissue contact on uneven surfaces and does not require prior knowledge of the needle's trajectory or additional sensors. The algorithm, evaluated in phantom tissue and in a realistic kidney mannequin, shows an average tool tracking accuracy of 1.65 mm and 1.17 mm, respectively, confirming its ability to reliably track a manually inserted tool during PCNL.

INDEX TERMS Ultrasound imaging, visual servoing, confidence maps, needle tracking, robot-assisted surgery, percutaneous nephrolithotomy

I. INTRODUCTION

PERCUTANEOUS nephrolithotomy (PCNL) is a type of minimally invasive surgery where the surgeon punctures the skin and passes down needles or other tools to remove large kidney stones through an incision in the patient's back. Unlike open surgery, in PCNL the surgeon relies on ultrasound (US) or fluoroscopy (real-time X-ray) images to track the location of the tool relative to the target. In North America, fluoroscopy has been the preferred imaging modality for needle guidance during kidney access. Although easy to operate, it exposes patients and interventionists to a significant level of radiation [1], [2]. Moreover, single-plane imaging and difficulties in visualizing organs adjacent to the kidney make kidney access challenging and pose the risk of accidental injury [3], [4]. In Asia and Europe, US-guided PCNL (usPCNL) is an emerging alternative to fluoroscopy. With US images the surgeon can detect radiolucent stones and delineate the adjacent viscera, the anterior, and the posterior calyces with greater accuracy than fluoroscopy, all in a radiation-free setting [5].

UsPCNL requires an increased level of dexterity compared to fluoroscopy. The surgeon must precisely coordinate the position and orientation of the US probe and needle to keep the latter in the imaging plane while simultaneously steering

it toward the calyx [6]. To alleviate the surgeon's cognitive workload, robotic assistance has been proposed for percutaneously steering a needle in a wide range of US-guided interventions [7]–[13]. For example, robotic assistance has been applied to kidney access, prostate brachytherapy, and ablation of liver and kidney tumours [14]–[20] and has also been considered for tool and organ imaging [21]–[25].

To control the US probe position and orientation based solely on real-time US images, different variations of visual servoing (VS) have been used. In VS, the goal is to determine the required probe's velocity so that the real-time image converges to a desired image of the needle or organ. To this end, a set of image features that provide quantitative information about a particular characteristic of the image are defined. These features are then compared with the desired features from a target image, and the error is calculated. The optimal probe velocity is then determined to minimize the error and make the real-time images converge to the desired images. Applications of VS include needle imaging during percutaneous cholecystectomy [26], scanning of peripheral arteries [27], and approaches using 3D US images [28]. In our previous work [29], we used VS to track the in-plane motion of a needle in 3-DOF in a water tank.

Robot-aided US imaging requires sufficient acoustic coupling between the US probe and the tissue. Force and impedance control are commonly employed to ensure acoustic coupling while limiting the force applied to tissue [30]–[32]. Force sensors are usually expensive and noisy, and the measured force depends strongly on the sensor orientation and US probe weight, requiring precise online calibration [33], [34]. Furthermore, contact force control is usually only implemented along the US probe's longitudinal axis [35]–[38], where the controller maintains a constant force but neglects the resulting image quality. In practice, the body surface is uneven and therefore it is desirable to orient the probe orthogonally to the surface. This requires more advanced approaches such as trajectory planning using 3D surface perception [39], human pose estimation using RGB-D cameras [40], [41], patient-to-MRI registration for trajectory planning [42], [43], contact posture estimation with optical waveguide force sensors [44], solving an online optimization problem for normal direction estimation [33], reinforcement learning [45], or task-space compliance control [46].

Rather than relying on force sensors, it is possible to quantify the probe/tissue acoustic coupling from quantitative metrics of the US image itself. Confidence maps, initially proposed by Karamalis et al. [47], are pixel-wise probability maps showing the reliability of detected features or regions within an image. Chatelain et al. [48] used confidence maps to orient an US probe during teleoperation to achieve optimal image quality and to maintain a target centred in the image in an automated imaging algorithm. With a similar approach, confidence map is also used by Jiang et al. [49] to optimize the in-plane orientation of the probe in an automatic positioning task. While useful to locate and image a fixed target in the tissue, confidence maps have not been combined with VS to track a moving object such as a needle.

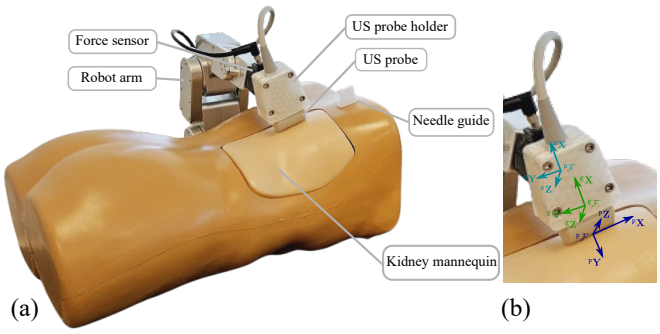


FIGURE 1. US-based needle tracking on a kidney mannequin. a) The setup with a robot arm, an US machine, and a needle guide. b) All coordinate frames: A force sensor frame ${}^{s}\mathcal{F}$ is defined at the centre of the sensor, a probe frame ${}^{p}\mathcal{F}$ is placed at the centre of probe imaging end, and a probe gravity frame ${}^{g}\mathcal{F}$ is defined at the probe's centre of mass.

In this paper, we extend the concept proposed in [48] to create a combined confidence map/VS algorithm to track a surgical needle using 2D longitudinal US images while maintaining acoustic coupling with an uneven surface. Although VS and confidence maps have been used separately for needle tracking and ensuring acoustic coupling, respectively, an approach where both algorithms work concurrently and without the need for a force sensor has not been proposed before. Our algorithm accounts for needle motion in 3 degrees of freedom (3-DOF), that is, two in-plane translations and one in-place rotation, and replaces the force control for vertical movement with a confidence-based feature.

The key contributions of this paper include 1) a novel confidence-based control strategy that dynamically adjusts the probe's position and orientation to maintain optimal acoustic contact and visibility while tracking a manually steered needle, 2) a sensorless visual tracking system that eliminates the need for calibration, filtering, or registration procedures, and 3) real-time performance at 25 fps, matching standard ultrasound imaging rates. To the best of our knowledge, this is the first time that confidence maps and VS are merged to track a manually steered needle with surface contact control while optimizing image quality. The algorithm operates solely based on longitudinal US images, without prior knowledge of the needle trajectory, tip position, or additional position sensors. This makes the system easier to integrate into clinical workflows compared to methods that rely on external sensors.

As in our previous work [29], an US probe is attached to a robotic manipulator and controlled to follow the longitudinal plane of a moving needle, with the configuration shown in Fig. 1. Image features are extracted from real-time US images. The algorithm determines the required linear and angular speeds of the US probe based on the difference between the desired and current image features, ensuring that the needle is always visible in the images. A major limitation

of our past work was that the tissue was assumed to be a flat surface. In this paper, the concept is expanded to an uneven surface and confidence maps are used in conjunction with VS to ensure acoustic contact during tracking. The paper is organized as follows. The next section introduces the unified VS and confidence map algorithm along with the robot control law. The experimental setup using a realistic kidney phantom and agar tissues is presented along with the results. Finally, we discuss the contributions of the work and recommendations for future work.

II. UNIFIED NEEDLE TRACKING AND SURFACE CONTACT CONTROLLER

During usPCNL, seeing the needle's tip is crucial for precise puncturing. Seeing the needle's shaft ensures that the needle is aligned with the imaging plane and that the end of the visible portion of the needle corresponds to the needle tip. Therefore, the objective of the controller is to track the needle using longitudinal US images, while maintaining acoustic coupling between the tissue and probe.

Consider an US probe attached to a robotic arm as in Fig. 1(a). The goal of the controller is to find the velocity of the robot's end effector required to keep the imaging plane aligned with the needle shaft. Let the velocity vector be:

$$\mathbf{v} = [v_x \quad v_y \quad \omega_z]^T, \tag{1}$$

where v_x and v_y are the translational velocities of the probe, and ω_z the rotational component of probe velocity, along the x, y, and z axes of a given reference frame, respectively.

Considering the VS algorithm with the velocity vector $\mathbf{v}$ defined in the probe frame ${}^{\mathrm{p}}\mathcal{F}$ and assuming $\mathbf{s} \in \mathbb{R}^3$ to be a vector of measurable and differentiable image features, these two vectors can be related as:

$$\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v}, \tag{2}$$

where $\mathbf{L}_s \in \mathbb{R}^{3\times 3}$ is the interaction matrix representing the relation between $\mathbf{v}$ and $\dot{\mathbf{s}}$. If the vector of desired features is $\mathbf{s}^*$, the error between the desired and current image features is $\mathbf{e} = \mathbf{s}^* - \mathbf{s}$, and the control action is:

$$\mathbf{v} = \mathbf{K}_p \mathbf{L}_s^{\dagger} \left(\mathbf{s}^* - \mathbf{s}\right), \tag{3}$$

where $\mathbf{K}_p \in \mathbb{R}^{3\times 3}$ is a positive diagonal matrix and $\mathbf{L}_s^{\dagger}$ is the pseudo-inverse of $\mathbf{L}_s$. The error can be shown to decrease exponentially, making the features extracted from the real-time images converge to the desired features.

Traditionally, the interaction matrix is made up of the partial derivatives of the image features with respect to the camera speed, and then inverted to find the control action. Here, we used a different approach where $\mathbf{L}_s^{\dagger}$ in (3) is arbitrarily chosen among possible generalized forms, such that the probe velocity can be directly written as a function of the image feature error [48]. In this way, it is not necessary to compute the inverse of $\mathbf{L}_s$, which provides more flexibility in the design of the control law. Fig. 2 depicts the proposed control approach. The next step is to define appropriate image features to establish the control law.

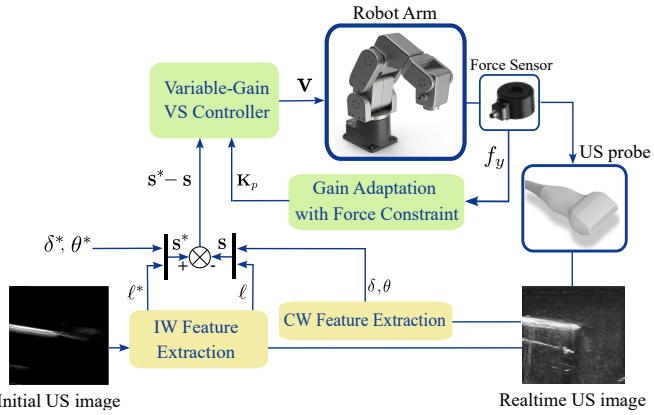


FIGURE 2. Block diagram of the proposed algorithm. Real-time US images are used to calculate intensity-weighted (IW) and confidence-weighted (CW) features, which are sent to the VS controller to determine the US probe's speed. The VS control gains are updated according to the current feature values and force constraints.

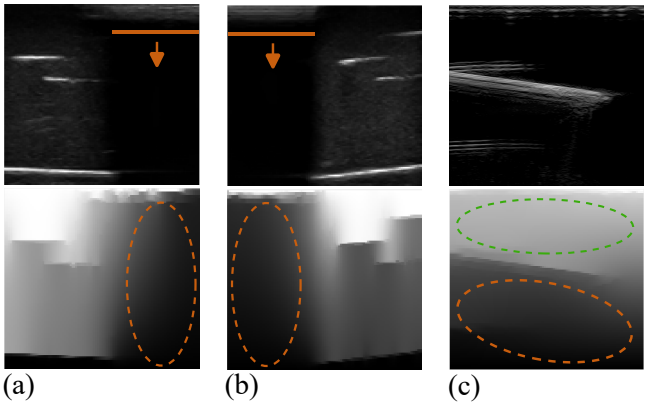


FIGURE 3. a) and b) US image with partial probe/tissue contact (first row). The orange line shows the area of no contact. The second row shows the corresponding confidence map indicating low confidence in the no-contact area. c) US image with full probe-to-tissue contact and its confidence map, showing the needle as a reflective object. The green line shows an area with high confidence as a result of acoustic coupling, and the orange line shows low confidence beneath the needle.

A. ULTRASOUND IMAGE FEATURES DEFINITION

Image moments of different orders are widely used as image-defining features in computer vision [50] and their choice is essential to achieve convergence of the controller [51]. We propose to use two image moments: 1) intensity-weighted calculated over the actual US image for the needle tracking, and 2) confidence-weighted, calculated over the confidence map of the US image to control the probe-tissue contact.

1) Intensity-Weighted Image Moments

The general form of an image moment of order $i + j$ in a greyscale image is:

$$m_{ij} = \iint x^i y^j I(x, y) dx dy, \tag{4}$$

where (x, y) is the coordinate of each image pixel and $I(x, y)$ is their intensity. In [29] moment-based image features were used to represent the needle pose in US images. Here, we will use the same approach for tracking the needle motion along the direction of needle insertion (x-axis). Since the needle shaft must be visible to locate its tip, the first feature is defined as the visible length ℓ of the needle shaft:

$$\ell = 4\sqrt{\frac{2}{m_{00}}\left[\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right]}, \quad (5)$$

where

$$\mu_{11} = m_{11} - \tfrac{m_{10}m_{01}}{m_{00}},$$
$$\mu_{20} = m_{20} - \tfrac{m_{10}^2}{m_{00}}, \quad \mu_{02} = m_{02} - \tfrac{m_{01}^2}{m_{00}}.$$

By choosing the first row of $\mathbf{L}_s^\dagger$ in (3) as $\mathbf{L}_\ell^\dagger = [-1\ 0\ 0]$ the control law for in-plane needle tracking can be defined as:

$$v_x = -k_\ell\ (\ell^* - \ell), \quad (6)$$

where ℓ^* is the desired ℓ, and k_ℓ is the control gain.

2) Confidence-Weighted Image Moments

A confidence map is a pixel-based probability map that shows which parts of the image are most reliable, helping assess the certainty of the data across different parts of the image. The complete algorithm can be found in [47]. The probability of each pixel in the map depends on the tissue properties but also on the path travelled by the acoustic signal from that point. In our application, two main factors can affect the confidence map. First, the loss of US signal due to weak or improper acoustic contact will create an area of low confidence below the section of the image where contact was lost. Second, the existence of a strongly reflective object, such as a needle, drops the confidence level beneath it. These effects can be seen in Fig. 3.

Chatelain *et al.* used the angular coordinate of the barycentre of confidence map to orient a 2D curvilinear probe and optimize image quality [48]. Inspired by their work, we define two confidence-based features to control the vertical and rotational motions of the probe. Confidence-weighted image moments of order $i + j$ may be derived from the confidence map similarly to (4) by replacing the pixels' intensity with confidence values:

$$cm_{ij} = \iint_{\mathcal{H}} x^i\ y^j\ \mathbf{C}(x, y)\ dx\ dy, \quad (7)$$

where $\mathbf{C}(x, y) \in [0, 1]$ is the confidence of each pixel, and $\mathcal{H}$ is a section of the image. Since these moments are solely used for contact control, variations of the needle position in the image should not affect them. Therefore, $\mathcal{H}$ is arbitrarily defined to not include the needle. Two corresponding image features can now be defined. The first feature is the normalized confidence δ in $\mathcal{H}$. It quantifies the contact between the probe and tissue as

$$\delta = (cm_{00})n^{-1}, \quad (8)$$

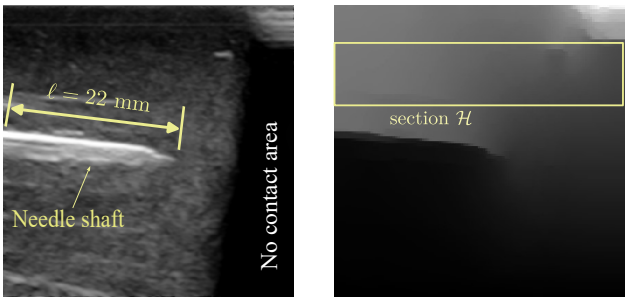


FIGURE 4. US image of the needle shaft (left) when the probe is in partial contact with the tissue and the confidence map (right) where image section $\mathcal{H}$ is shown. In this image, $\delta = 0.42$, and $\theta = -10°$.

where n is the total number of pixels in $\mathcal{H}$. When the US probe makes full contact with the tissue $\delta \to 1$, indicating higher confidence. As the probe loses contact, $\delta \to 0$ indicates lower confidence. Therefore, we can use δ to control the probe vertical displacement by choosing $\mathbf{L}_\delta^\dagger = [0\ \ -1\ \ 0]$, as the second row of $\mathbf{L}_s^\dagger$ in (3), and defining the probe velocity along the y axis as:

$$v_y = -k_\delta\ (\delta^* - \delta), \quad (9)$$

where δ^* is the desired δ and $k_\delta > 0$ is a control gain. The second feature gives the orientation of the probe with respect to a vector normal to the tangent of the tissue surface as:

$$\theta = \arctan\left(\frac{cm_{10} - (x_{mid})\ (cm_{00})}{cm_{01}}\right) \quad (10)$$

where x_{mid} is the horizontal middle point of the image. The angle θ can be employed for controlling the probe's in-plane orientation by choosing the third row of $\mathbf{L}_s^\dagger$ in (3) as $\mathbf{L}_\theta^\dagger = [0\ \ 0\ \ -1]$, and defining the rotational component of probe velocity about the z axis as:

$$w_z = -k_\theta\ (\theta^* - \theta), \quad (11)$$

where θ^* is the desired θ, and $k_\theta > 0$ is a control gain. If $\theta^* = 0$, the probe is controlled to align with a vector normal to the tissue surface. Since the above is only valid when the probe is in contact with the tissue, δ must be greater than a minimum value and (11) is reformulated with a Sigmoid activation term that depends on δ as:

$$w_z = -k_\theta\ (\theta^* - \theta)\ \frac{1}{1 + \mathrm{e}^{\zeta(\delta_{min} - \delta)}}, \quad (12)$$

where δ_{min} is a predefined threshold and $\zeta >> 1$ is a coefficient of the sigmoid function. Fig. 4 depicts how these three image features are defined based on an US image.

B. FORCE CONSTRAINT

While the combined visual servoing and confidence map algorithm does not require any sensors other than an US probe, in this section we propose the *optional* addition to constraint the vertical force f_y the US probe applies to the tissue. This inevitably requires the use of a force sensor connected between the US probe and the robot. The

Algorithm 1 Confidence-based VS for needle tracking.

Input: real-time 2D US image
Output: probe velocities $\mathbf{v}$
Initializing:
Set desired features δ^* and $\theta^*(=0)$
Set controller parameters: k_ℓ, k_δ, k_θ, ζ, δ_{min}, and f_{max}
Wait until the inserted needle becomes visible
Calculate desired needle length ℓ^* (5)
while needle is manually inserted **do**
  Capture real-time US image
  Calculate the confidence map
  Extract features (5), (8), and (10)
  Update variable gain matrix $\mathbf{K}_p$
  Calculate probe velocities $\mathbf{v}$ (3)
end while

maximum force constraint can be applied by multiplying the control law (9) by a Softplus activation function as:

$$v_y = -k_\delta\ (\delta^* - \delta)\ \log\left[1 + \mathrm{e}^{\zeta(f_{max} - f_y)}\right]. \tag{13}$$

The additional term in the above equation acts as a variable control gain that increases with the difference between the vertical force f_y and maximum allowed force f_{max}. Force f_y is the second component of the force tensor in the probe frame ${}^{\mathtt{p}}\mathcal{F}$, that is $f_{\mathtt{y}} = [0\ 1\ 0\ 0\ 0\ 0]\ {}^{\mathtt{p}}\mathbf{f}_{\mathtt{ext}}$, where:

$${}^{\mathtt{p}}\mathbf{f}_{\mathtt{ext}} = {}^{\mathtt{p}}_{\mathtt{s}}\mathbf{T}\left({}^{\mathtt{s}}\mathbf{f}_{\mathtt{raw}} - {}^{\mathtt{s}}_{\mathtt{g}}\mathbf{T}\ {}^{\mathtt{g}}\mathbf{f}_{\mathtt{g}} - {}^{\mathtt{s}}\mathbf{f}_{\mathtt{offset}}\right), \tag{14}$$

and ${}^{\mathtt{p}}_{\mathtt{s}}\mathbf{T}$ and ${}^{\mathtt{s}}_{\mathtt{g}}\mathbf{T}$ are the force/torque transformation matrices from ${}^{\mathtt{s}}\mathcal{F}$ to ${}^{\mathtt{p}}\mathcal{F}$ and from ${}^{\mathtt{g}}\mathcal{F}$ to ${}^{\mathtt{s}}\mathcal{F}$, respectively, with the coordinate frames given in Fig. 1(b). For any two coordinate frames a and b, this transformation matrix is defined as:

$${}^{\mathtt{b}}_{\mathtt{a}}\mathbf{T} = \begin{bmatrix} {}^{\mathtt{b}}_{\mathtt{a}}\mathbf{R} & 0 \\ [{}^{\mathtt{b}}\mathbf{t}_{\mathtt{a}}]_\times\ {}^{\mathtt{b}}_{\mathtt{a}}\mathbf{R} & {}^{\mathtt{b}}_{\mathtt{a}}\mathbf{R} \end{bmatrix} \tag{15}$$

where ${}^{\mathtt{b}}_{\mathtt{a}}\mathbf{R}$ and ${}^{\mathtt{b}}\mathbf{t}_{\mathtt{a}}$ are the rotation matrix and the translation vector relating two coordinate frames. Here, $[{}^{\mathtt{b}}\mathbf{t}_{\mathtt{a}}]_\times$ represents the cross product or the skew-symmetric matrix associated with vector ${}^{\mathtt{b}}\mathbf{t}_{\mathtt{a}}$.

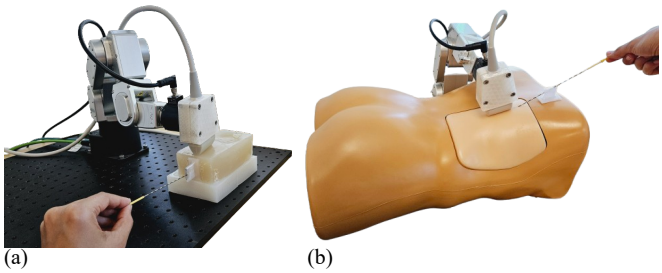
Finally, needle tracking and contact control can be formulated in a unified control algorithm taking the form of an extended VS with variable gains, as in (3) where the feature vector is $\mathbf{s} = [\ell \quad \delta \quad \theta]^T$, $\mathbf{L}_s^\dagger$ is a negative identity matrix with its 3 rows defined by $\mathbf{L}_\ell^\dagger$, $\mathbf{L}_\delta^\dagger$, and $\mathbf{L}_\theta^\dagger$, and $\mathbf{K}_p$ is a diagonal positive definite matrix whose values are:

$$k_{1,1} = k_\ell, \qquad k_{2,2} = k_\delta \log\left[1 + \mathrm{e}^{\zeta(f_{max} - f_y)}\right],$$
$$k_{3,3} = k_\theta \left[1 + \mathrm{e}^{\zeta(\delta_{min} - \delta)}\right]^{-1}. \tag{16}$$

The overall algorithm is briefly explained in Algorithm 1.

III. EXPERIMENTAL VALIDATION

To validate the proposed method we used the setup shown in Fig. 5, where a 40 mm US probe (L15-7H40-A5 from Telemed Ultrasound, Vilnius, Lituania) is attached to the end-effector of a 6-DOF robot arm (Meca500 from Mecademic, Montréal, Canada) through a force/torque sensor (Medusa FT from Bota Systems, Zurich, Switzerland). The US machine streams images at 50 Hz according to the parameters given in Table 1. The algorithm is implemented on an Intel(R) Core i7-9700K computer with a 3.20 GHz CPU and 128 GB of RAM. Matlab is used for US image acquisition and robot control, with a sampling rate of 25 Hz. The robot communicates with the computer via EtherCAT/TwinCAT3. Two different scenarios are used.



(a) (b)

FIGURE 5. Scenario 1 (a) and 2 (b) using a 5% agar phantom tissue and kidney mannequin, respectively. The needle is manually steered while the robot autonomously orients the US probe to track the needle shaft and ensure acoustic coupling between the probe and the tissue.

TABLE 1. Ultrasound image acquisition parameters

frequency	gain	focus	depth	dyn. range	power
10 MHz	80 %	14-21 mm	40 mm	72 dB	−4 dB

Scenario 1 - Agar phantom experiments: In this scenario, we use a phantom made of 5% agar with an arbitrarily designed uneven surface; see Fig. 5(a). We ran 6 trials, during which the needle was inserted from different points in the tissue so that the topology of the tissue surface above the needle was different in each trial. The needle is manually inserted along a 10 cm path for 30 s at a variable speed and with some random back-and-forth movements.

Scenario 2 - Kidney mannequin experiments: In this scenario, see Fig. 5(b), we use a renal biopsy US training mannequin (CAE Blue Phantom, USA) to provide realistic US images and replicate the human anatomy, including ribs, the right kidney with surrounding tissue, and the immediate renal system. The needle is manually inserted along a 7 cm path for 60 s at a variable speed and with random stops and back-and-forth movements. We ran 6 trials, each from different insertion points so that the surface over which the US probe slid was different in each trial.

To constrain the 18-G diamond-tip needle to in-plane motion, a guide was fixed close to the insertion point in the tissue so that the needle could not be moved laterally. Once a small portion of the needle was inserted into the tissue, the position of the US probe was adjusted until the inserted portion of the needle became visible in the image. Tracking then started as the needle was further inserted manually. The desired needle length and normalized confidence were set to

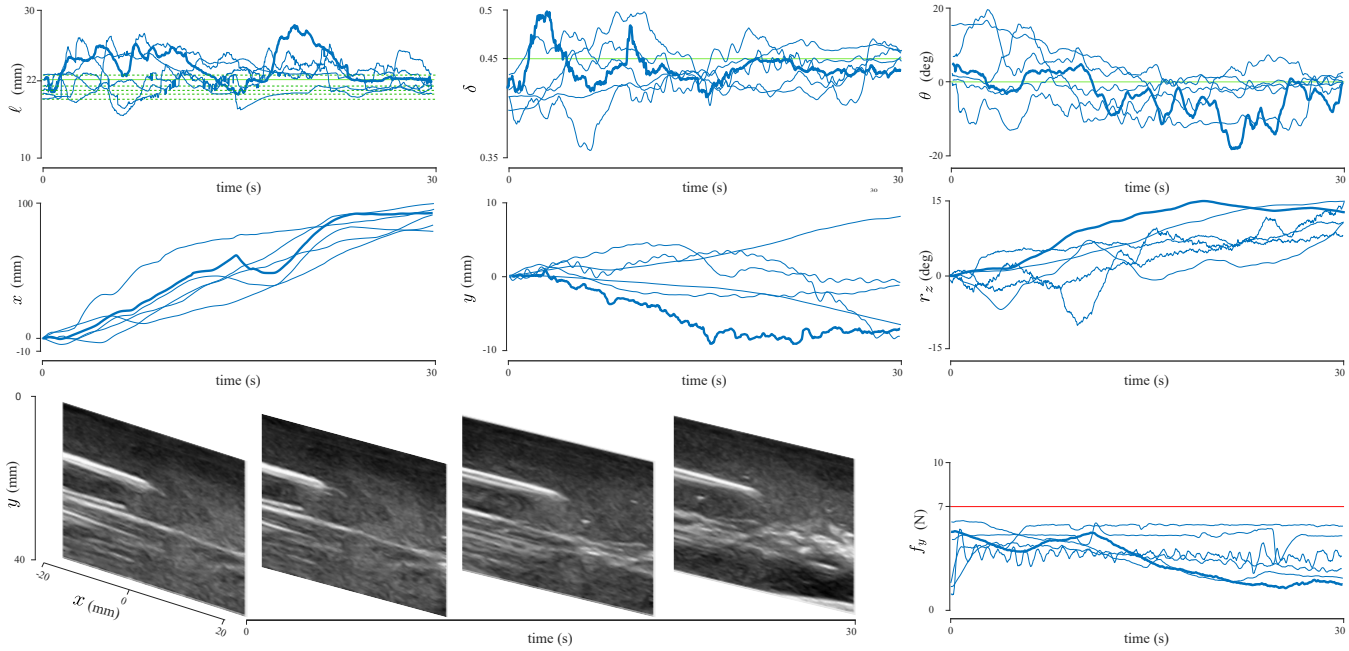


FIGURE 6. Results from 6 trials in Scenario 1. The top 3 panels show the measured and desired needle length ℓ, normalized confidence δ, and orientation angle θ. The needle length is calculated from (5) and confidence-based features from (8) and (10). The middle panels show the US probe displacement in ${}^{\mathrm{p}}\mathcal{F}$. On the bottom left are samples of the US images captured at 4 time steps and on the bottom right the force applied to the tissue.

$\ell^* = 17 - 22$ mm and $\delta^* = 0.45$. The other controller parameters were selected as follows: $k_\ell = 0.4$, $k_\delta = 1.6$, $k_\theta = 0.1$, $\zeta = 100$, $\delta_{min} = 0.35$, and $f_{max} = 7$ N.

IV. RESULTS AND DISCUSSION

Fig. 6 shows the results of the 6 trials in Scenario 1. The first row shows the real-time and the desired image features. Since the desired needle length ℓ^* was captured at the beginning of each test (shown by the green dotted lines), ℓ^* varies across trials. As we can see, while the needle is tracked the features are kept close to the desired values. The second row of Fig. 6 shows the measured displacement of the US probe in ${}^{\mathrm{p}}\mathcal{F}$, where x is the horizontal displacement of the probe along its x-axis, y is the probe's vertical movement following the uneven surface, and r_z is the probe's rotation around its z-axis. The first image feature ℓ controls the probe's motion along the x-axis to keep the needle axis within the imaging plane. In contrast, the vertical displacement y and the rotation r_z are guided by confidence-based features to maintain full probe-tissue contact on uneven surfaces. As a result, the y and r_z components are different between distinct surface topologies. While the probe travelled approximately 100 mm along the x-axis in all trials, the exact motion trajectories differed due to variations in the probe's path and the local surface geometry and interaction dynamics encountered in each trial. Four US images captured at different time steps during one of the trials are also shown (corresponding to the lines in bold in the other plots). The bottom right panel of Fig. 6 shows how the controller guaranteed that the force was kept under the maximum allowed value. Fig. 7(a) provides a statistical summary of the test results in Scenario 1, showing the range of calculated features in each test. As shown in the plot, all three features exhibit variations around their desired values. The variation in needle length is primarily due to the needle's random motion within the tissue. In contrast, the variations in the other two features are mainly attributed to the movement of the ultrasound probe over the uneven tissue surface, which occasionally leads to poor acoustic coupling. Despite these challenges, the median error values and interquartile ranges demonstrate that the proposed method is robust in maintaining feature stability under varying conditions.

Fig. 8 shows the results of Scenario 2. The high variation in the measured features is the result of imaging noise in a realistic scenario. Yet, these features fluctuate around the desired values ensuring successful tracking. While the probe moved around 70 mm along the x-axis in all trials of this scenario, the motion trajectories varied due to trial-specific paths and differences in surface geometry. The vertical force applied to the tissue remained below the maximum allowed range. Fig. 7(b) shows a statistical summary of the results.

Table 2 provides a summary of experimental results with the average values of medians and interquartile ranges of features in all tests in both scenarios. The tracking error in both test scenarios remains around 1.65 mm and 1.17 mm respectively, which is equivalent to less than 8% and 5% variation in the visible needle length. In comparison, it is significantly smaller than the allowable needle placement error of 2.7 mm in targeted percutaneous surgeries [52] and the freehand needle positioning accuracy of 4.6 ± 2.1 mm

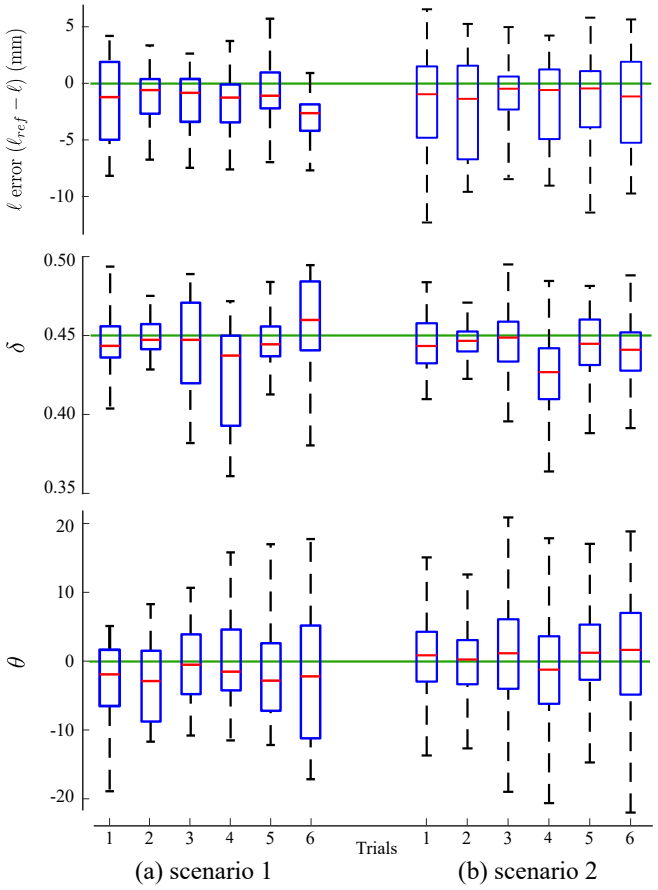


FIGURE 7. Calculated features in 6 trials of each scenario. The red line is the median value, and the box shows the interquartile range.

in usPCNL [53], which confirms the functionality of the proposed method in tracking the needle based only on real-time US images.

A quantitative comparison of the achieved accuracy with other reported methods is provided in Table 3. The imaging modality in all methods is longitudinal 2D US. In contrast to our paper, three out of the five compared methods rely on additional tracking sensors, and the other two on deep learning and offline calibration. Although all methods achieved comparable tracking accuracy, the proposed approach relies solely on 2D US images and eliminates the need for additional sensors such as electromagnetic (EM) or optical tracking systems that are widely used for needle tracking [54]. Moreover, the test environment in all these methods had a flat surface, and none of them dealt with the probe-tissue acoustic contact in case of uneven surface structures.

Previously reported methods addressing the probe-tissue contact problem attempt to maintain a constant vertical force but neglect the resulting image quality [30]–[37] or rely on additional equipment to detect probe/tissue contact [39]–[44]. In contrast, our method accounts for the unknown tissue topology without additional sensors. Since the controller monitors image quality, our method can provide better performance than force sensor-based approaches for needle tracking and tissue imaging. Furthermore, contact control with optimal probe orientation eliminates problems associated with force sensor calibration.

TABLE 2. Statistical summary of the experimental results

	$\ell^* - \ell$ (mm)		δ		θ (deg)	
Scenario	Med. (%[1])	IQR	Med. (%[1])	IQR	Med.	IQR
Sc. 1	-1.65 (8%)	4.33	0.44 (2%)	0.04	-2.59	10.96
Sc. 2	-1.17 (5%)	6.58	0.43 (4%)	0.03	1.02	9.63

[1]Median error relative to desired values.

Finally, unlike other methods that rely on additional sensors to detect the position of the needle [7], [11], [13], [14], our method does not need positional information of the needle's base or tip to track it. The proposed method works at a 25 fps rate, a similar frame rate to most US machines, which meets the requirement of real-time image-guided interventions [55]. This frame rate is also noticeably higher than many other tracking methods (15.7 fps in [12], 18 fps in [13], and 15 fps in [56]), which makes the proposed method suitable for real-time applications.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a unified confidence map and VS method for in-plane needle tracking in a tissue with uneven surfaces. The controller tracks the position of a needle inserted into the tissue while ensuring sufficient acoustic coupling between the tissue and the probe. It adjusts the probe's position and orientation in real-time to optimize image quality and may use an optional controller to limit the maximum contact force. The US probe is controlled in 3-DOF—longitudinal motion, vertical translation, and in-plane rotation—to follow the distal end of the needle shaft and maximize the transducer's surface contact with the tissue. The control law determines the US probe's velocity based on features obtained from real-time longitudinal US images and their corresponding confidence maps, without requiring prior knowledge of the needle trajectory. To the best of our knowledge, this is the first unified confidence map and VS algorithm that 1) ensures probe-tissue contact, 2) optimizes image quality during needle tracking, and 3) operates without additional sensors or trajectory information.

The method was validated experimentally in multiple trials across two distinct scenarios, successfully maintaining needle visibility in US images. For such an algorithm to be used in real-time, low complexity and robustness are essential. While previous methods have used depth cameras or force control for surface modelling and contact control, our method relies solely on widely accessible 2D US images with a computational time below 40 ms, which is suitable for real-time implementation without added resources.

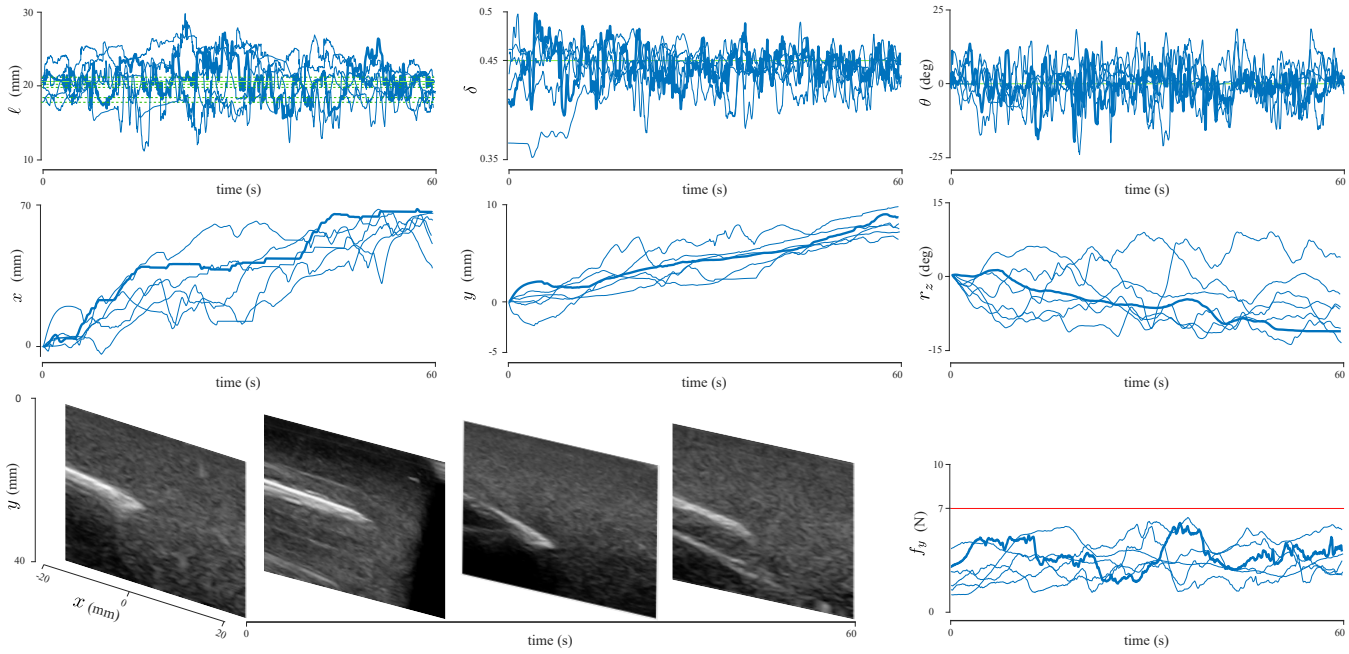


FIGURE 8. Results from 6 experimental trials in Scenario 2. The top 3 panels show the measured and desired needle length ℓ, normalized confidence δ, and orientation θ. The needle length is calculated in (5) and confidence-based features in (8) and (10). The middle panels show the US probe displacement in ${}^{\mathrm{p}}\mathcal{F}$. The bottom left panel shows samples of the US images captured at 4 time steps for the trial corresponding to the bold lines in the other panels. The bottom right shows the measured probe/tissue contact force.

TABLE 3. Comparison of the needle tracking/placement error through different methods

	tracking method	error [mm]	test environment	considerations
Yan et al. [12]	deep learning (DL)	1.20 ± 1.30	phantom and tissue	motorized needle insertion / known velocity
Seitel et al. [14]	EM tracking system	5	gelatin phantom	needle targeting error is reported
Che et al. [11]	optical tracking + DL	1.17 ± 0.70	ex-vivo liver tissue	US probe position is fixed
Baker et al. [13]	fibre-optic tracking sensor	1.10 ± 0.70	ex-vivo bovine tissue	US probe position is fixed
Zheng et al. [10]	simultaneous US-robot calibration	0.75 ± 0.27	tofu phantom	requires prior sample data for calibration
Proposed method	confidence-based visual servoing	1.65 / 1.17	agar / kidney mannequin	manual needle insertion / moving US probe

Future work will focus on integrating organ imaging, segmentation, and optimal path planning. Improved segmentation will reduce noise and enhance controller performance. Additionally, a new control law will be explored to address out-of-plane needle motion.

REFERENCES

[1] D. T. Tzou, T. O. Tailly, and K. L. Stern, "Ultrasound-guided pcnl—why are we still performing exclusively fluoroscopic access?" *Curr Urol Rep*, vol. 24, no. 7, pp. 335–343, 2023.

[2] R. Geraghty *et al.*, "Best practice in interventional management of urolithiasis: an update from the european association of urology guidelines panel for urolithiasis 2022," *Eur Urol Focus*, vol. 9, no. 1, pp. 199–208, 2023.

[3] B. Lojanapiwat, "The ideal puncture approach for pcnl: Fluoroscopy, ultrasound or endoscopy?" *Indian J Urol*, vol. 29, no. 3, p. 208, 2013.

[4] Y.-H. Yang, Y.-C. Wen, K.-C. Chen, and C. Chen, "Ultrasound-guided versus fluoroscopy-guided percutaneous nephrolithotomy: a systematic review and meta-analysis," *World J Urol*, vol. 37, pp. 777–788, 2019.

[5] F. C. Ng *et al.*, "Ultrasound-guided percutaneous nephrolithotomy: Advantages and limitations," *Investig Clin Urol*, vol. 58, no. 5, pp. 346–352, 2017.

[6] C. Chu *et al.*, "Ultrasound-guided renal access for percutaneous nephrolithotomy: a description of three novel ultrasound-guided needle techniques," *J Endourol*, vol. 30, no. 2, pp. 153–158, 2016.

[7] I. Boekestijn *et al.*, "Value-assessment of computer-assisted navigation strategies during percutaneous needle placement," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 10, pp. 1775–1785, 2022.

[8] F. Ferraguti, S. Farsoni, and M. Bonfè, "Augmented reality and robotic systems for assistance in percutaneous nephrolithotomy procedures: recent advances and future perspectives," *Electronics*, vol. 11, no. 19, p. 2984, 2022.

[9] Y. Lin, S. Chen, W. Xu, X. Zhu, and Q. Cao, "Robotic system for accurate percutaneous puncture guided by 3d–2d ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 18, no. 2, pp. 217–225, 2023.

[10] G. Zheng *et al.*, "A simultaneous ultrasound-robot calibration approach for dual-robot intervention by solving the axp = ycq problem," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, 2024.

[11] H. Che *et al.*, "Improving needle tip tracking and detection in ultrasound-based navigation system using deep learning-enabled approach," *IEEE J of Biomed. Health Inform.*, vol. 28, no. 5, pp. 2930–2942, 2024.

[12] W. Yan *et al.*, "Learning-based needle tip tracking in 2d ultrasound by fusing visual tracking and motion prediction," *Medical Image Analysis*, vol. 88, p. 102847, 2023.

[13] C. Baker *et al.*, "Intraoperative needle tip tracking with an integrated fibre-optic ultrasound sensor," *Sensors*, vol. 22, no. 23, p. 9035, 2022.

[14] A. Seitel *et al.*, "Miniaturized electromagnetic tracking enables efficient ultrasound-navigated needle insertions," *Scientific Reports*, vol. 14, no. 1, p. 14161, 2024.

[15] W. Zhang, S. Wang, L. Xie, and R. Yang, "An autonomous liver biopsy robotic system under respiratory motion," *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–10, 2025.

[16] M. Pavone *et al.*, "Ultrasound-guided robotic surgical procedures: a systematic review," *Surg Endosc*, vol. 38, no. 5, pp. 2359–2370, 2024.

[17] O. Wilz, B. Kent, B. Sainsbury, and C. Rossa, "Multiobjective trajectory tracking of a flexible tool during robotic percutaneous nephrolithotomy," *IEEE Robot Autom Lett*, vol. 6, no. 4, pp. 8110–8117, 2021.

[18] J. Lazarus, M. Asselin, and L. Kaestner, "Optically tracked needle for ultrasound-guided percutaneous nephrolithotomy puncture: A preliminary report," *J Endourol*, vol. 35, no. 12, pp. 1733–1737, 2021.

[19] T. de Baere *et al.*, "Robotic assistance for percutaneous needle insertion in the kidney: preclinical proof on a swine animal model," *Eur Radiol Exp*, vol. 6, no. 1, p. 13, 2022.

[20] A. Kuntz *et al.*, "Autonomous medical needle steering in vivo," *Science Robotics*, vol. 8, no. 82, p. eadf7614, 2023.

[21] M. Khadem, C. Rossa, R. S. Sloboda, N. Usmani, and M. Tavakoli, "Ultrasound-guided model predictive control of needle steering in biological tissue," *J Med Robot Res*, vol. 1, no. 01, p. 1640007, 2016.

[22] B. Xu and S. Y. Ko, "3d feedback control using fuzzy logic for a curvature-controllable steerable bevel-tip needle," *Mechatronics*, vol. 68, p. 102368, 2020.

[23] B. Fallahi, C. Rossa, R. S. Sloboda, N. Usmani, and M. Tavakoli, "Sliding-based image-guided 3d needle steering in soft tissue," *Control Eng Pract*, vol. 63, pp. 34–43, 2017.

[24] H. Yang, C. Shan, A. F. Kolen, and P. H. de With, "Medical instrument detection in ultrasound: a review," *Artif. Intell. Rev.*, vol. 56, no. 5, pp. 4363–4402, 2023.

[25] S. Zhang *et al.*, "Ultra-malin: Robotic ultrasound mapping and localization via implicit neural representation," *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–12, 2025.

[26] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata, "An ultrasound-driven needle-insertion robot for percutaneous cholecystostomy," *Phys. Med. Biol.*, vol. 49, no. 3, pp. 441–455, 2004.

[27] F. von Haxthausen *et al.*, "Visual servoing for semi-automated 2d ultrasound scanning of peripheral arteries," in *AUTOMED-Automation in Medical Engineering*, 2020.

[28] C. Nadeau, A. Krupa, J. Petr, and C. Barillot, "Moments-based ultrasound visual servoing: From a mono-to multiplane approach," *IEEE Trans Robot*, vol. 32, no. 6, pp. 1558–1564, 2016.

[29] H. Mazdarani, A. Cotton, and C. Rossa, "2d ultrasound-guided visual servoing for in-plane needle tracking in robot-assisted percutaneous nephrolithotomy," in *2023 IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, 2023, pp. 1786–1791.

[30] C. Nadeau and A. Krupa, "Intensity-based ultrasound visual servoing: Modeling and validation with 2-d and 3-d probes," *IEEE Trans Robot*, vol. 29, no. 4, pp. 1003–1015, 2013.

[31] J. Kim *et al.*, "Development of a control algorithm for the ultrasound scanning robot (nccusr) using ultrasound image and force feedback," *Int. J. Med. Robot*, vol. 13, no. 2, p. e1756, 2017.

[32] F. Suligoj, C. M. Heunis, J. Sikorski, and S. Misra, "Robust–an autonomous robotic ultrasound system for medical imaging," *IEEE access*, vol. 9, pp. 67 456–67 465, 2021.

[33] Z. Jiang *et al.*, "Automatic force-based probe positioning for precise robotic ultrasound acquisition," *IEEE Trans Ind Electron*, vol. 68, no. 11, pp. 11 200–11 211, 2020.

[34] J. Jiang *et al.*, "Force tracking control method for robotic ultrasound scanning system under soft uncertain environment," *Actuators*, vol. 13, no. 2, p. 62, 2024.

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

[36] R. Finocchi *et al.*, "Co-robotic ultrasound imaging: A cooperative force control approach," in *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10135, 2017, pp. 270–280.

[37] T.-Y. Fang *et al.*, "Force-assisted ultrasound imaging system through dual force sensing and admittance robot control," *Int J Comput Assist Radiol Surg*, vol. 12, pp. 983–991, 2017.

[38] H. Mazdarani, B. Sainsbury, J. Watterson, R. Hibbert, and C. Rossa, "Ultrasound-based visual servoing for out-of-plane longitudinal needle tracking in robot-aided percutaneous nephrolithotomy," *IEEE Access*, vol. 13, pp. 69 090–69 102, 2025.

[39] J. Tan *et al.*, "Fully automatic dual-probe lung ultrasound scanning robot for screening triage," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 70, no. 9, pp. 975–988, 2022.

[40] Q. Huang, J. Lan, and X. Li, "Robotic arm based automatic ultrasound scanning for three-dimensional imaging," *IEEE Trans Ind Electron*, vol. 15, no. 2, pp. 1173–1182, 2018.

[41] X. Ma, Z. Zhang, and H. Zhang, "Autonomous scanning target localization for robotic lung ultrasound imaging," in *IEEE Int. Conf. Intell. Robots Syst. (IROS)*, 2021, pp. 9467–9474.

[42] S. Virga *et al.*, "Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms," in *2016 IEEE Int. Conf. Intell. Robots Syst. (IROS)*, 2016, pp. 508–513.

[43] C. Hennersperger *et al.*, "Towards mri-based autonomous robotic us acquisitions: a first feasibility study," *IEEE Trans Med Imaging*, vol. 36, no. 2, pp. 538–548, 2016.

[44] S. Chen *et al.*, "Automatic ultrasound scanning robotic system with optical waveguide-based force measurement," *Int J Comput Assist Radiol Surg*, vol. 16, no. 6, pp. 1015–1025, 2021.

[45] G. Ning, J. Chen, X. Zhang, and H. Liao, "Force-guided autonomous robotic ultrasound scanning control method for soft uncertain environment," *Int J Comput Assist Radiol Surg*, vol. 16, no. 12, pp. 2189–2199, 2021.

[46] J. Wang *et al.*, "Task space compliant control and six-dimensional force regulation toward automated robotic ultrasound imaging," *IEEE Trans Autom Sci Eng*, vol. 21, no. 3, pp. 3652–3663, 2023.

[47] A. Karamalis, W. Wein, T. Klein, and N. Navab, "Ultrasound confidence maps using random walks," *Med Image Anal*, vol. 16, no. 6, pp. 1101–1112, 2012.

[48] P. Chatelain, A. Krupa, and N. Navab, "Confidence-driven control of an ultrasound probe," *IEEE Trans Robot*, vol. 33, no. 6, pp. 1410–1424, 2017.

[49] Z. Jiang *et al.*, "Automatic normal positioning of robotic ultrasound probe based only on confidence map optimization and force measurement," *IEEE Robot Autom Lett*, vol. 5, no. 2, pp. 1342–1349, 2020.

[50] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans Robot*, vol. 21, no. 6, pp. 1116–1127, 2005.

[51] F. Chaumette, "Image moments: a general and useful set of features for visual servoing," *IEEE Trans Robot*, vol. 20, no. 4, pp. 713–723, 2004.

[52] T. L. De Jong *et al.*, "Needle placement errors: do we need steerable needles in interventional radiology?" *Medical Devices: Evidence and Research*, pp. 259–265, 2018.

[53] I. M. Spenkelink, J. Heidkamp, Y. Avital, and J. J. Fütterer, "Evaluation of the performance of robot assisted ct-guided percutaneous needle insertion: Comparison with freehand insertion in a phantom," *European Journal of Radiology*, vol. 162, p. 110753, 2023.

[54] Z. Cheng *et al.*, "Sensing technologies for guidance during needle-based interventions," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–15, 2024.

[55] S. B. Mondal *et al.*, "Real-time fluorescence image-guided oncologic surgery," *Advances in cancer research*, vol. 124, pp. 171–211, 2014.

[56] C. Mwikirize *et al.*, "Time-aware deep neural networks for needle tip localization in 2d ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, pp. 819–827, 2021.