



**HAL**  
open science

## Confidence-Driven Control of an Ultrasound Probe

Pierre Chatelain, Alexandre Krupa, Nassir Navab

► **To cite this version:**

Pierre Chatelain, Alexandre Krupa, Nassir Navab. Confidence-Driven Control of an Ultrasound Probe. IEEE Transactions on Robotics, 2017, 33 (6), pp.1410-1424. hal-01551431

**HAL Id: hal-01551431**

**<https://inria.hal.science/hal-01551431>**

Submitted on 30 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Confidence-Driven Control of an Ultrasound Probe

Pierre Chatelain, Alexandre Krupa, *Member, IEEE*, and Nassir Navab

**Abstract**—We present a control framework for optimizing the image quality during robotic ultrasound acquisitions. The quality of the ultrasound signal across the field of view is represented by a confidence map that is computed online from the B-mode frames, following a model of sound propagation. Moments extracted from this confidence map are used to design a control law for optimizing imaging quality, based on the task function approach. The proposed *confidence control* is combined with force and position control to build two illustrative applications. First, we use force control and confidence control in order to maintain a correct pressure and a good orientation of the probe during teleoperation. Thus, control is shared between a human operator and the robot. Then, we add an automatic positioning task, so that the quality is optimized while maintaining a target in the image center. We show experimentally that confidence-driven control can effectively optimize the acoustic window in real-time. In addition, we show that it can improve the tracking robustness, by preventing the target from being shadowed. Finally, we present the results of experiments performed on a human volunteer.

**Index Terms**—Image quality, medical robotics, robot control, ultrasonography, visual servoing.

## NOMENCLATURE

|                    |  |
|--------------------|--|
| $a$                | Scalar.  |
| $\mathbf{a}$       | Vector.  |
| $\mathbf{A}$       | Matrix.  |
| $\mathbf{I}_n$     | Identity matrix of size $n$ .  |
| $\Omega_{2D}$      | $= [r_{\min}, r_{\max}] \times [\theta_{\min}, \theta_{\max}]$ . 2D field of view. |
| $\Omega_{3D}$      | $= \Omega_{2D} \times [\phi_{\min}, \phi_{\max}]$ . 3D field of view.              |
| $O$                | Imaging center.  |
| $M$                | Intersection between the motor axis and the central scan line.                     |
| $I$                | Ultrasound image.  |
| $\mathbb{R}$       | Set of real numbers.   |
| $\mathbb{R}^+$     | Set of positive real numbers.  |
| $\mathcal{F}_p$    | Probe frame.   |
| $\mathcal{F}_{pc}$ | Contact frame.   |
| $\mathcal{F}_s$    | Force sensor frame.  |
| $\mathcal{F}_g$    | Gravity frame.   |

|                               |  |
|-------------------------------|--|
| $\mathcal{F}_r$               | Reference frame.   |
| $\mathbf{v}$                  | $= (v_x \ v_y \ v_z \ \omega_x \ \omega_y \ \omega_z)^\top$ . Sensor velocity screw. |
| $\mathbf{s}$                  | Visual feature vector.   |
| $\mathbf{s}^*$                | Desired visual feature vector.   |
| $\mathbf{e}$                  | $= \mathbf{s} - \mathbf{s}^*$ . Visual feature error.                                |
| $\dot{\mathbf{e}}^*$          | Desired variation of $\mathbf{e}$ .  |
| $\lambda$                     | Control gain.  |
| $\mathbf{L}_s$                | Interaction matrix associated to $\mathbf{s}$ .                                      |
| $\mathbf{L}_s^g$              | Generalized inverse of $\mathbf{L}_s$ .  |
| $\ker(\mathbf{L}_s)$          | Kernel of $\mathbf{L}_s$ .   |
| $\mathbf{P}_s$                | Projection operator onto $\ker(\mathbf{L}_s)$ .                                      |
| ${}^a\mathbf{H}_a$            | Vector of forces and torques expressed in $\mathcal{F}_a$ .                          |
| ${}^a\mathbf{F}_b$            | Twist transformation matrix from $\mathcal{F}_b$ to $\mathcal{F}_a$ .                |
| $m_p$                         | Mass of the probe.   |
| $g$                           | $\approx 9.81 \text{ m} \cdot \text{s}^{-2}$ . Standard acceleration due to gravity. |
| $\mathbf{S}_y$                | $= (0 \ 1 \ 0 \ 0 \ 0 \ 0)$ . Selection vector.                                      |
| $k$                           | Stiffness.   |
| $\mathbf{t}$                  | $= (x_t, y_t, z_t)$ . Target coordinates.  |
| $C$                           | Confidence map.  |
| $\mathbf{c}$                  | Vector representation of $C$ .   |
| $\mathbf{p}, \mathbf{q}$      | $\in \Omega$ . Pixels in the field of view.  |
| $\tilde{I}(\mathbf{p})$       | Attenuated intensity.  |
| $w_r$                         | Radial propagation weight.   |
| $w_l$                         | Lateral propagation weight.  |
| $\Lambda$                     | Graph Laplacian matrix.  |
| $\mathcal{N}(\mathbf{p})$     | Set of neighbors of $\mathbf{p}$ .   |
| $\theta_c, \phi_c$            | Angular confidence moments.  |
| $\Delta R$                    | Offset between $O$ and $M$ .   |
| $C_\Omega$                    | Sum of confidence over $\Omega$ .  |
| $r_t, \theta_t, \phi_t$       | Target polar coordinates.  |
| $\mathcal{R}$                 | Region of interest.  |
| $\mathbf{s}_{\text{track}}$   | Vector of intensities in the tracked region of interest.                             |
| $\mathbf{s}_{\text{track}}^*$ | Vector of intensities in the initial region of interest.                             |
| $\rho$                        | Normalized correlation.  |

Manuscript received November 7, 2016; revised April 4, 2017; accepted June 8, 2017. This work was supported in part by the Bayerische Forschungstiftung funding for project RoBildOR. This paper was presented in part at the IEEE International Conference on Robotics and Automation, Seattle, WA, May 2015 [1], and in part at the IEEE International Conference on Robotics and Automation, Stockholm, Sweden, May 2016 [2].

P. Chatelain was with IRISA, University of Rennes 1, 35042 Rennes, France, and also with Computer Aided Medical Procedures, Technical University of Munich, 85748 Garching bei München, Germany. He is now with the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK (e-mail: pierre.chatelain@eng.ox.ac.uk).

A. Krupa is with Inria Rennes – Bretagne Atlantique, IRISA, 35042 Rennes, France (e-mail: alexandre.krupa@inria.fr).

N. Navab is with Computer Aided Medical Procedures, Technical University of Munich, 85748 Garching bei München, Germany, and also with Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: nassir.navab@in.tum.de).

## I. INTRODUCTION

**M**EDICAL ultrasound is used increasingly for both diagnostic and interventional imaging. Indeed, ultrasound is attractive for being non-radioactive, and for providing real-time imaging. The cost of medical ultrasonography equipment is also reduced compared to that of other imaging devices, which further motivates the development of ultrasound-based imaging solutions for many clinical applications. A major limitation for the use of ultrasound is the quality of the images, which remains largely dependent on the operator's experience. Indeed, the proper positioning of the probe is a

critical factor to obtain a good image quality. The incidence angle of the ultrasound beam, the contact force, and also the distribution of ultrasound gel can significantly impact the quality of the acquired images. In particular, finding the optimal acoustic window to observe a location behind a strongly shadowing object, such as bone or gas, can be challenging. Moreover, manual manipulation of an ultrasound probe during a prolonged examination is cumbersome. Several studies have reported a prevalence of musculoskeletal disorders among sonographers [3], [4].

Robotized ultrasonography is a promising way to overcome the aforementioned limitations. Indeed, the use of a robot removes the need to manually hold the probe, and can improve the repeatability of examinations. The first robotic ultrasound systems were developed for the purpose of artery screening, in order to detect atheromatous plaques [5]–[7]. In parallel, several systems for robotized tele-echography have been designed [7]–[12], to allow specialists to remotely examine patients located in medically isolated sites. These systems use force control to limit or maintain constant the force applied to the patient. To further assist the sonographer, image-guided control methods have also been developed. Abolmaesumi *et al.* [10] control the in-plane motion of an ultrasound probe, in order to compensate for undesired motion of tracked image features. Out-of-plane control of a 2D probe can also be achieved using image moments [13], ultrasound speckle correlation [14], or block matching [15]. In [16], an intensity-based visual servoing method is proposed to control a 2D or a 3D probe. Ultrasound-based visual servoing can be used for motion compensation, but also in a shared control framework for tele-echography [17].

A particularity of ultrasound imaging is that the image content depends not only on the pose of the probe relative to the object of interest, but also on the quality of the acoustic coupling between the probe and the patient’s skin. In addition, the complete path traveled by the ultrasound beam between the transducer and a given point influences the quality of the image at this point, due mainly to attenuation and shadowing. This leads to the concept of *acoustic window*, which can be defined as the area between the ultrasound transducer and the target anatomy. Depending on the presence of acoustic attenuators in this area, different acoustic windows (i.e., different positions and orientations of the transducer) can yield different imaging quality levels for a same target. Finding a good acoustic window is crucial to get a satisfying image quality. While force and position control of ultrasound probes has been extensively studied, an adaptive control of image quality is still lacking. Such control first requires a way to quantify the quality of the ultrasound signal. Statistical quality measures, such as entropy, which are commonly used in image analysis, can also be applied to ultrasound images [18]. Kuhlemann [19] uses entropy as a measure of signal quality to detect acoustic shadows, and subsequently adapts the orientation of an ultrasound probe held by a robotic manipulator. However, such statistical quality measures are quite general, and they ignore the specific properties of ultrasound imaging. Another notable example is [20], where a measure of ultrasound image quality is learned from expert ratings, and used to optimize imaging

parameters. But the quality evaluation remains global, at the image level. Hellier *et al.* [21] propose a method to detect acoustic shadows within ultrasound images. This method, based on a statistical analysis of ultrasound scan-lines and a model of ultrasound noise, provides a binary mask indicating shadowed regions. The *ultrasound confidence map* framework, recently introduced by Karamalis *et al.* [22], allows the estimation of a continuous pixelwise measure of quality for ultrasound images. The confidence map is computed from the images, based on a graphical model which simulates the physical properties of sound propagation in soft tissues. It can be estimated in real-time, which is crucial when considering control applications. The confidence map is an efficient tool to detect acoustic shadows, and it can be used to improve image reconstruction or mosaicing [23]. In [24], the authors propose a visualization scheme where the confidence map is merged with the ultrasound image. Results from a user study with ultrasound experts show that the confidence map visualization helps with (i) correct interpretation and (ii) optimizing the acoustic window.

In this paper, we present a confidence-driven control framework for robotized ultrasound probes. A servoing purely based on ultrasound images would ignore the physics of ultrasound and would not be able to optimize the acoustic window. A knowledgeable and experienced sonographer becomes aware of the physics behind the acquisition. Thus, he/she can make sure of correct contact to the body surface and find the best position and orientation of the probe to achieve an optimal ultrasound image quality. In order to imitate such intelligent expert behavior, we use the ultrasound confidence map as a signal to design a servo control task. To this end, we define specific features extracted from the confidence map, and we propose an analysis of their variations with respect to the probe’s motion. The resulting *confidence-driven control* is the main contribution of this paper. Then, we use the redundancy framework [25] to combine the confidence control task with other functional tasks (force control, target tracking), in a hierarchical manner. We show, through robotic experiments with both 2D and 3D probes, how our control framework can be effectively used to optimize the imaging quality (i) in tele-echography, with a shared human/robot control, or (ii) during the autonomous tracking of a region of interest.

The methods presented herein are built on preliminary work presented in [1], [2]. We introduce several novel contributions. As a major contribution, the methods previously introduced for the in-plane control of a 2D probe are extended to the full control of a 3D probe. This enables a complete control of the probe for the optimization of image quality. In Addition, we also present new experimental results to validate the proposed approach. In particular, we introduce a measure of the quality of target tracking. By comparing the values taken by this measure in experiments with and without confidence control, we show that confidence control increases the robustness of target tracking. Moreover, we report the results of experiments performed on a human volunteer. These results show that the proposed framework is applicable in real conditions.

This paper is organized as follows. We first introduce the required background on ultrasound imaging and control. The

force and image control are briefly addressed in Sections II-B3 and II-B4, respectively. Then, we expose our confidence-driven control framework in Section III. Experimental results are presented in Section IV.

## II. ROBOTIZED ULTRASONOGRAPHY

### A. Ultrasound Imaging

Ultrasound images are generated by sending focused ultrasound waves into the body and listening to the echoes reflected by tissue interfaces. For B-mode imaging, which is the most popular ultrasound mode, the electric signals generated by these echoes are processed to form a 2D gray-level image. The time delay between the emission and the reception of an ultrasonic wave is proportional to the distance traveled by the ultrasound. Therefore, geometrically, the B-mode image corresponds to a section of the body. For example, the field of view of a convex transducer can be represented in polar coordinate by the region  $\Omega_{2D} = [r_{min}, r_{max}] \times [\theta_{min}, \theta_{max}]$ , as represented in Fig. 1(left). We note  $O$  the imaging center, that is, the intersection of all scan lines. Then  $r_{min}$  is the radius of the transducer, and  $r_{max}$  is the distance between  $O$  and the farthest imaged points. The imaging depth is defined as  $d = r_{max} - r_{min}$ . The angles  $\theta_{max}$  and  $\theta_{min}$  correspond to the lateral limits of the field of view. We adopt the convention that  $\theta_{min} = -\theta_{max}$ , so that the line  $\theta = 0$  corresponds to the central scan line.

Some ultrasound probes, with a motorized transducer or a matrix array transducer, can also image a volume. For a motorized convex probe, the field of view can be noted  $\Omega_{3D} = [r_{min}, r_{max}] \times [\theta_{min}, \theta_{max}] \times [\phi_{min}, \phi_{max}]$ , where the additional coordinate  $\phi$  represents the rotation around the motor's axis. In Fig. 1, the intersection between the motor's axis and the central scan line is represented by the point  $M$ . Similar to the 2D case, we adopt the convention that  $\phi_{min} = -\phi_{max}$ , so that the plane  $\phi = 0$  corresponds to the central frame. Note that, in practice, the continuous sweeping motion of the transducer induces a distortion of the acquired frames [26]. For the sake of simplicity, we omit the effect of the distortion in this paper. However, it should be taken into account if a geometrically precise reconstruction of the acquired volumes is needed. Up to the distortion effect, the geometry of the acquired volume corresponds to a rigid mapping of the physical scene. In this paper, we consider, unless otherwise stated, the case of a 3D ultrasound probe, which is the most complete one. The equations for the in-plane control of a 2D probe can be easily deduced by setting  $\phi = 0$ . We denote  $I: \Omega \rightarrow \mathbb{R}$  the ultrasound image (or volume), with  $\Omega = \Omega_{2D}$  or  $\Omega_{3D}$  according to the context. For a point  $\mathbf{p} \in \Omega$ ,  $I(\mathbf{p})$  is the image intensity at  $\mathbf{p}$ .

### B. Ultrasound Probe Control

We consider an ultrasound probe manipulated by a robotic system. The specific design of the system is beyond the scope of this paper. We rather propose a generic control strategy for the probe placement, which could be applied to different robotic systems, depending on the application. Possible designs are, for instance, fixed-base robotic arms [6], patient-mounted robots [11], [27], or hand-held robotic devices [28].

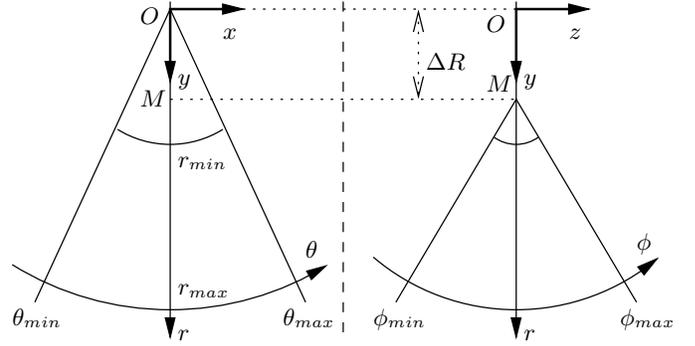


Fig. 1. Imaging geometry of a three-dimensional convex ultrasound probe. (left) Front view, also corresponding to the imaging geometry of a 2D probe. (right) Side view, showing the range of the motor.

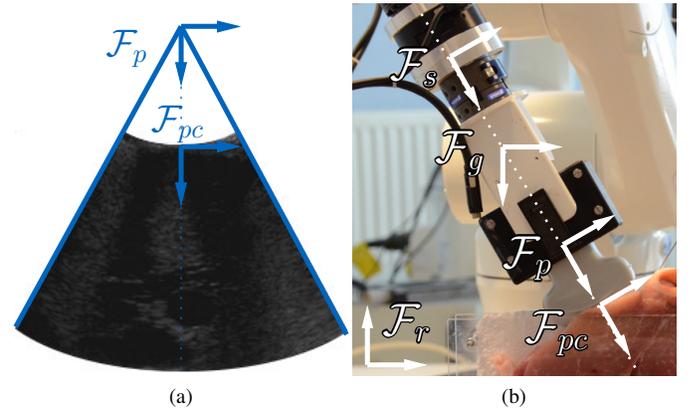


Fig. 2. Coordinate frames attached to the ultrasound probe, represented (a) on the ultrasound image and (b) on a picture of the robotic system.

Our only assumptions are that (i) the robotic system can fully control the pose of the probe, and (ii) it provides a measure of the external forces and torques exerted at its end-effector.

1) *Coordinate Frames*: Let us consider the example of a robotic manipulator with at least 6 degrees of freedom, equipped with a force sensing device. The ultrasound probe is rigidly attached to the end-effector of the manipulator. We define the following frames linked to the robot (see Fig. 2):

- The *probe frame*  $\mathcal{F}_p$ , attached to the imaging center  $O$  of the convex probe, such that its  $y$ -axis coincides with the central scan line, oriented towards the object, and the plane  $(O, x, y)$  corresponds to the central frame.
- The *contact frame*  $\mathcal{F}_{pc}$ , attached to contact point between the probe and the patient, and oriented like  $\mathcal{F}_p$ .
- The *sensor frame*  $\mathcal{F}_s$ , in which the measured force/torque tensor is expressed, and oriented as well like  $\mathcal{F}_p$ .
- The *gravity frame*  $\mathcal{F}_g$ , attached to the probe's center of mass, with its  $y$ -axis in the direction of the gravity force.
- The *reference frame*  $\mathcal{F}_r$ , a fixed frame attached to the environment.

For any two frames  $\mathcal{F}_a$  and  $\mathcal{F}_b$ , we note

$${}^a\mathbf{F}_b = \begin{bmatrix} {}^a\mathbf{R}_b & \mathbf{0}_{3 \times 3} \\ [{}^a\mathbf{t}_b]_{\times} & {}^a\mathbf{R}_b \end{bmatrix} \quad (1)$$

the twist transformation matrix by which one can express in

$\mathcal{F}_a$  a force/torque vector known in  $\mathcal{F}_b$ .  ${}^a\mathbf{R}_b \in \text{SO}_3$  is the matrix of rotation from  $\mathcal{F}_a$  to  $\mathcal{F}_b$ ,  ${}^a\mathbf{t}_b \in \mathbb{R}^3$  is the translation vector from  $\mathcal{F}_a$  to  $\mathcal{F}_b$ , and  $[{}^a\mathbf{t}_b]_\times \in \mathbb{R}^{3 \times 3}$  is skew-symmetric matrix built from  ${}^a\mathbf{t}_b$ .

2) *Control Tasks*: Let  $\mathbf{s} \in \mathbb{R}^p$  be a  $p$ -dimensional vector (with  $p \geq 1$ ) of sensor features to be regulated, and  $\mathbf{v}$  the instantaneous velocity screw of the sensor relative to the environment, expressed in the probe frame  $\mathcal{F}_p$ . The relation between  $\mathbf{v} = (v_x \ v_y \ v_z \ \omega_x \ \omega_y \ \omega_z)^\top$  and the variations of  $\mathbf{s}$  can be represented by an interaction matrix  $\mathbf{L}_s \in \mathbb{R}^{p \times 6}$  [29] such that

$$\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v}. \quad (2)$$

Letting  $\mathbf{s}^*$  be the desired value of the feature vector, the feature error is defined as  $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$ . To obtain an exponential decrease of this error, we define the desired error variation as  $\dot{\mathbf{e}}^* = -\lambda \mathbf{e}$ . The general form of a control law realizing this task is

$$\mathbf{v} = \mathbf{L}_s^g \dot{\mathbf{e}}^* + \mathbf{v}', \quad (3)$$

where  $\mathbf{L}_s^g$  is a generalized inverse of  $\mathbf{L}_s$ , and  $\mathbf{v}' \in \ker(\mathbf{L}_s)$ .

When  $\ker(\mathbf{L}_s) \neq \{\mathbf{0}\}$ , the robotic system is said redundant with respect to the task to realize, and  $\mathbf{v}'$  can be adapted to fulfill a secondary task without disturbing the primary task. This property is exploited in the *redundancy formalism* [25], which allows one to build a control law realizing a series of tasks with a given order of priority. Given a stack of tasks  $(\mathbf{e}_i^*, \mathbf{L}_i)$ , with  $i \in [0, N]$ , ordered according to priority, such a control law can be constructed recursively as:

$$\begin{aligned} \mathbf{v}_0 &= \mathbf{0} \\ \mathbf{v}_i &= \mathbf{v}_{i-1} + \mathbf{P}_{i-1} (\mathbf{L}_i \mathbf{P}_{i-1})^g (\dot{\mathbf{e}}_i^* - \mathbf{L}_i \mathbf{v}_{i-1}), \end{aligned} \quad (4)$$

where  $\mathbf{P}_{i-1}$  is a projection operator on the null-space of  $(\mathbf{L}_0, \dots, \mathbf{L}_{i-1})$ , defined as:

$$\begin{aligned} \mathbf{P}_0 &= \mathbf{I}_6 \\ \mathbf{P}_i &= \mathbf{P}_{i-1} - \mathbf{P}_{i-1} (\mathbf{L}_i \mathbf{P}_{i-1})^g \mathbf{L}_i \mathbf{P}_{i-1}. \end{aligned} \quad (5)$$

3) *Force Control*: Controlling the contact force applied to the body is crucial for the patient's safety. It is also an important factor for the quality of the acquired images, which requires a good contact. Therefore, we define the force control as the task of highest priority. We recall in this section the force control approach proposed in [16]. This approach has the merit of integrating well with the visual servoing formulation. However, other force control methods could be considered as well. We assume that the force sensing system provides a measure of the external forces  $\mathbf{f}$  and torques  $\boldsymbol{\tau}_s$ , expressed in the sensor frame  $\mathcal{F}_s$ , which we represent with a 6-dimensional vector  ${}^s\mathbf{H}_s = (\mathbf{f}^\top \ \boldsymbol{\tau}_s^\top)^\top$ . Given the  $6 \times 6$  twist transformation matrices  ${}^{pc}\mathbf{F}_s$  and  ${}^s\mathbf{F}_g$  allowing the transformation of a force/torque vector respectively from the sensor frame  $\mathcal{F}_s$  to the probe contact frame  $\mathcal{F}_{pc}$ , and from the probe's inertial frame  $\mathcal{F}_g$  to  $\mathcal{F}_s$ , the force/torque vector in the contact frame can be written

$${}^{pc}\mathbf{H}_{pc} = {}^{pc}\mathbf{F}_s ({}^s\mathbf{H}_s - {}^s\mathbf{F}_g {}^g\mathbf{H}_g), \quad (6)$$

where  ${}^g\mathbf{H}_g \in \mathbb{R}^6$  is the gravity force, expressed in  $\mathcal{F}_g$ , that is

$${}^g\mathbf{H}_g = (0 \ 0 \ m_p g \ 0 \ 0 \ 0)^\top, \quad (7)$$

with  $m_p$  the mass of the probe, and  $g \approx 9.81 \text{ m} \cdot \text{s}^{-2}$  the standard acceleration due to gravity.

As we are interested in controlling the force only along the  $y$ -axis of the probe, we define the force feature as the  $y$  component of the force tensor in the probe contact frame:

$$s_f = \mathbf{S}_y {}^{pc}\mathbf{H}_{pc}, \quad (8)$$

where  $\mathbf{S}_y = (0 \ 1 \ 0 \ 0 \ 0 \ 0)$ . Considering soft tissues as an elastic body of constant stiffness  $k$ , the interaction matrix for the force feature  $s_f$  can be approximated as

$$\mathbf{L}_f = [0 \ k \ 0 \ 0 \ 0 \ 0]. \quad (9)$$

Let us define a constant desired contact force  $s_f^*$ , and the force feature error  $e_f = s_f - s_f^*$ . The formulation given by (3) leads to the following control law for the force control task:

$$\mathbf{v}_f = -\frac{\lambda_f}{k} (\mathbf{S}_y {}^{pc}\mathbf{F}_s ({}^s\mathbf{H}_s - {}^s\mathbf{F}_g {}^g\mathbf{H}_g) - s_f^*) \mathbf{S}_y^\top, \quad (10)$$

where  $\lambda_f$  is the force control gain. This control law requires an estimation of the probe's mass  $m_p$ , the transformations  ${}^{pc}\mathbf{F}_s$  and  ${}^s\mathbf{F}_g$ , and the contact stiffness  $k$ .

4) *Ultrasound-Based Visual Servoing*: Due to the specific geometry of a B-mode frame, ultrasound is usually considered as a visual modality, and it can be treated as such in a control framework. A major difference with a camera projection model is that a B-mode frame corresponds to a planar section of the environment. In the particular case of a 3D ultrasound transducer, the reconstructed volume corresponds to a scaled version of the scene, so that position-based visual servoing can be applied directly. Let us assume that a tracking algorithm provides the position of a physical target  $\mathbf{t} = (x_t, y_t, z_t)$  in the probe frame, and that this target is immobile in the reference frame. Then, the interaction matrix that relates the variations of the target's position to the velocity screw  $\mathbf{v}$  of the probe can be written

$$\mathbf{L}_t = \begin{bmatrix} -1 & 0 & 0 & 0 & -z_t & y_t \\ 0 & -1 & 0 & z_t & 0 & -x_t \\ 0 & 0 & -1 & -y_t & x_t & 0 \end{bmatrix}. \quad (11)$$

A target centering task can be formalized as the regulation of  $\mathbf{s}_t = (x_t \ z_t)^\top$  to  $\mathbf{s}_t^* = \mathbf{0}$ . Note that the  $y$ -axis corresponds to the depth direction, so that it is not desirable to constrain the  $y$ -coordinate of the target. The interaction matrix for the target centering task is simply deduced from (11) as

$$\mathbf{L}_t = \begin{bmatrix} -1 & 0 & 0 & 0 & -z_t & y_t \\ 0 & 0 & -1 & -y_t & x_t & 0 \end{bmatrix}, \quad (12)$$

which defines a virtual linkage of class 4 [29]. Among the possible generalized inverses of  $\mathbf{L}_t$ , we can choose arbitrarily

$$\mathbf{L}_t^g = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}^\top, \quad (13)$$

which corresponds to a pure translation motion. Then, noting  $\mathbf{e}_t = \mathbf{s}_t - \mathbf{s}_t^*$  the visual feature error, (3) leads to

$$\mathbf{v}_t = -\lambda_t \mathbf{L}_t^g \mathbf{e}_t. \quad (14)$$

### III. CONFIDENCE-DRIVEN CONTROL

#### A. Objectives

We define two different scenarios for confidence-driven control. In the global scenario, the confidence-driven control aims at maintaining an orientation which globally optimizes the quality of the acquired images, without considering any particular anatomy. A possible application of this scenario is tele-echography, where the control of the probe is shared with a clinician. In this example, the goal of the confidence-control is to ease the manipulation of the probe by automating its orientation, in order to ensure a good image quality. The tasks involved in this scenario can be summarized as follows, in decreasing order of priority:

- 1) Maintaining a constant contact force.
- 2) Optimizing the quality of the ultrasound image.
- 3) Applying the commands of the operator.

In the target-specific scenario, the confidence-driven control aims at optimizing the image quality for a specific anatomical target. In this scenario, we can consider applications where the ultrasound acquisition is fully automatic, once the target is defined. The tasks involved in this scenario can be defined as follows, in decreasing order of priority:

- 1) Maintaining a constant contact force.
- 2) Maintaining the target centered in the image.
- 3) Optimizing the image quality at the target location.

#### B. Ultrasound Confidence Maps

Due to the specific process of ultrasound image formation, the resulting images are not uniform in terms of quality. More precisely, ultrasound waves undergo acoustic attenuation, absorption and diffraction when traveling through biological tissues. The variation of image quality can be particularly important when an abrupt change in acoustic impedance occurs on the path of the wave. A commonly observed artifact is *shadowing*, which corresponds to an important diminution of the amplitude of the ultrasound wave due to the presence of a strong attenuator. We represent the quality of the image by a *confidence map*  $C : \Omega \rightarrow [0, 1]$ , such that, for a pixel  $\mathbf{p} \in \Omega$ ,  $C(\mathbf{p})$  is a measure of the confidence in the image intensity  $I(\mathbf{p})$ . Note that the purpose of the confidence map is to estimate differences in signal quality within the image, rather than to evaluate the quality of the image as a whole.

1) *Graphical Representation*: Karamalis *et al.* [22] define  $C(\mathbf{p})$  as the probability that an echo originating from  $\mathbf{p}$  reaches one of the transducer elements. This probability map is inferred from a graphical representation of the image (Fig. 3). Specifically, a graph is constructed, whose nodes represent the ultrasound image samples, and edges are weighted according to ultrasound propagation constraints. Thus, the graph is a simplified model of the physics of ultrasound propagation in soft tissues. In particular, it models attenuation and absorption, which are the main phenomena governing signal quality. For two adjacent samples  $\mathbf{p}, \mathbf{q}$  along a scan line, the propagation weight between  $\mathbf{p}$  and  $\mathbf{q}$  is defined as

$$w_r(\mathbf{p}, \mathbf{q}) = \exp(-\beta|\check{I}(\mathbf{p}) - \check{I}(\mathbf{q})|), \quad (15)$$

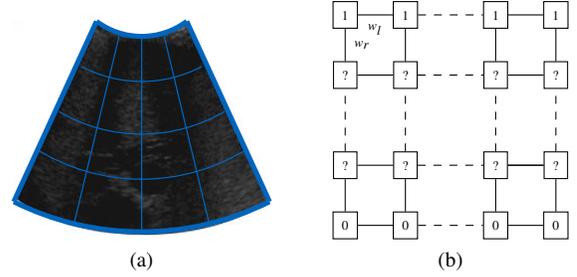


Fig. 3. Graph layout for the confidence map computation. (a) Ultrasound image with the scan lines geometry highlighted in blue. (b) The corresponding 4-connected lattice. The first row corresponds to the pixels which are the closest to the transducer array. Their confidence is set to 1. The last row corresponds to the pixels which are the farthest away from the transducer array. Their confidence is set to 0. Each edge is given a weight, which constrains the ultrasound propagation along the radial ( $w_r$ ) and lateral ( $w_l$ ) directions.

where  $\check{I}$  is the attenuated intensity, defined as

$$\check{I}(\mathbf{p}) = I(\mathbf{p}) \exp(-\alpha d(\mathbf{p})), \quad (16)$$

$d(\mathbf{p})$  being the depth of  $\mathbf{p}$ . The parameters  $\alpha$  and  $\beta$  control the importance of acoustic attenuation and absorption, respectively. The ultrasound beam width is modeled by allowing lateral propagation as well. For two samples  $\mathbf{p}, \mathbf{q}$  at the same depth and on adjacent scan lines, the propagation weight between  $\mathbf{p}$  and  $\mathbf{q}$  is defined as

$$w_l(\mathbf{p}, \mathbf{q}) = \exp(-\beta(|\check{I}(\mathbf{p}) - \check{I}(\mathbf{q})| + \gamma\|\mathbf{p} - \mathbf{q}\|)), \quad (17)$$

where  $\gamma$  is a parameter related to the beam width.

2) *Resolution*: Once the propagation graph is constructed, the confidence estimation problem is solved using the random walks algorithm [30]. This algorithm requires boundary conditions, which are defined as follows:

- The confidence is 1 at the minimal depth.
- The confidence is 0 at the maximal depth.

Let  $n$  be the number of pixels in the image. We note  $(\mathbf{p}_i)_{i \in [1, n]}$  the list of pixels (in an arbitrary order). The weights of (15) and (17) are used to define the graph Laplacian matrix  $\mathbf{\Lambda} \in \mathbb{R}^{n \times n}$  such that, for  $(i, j) \in [1, n]^2$ ,

$$\Lambda_{i,j} = \begin{cases} \sum_{\mathbf{q} \in \mathcal{N}(\mathbf{p}_i)} w(\mathbf{p}_i, \mathbf{q}) & \text{if } i = j \\ -w(\mathbf{p}_i, \mathbf{p}_j) & \text{if } \mathbf{p}_j \in \mathcal{N}(\mathbf{p}_i) \\ 0 & \text{otherwise} \end{cases}, \quad (18)$$

where  $w(\mathbf{p}, \mathbf{q})$  refers to  $w_r$  or  $w_l$  according to the configuration of  $\mathbf{p}$  and  $\mathbf{q}$ , and  $\mathcal{N}(\mathbf{p})$  is the set of neighbors of  $\mathbf{p}$ . Let  $\mathbf{c} \in \mathbb{R}^n$  be a vector representation of the confidence map. We assume that  $\mathbf{c}$  is ordered in such a way that known ( $K$ ) values appear before unknown ( $U$ ) values:  $\mathbf{c} = (\mathbf{c}_K^\top \ \mathbf{c}_U^\top)^\top$ . Thus, the Laplacian matrix can be written

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_K & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{\Lambda}_U \end{bmatrix}, \quad (19)$$

where  $\mathbf{\Lambda}_K$  is the Laplacian matrix of the subgraph consisting of nodes with known confidence values only,  $\mathbf{\Lambda}_U$  is the Laplacian matrix of the subgraph consisting of nodes with unknown confidence values, and  $\mathbf{B}$  corresponds to the transitions between

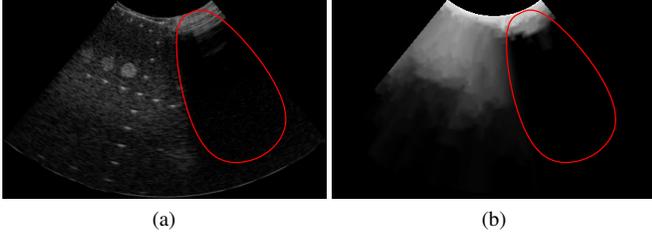


Fig. 4. Drop-out artifact (in red), induced by a weak contact between the ultrasound probe and the skin. (a) Ultrasound image. (b) Confidence map.

these subgraphs. The confidence map is obtained by solving for  $\mathbf{c}_U$  the sparse linear system

$$\mathbf{A}_U \mathbf{c}_U = -\mathbf{B}^\top \mathbf{c}_K, \quad (20)$$

as described in [30].

### C. Confidence Variations

The quality  $C(\mathbf{p})$  at a given pixel  $\mathbf{p}$  depends on the path traveled by the sound to reach  $\mathbf{p}$ , and not only on local tissues properties. There are two important examples which show the dependence of the confidence to the global situation:

- When a drop-out artifact occurs (Fig. 4). That is, when the ultrasound signal is lost due to a bad acoustic coupling (weak contact or a lack of gel).
- When a strongly shadowing object, such as a bone, is present in the image, inducing a local drop in the ultrasound signal confidence.

In these cases, re-orienting the probe can improve the quality of the ultrasound image.

Unlike the case of target position control, we do not have an analytic way to link the variations of the confidence map to the velocity screw. Therefore, we propose to construct a specific feature, for which we can approximate the interaction matrix. We start with some general observations before to detail the choice of the confidence feature:

- A rotation around the imaging center of the probe is equivalent to a shift in the scan lines, so that, under a pure rotation around the imaging center, the confidence map remains unchanged, up to a translation in polar space.
- A translation in the direction normal to the surface influences the global quality of the image, but does not influence the angular distribution of the confidence.
- Translations at the patient's surface are more difficult to predict. When scanning a relatively uniform region, surface translations will not have a large impact on the confidence map. However, in the presence of tissues with different acoustic properties, the impact of such translations could be larger. In practice, the influence of the surface contact on the quality of the image is more important than that of internal interfaces, so that the effect of surface translations can be ignored in the design of the control law.

1) *2D Case:* Let us first consider the acquisition of a single ultrasound frame by a convex probe. An in-plane rotation around the imaging center is equivalent to a shift in the scan lines, so that, neglecting the effect of translations,

$$\dot{C}(r, \theta, t) = \omega_z \frac{\partial C}{\partial \theta}(r, \theta, t). \quad (21)$$

It follows that a good feature to represent the distribution of the confidence should depend only on the angular direction. A natural choice would be to use the direction in which the confidence is maximum. However, this feature would be unstable, due to possible discontinuities of the argmax function. Instead, we consider the angular coordinate of the confidence-weighted barycenter, which consists in a weighted average of all values of the confidence map, and is therefore more stable. This angular moment can be defined as

$$\theta_c = \frac{1}{C_\Omega} \iint_{(r, \theta) \in \Omega} \theta C(r, \theta, t) r dr d\theta, \quad (22)$$

where  $C_\Omega = \iint_{(r, \theta) \in \Omega} C(r, \theta) r dr d\theta$ . Thus, if we assume that the mean image confidence is constant ( $\dot{C}_\Omega = 0$ ), we obtain from (21) that, as shown in the Appendix,

$$\dot{\theta}_c = -\omega_z. \quad (23)$$

2) *3D Case:* We now consider a wobbler probe, which sweeps the transducer back and forth around the motor's axis. Given the geometry of the probe, the Cartesian coordinates  $(x, y, z)$  of a point in the probe frame  $\mathcal{F}_p$  can be expressed in pseudo-spherical coordinates  $(r, \theta, \phi)$ , such as [26]

$$x = r \sin \theta, \quad (24)$$

$$y = (r \cos \theta - \Delta R) \cos \phi + \Delta R, \quad (25)$$

$$z = (r \cos \theta - \Delta R) \sin \phi, \quad (26)$$

where  $\Delta R$  is the offset between the imaging center  $O$  and the axis of the probe's motor (see Fig. 1). Note that a rotation around the motor axis is equivalent to a shift in the frames. Similar to (22), we compute the pseudo-spherical coordinates of the confidence-weighted barycenter in 3D:

$$\theta_c = \frac{1}{C_\Omega} \iiint_{\Omega} \theta C(r, \theta, \phi, t) |\det(J)| dr d\theta d\phi, \quad (27)$$

$$\phi_c = \frac{1}{C_\Omega} \iiint_{\Omega} \phi C(r, \theta, \phi, t) |\det(J)| dr d\theta d\phi, \quad (28)$$

where

$$|\det(J)| = r \cos^2 \theta \cos^2 \phi (r \cos \theta - \Delta R) \quad (29)$$

is the determinant of the Jacobian of the transformation from Cartesian to pseudo-spherical coordinates, and  $C_\Omega = \iiint_{\Omega} C(r, \theta, \phi) |\det(J)| dr d\theta d\phi$ . Under the same assumptions as in the 2D case, the relation between the variations of  $\phi_c$  and the velocity screw in the probe frame can be written:

$$\dot{\phi}_c = \omega_x. \quad (30)$$

#### D. General Confidence Control

We first consider the teleoperation scenario, where no specific target is defined. In this case, the confidence control task is to maintain an optimal orientation for the probe, and the control of its position is left to the clinician. To this end, we define the confidence feature simply as

$$\mathbf{s}_c = \begin{pmatrix} \theta_c \\ \phi_c \end{pmatrix}, \quad (31)$$

and its desired value  $\mathbf{s}_c^* = \mathbf{0}$ , corresponding to the central scan line. The goal behind this choice is to balance the distribution of the confidence, and to maintain a high confidence in the center of the volume. From (23) and (30), we can express the interaction matrix as

$$\mathbf{L}_c = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (32)$$

For teleoperation, the control tasks are, in decreasing order of priority, the force control task ( $\mathbf{e}_f^*, \mathbf{L}_f$ ), the confidence control task ( $\mathbf{e}_c^*, \mathbf{L}_c$ ), and the execution of the clinician's commands ( $\mathbf{e}_{op}^*, \mathbf{L}_{op}$ ). From (4), the full control screw  $\mathbf{v}_{gen}$  can be decomposed as

$$\mathbf{v}_{gen} = \mathbf{v}_f + \mathbf{v}_c + \mathbf{v}_{op}, \quad (33)$$

where the force control screw component  $\mathbf{v}_f$  is defined in (10). According to (9), the projection operator on  $\ker(\mathbf{L}_f)$  is

$$\mathbf{P}_f = \text{diag}(1, 0, 1, 1, 1, 1), \quad (34)$$

and we see that the force and confidence tasks are decoupled, because  $\mathbf{L}_c \mathbf{P}_f = \mathbf{L}_c$  and  $\mathbf{L}_c \mathbf{v}_f = \mathbf{0}$ . This was expected, because the two tasks involve different axes. Therefore, the contribution of the confidence task to the control screw is simply

$$\mathbf{v}_c = \mathbf{L}_c^g \mathbf{e}_c^*, \quad (35)$$

where the generalized inverse of  $\mathbf{L}_c$  is defined as

$$\mathbf{L}_c^g = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^\top. \quad (36)$$

The projection operator on the null-space of  $(\mathbf{L}_f, \mathbf{L}_c)$  is  $\mathbf{P}_c = \text{diag}(1, 0, 1, 0, 1, 0)$ . Therefore, the operator controls the 3 remaining degrees of freedom, which correspond to the translations on the patient's surface and the rotation around the depth axis.

#### E. Target-Specific Confidence Control

We now consider the target tracking scenario, where the goal is to maintain the visibility of a specific anatomical target. The target is tracked in the ultrasound image, and we note the coordinates of its barycenter in the probe frame  $(x_t, y_t, z_t)$  in Cartesian coordinates, and  $(r_t, \theta_t, \phi_t)$  in pseudo-spherical coordinates. The task is to maintain a good image quality at the location of the target. Therefore, we define the feature vector for target-specific confidence control as

$$\mathbf{s}_c = \begin{pmatrix} \theta_c - \theta_t \\ \phi_c - \phi_t \end{pmatrix}, \quad (37)$$

which corresponds to the two angles between the target and the confidence barycenter. The desired value of this feature

vector is  $\mathbf{s}_c^* = \mathbf{0}$ . With this new choice of visual features, the interaction matrix  $\mathbf{L}_c$  has to be modified to account for the variations of  $\theta_t$  and  $\phi_t$ , so that

$$\mathbf{L}_c = \begin{bmatrix} -\cos \theta_t & 0 \\ \frac{\sin \theta_t}{r_t} \cos \phi_t & \frac{\sin \phi_t}{r_t \cos \theta_t - \Delta R} \\ \frac{r_t}{\sin \theta_t} \sin \phi_t & \frac{\cos \phi_t}{r_t \cos \theta_t - \Delta R} \\ \frac{\Delta R \sin \theta_t}{r_t} \sin \phi_t & -\frac{r_t \cos \theta_t - \Delta R}{\Delta R \cos \theta_t} \\ \frac{\Delta R \cos \theta_t - r_t}{r_t} \sin \phi_t & -\frac{r_t \cos \theta_t - \Delta R}{r_t \sin \theta_t \cos \phi_t} \\ \frac{\Delta R \cos \theta_t - r_t}{r_t} (1 - \cos \phi_t) & \frac{r_t \sin \theta_t \cos \phi_t}{r_t \sin \theta_t \sin \phi_t} \\ & \frac{r_t \cos \theta_t - \Delta R}{r_t \cos \theta_t - \Delta R} \end{bmatrix}^\top. \quad (38)$$

Note that, setting  $\phi_t = 0$  in the first line of  $\mathbf{L}_c$ , we obtain the interaction matrix for the case of a 2D probe, as in [2].

The three control tasks that we consider are, in decreasing order of priority, the force control task ( $\mathbf{e}_f^*, \mathbf{L}_f$ ), the target centering task ( $\mathbf{e}_t^*, \mathbf{L}_t$ ), and the confidence control task ( $\mathbf{e}_c^*, \mathbf{L}_c$ ). Therefore, the control screw  $\mathbf{v}_{spe}$  for the target-specific confidence control can be decomposed as

$$\mathbf{v}_{spe} = \mathbf{v}_f + \mathbf{v}_t + \mathbf{v}_c. \quad (39)$$

The target tracking component of the control screw is obtained from (4) as

$$\mathbf{v}_t = \mathbf{P}_f (\mathbf{L}_t \mathbf{P}_f)^g (\mathbf{e}_t^* - \mathbf{L}_t \mathbf{v}_f), \quad (40)$$

which simplifies to

$$\mathbf{v}_t = \mathbf{L}_t^g \mathbf{e}_t^*, \quad (41)$$

since this task is decoupled from the force control task. Using the generalized inverse  $\mathbf{L}_t^g$  defined in (13), the projection operator on  $\ker(\mathbf{L}_f, \mathbf{L}_t)$  is obtained from (5) as

$$\mathbf{P}_{f,t} = \begin{bmatrix} 0 & 0 & 0 & 0 & -z_t & y_t \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -y_t & x_t & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (42)$$

Then, the component of the velocity screw corresponding to the quality optimization task can be written:

$$\mathbf{v}_c = \mathbf{P}_{f,t} (\mathbf{L}_c \mathbf{P}_{f,t})^g (\mathbf{e}_c^* - \mathbf{L}_c (\mathbf{v}_f + \mathbf{v}_t)). \quad (43)$$

To simplify the resulting expression, let us consider the interaction matrix at the desired pose, as it is usually done in the literature [31]. We define

$$\mathbf{L}_c^* = \mathbf{L}_c \Big|_{\substack{\theta_t = 0 \\ \phi_t = 0}} = \begin{bmatrix} -\frac{1}{r_t} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta R - r_t} & \frac{\Delta R}{\Delta R - r_t} & 0 & 0 \end{bmatrix} \quad (44)$$

to replace  $\mathbf{L}_c$  in (43), and we obtain the final velocity screw for the stack of tasks:

$$\mathbf{v}_{spe} = \begin{pmatrix} \lambda_t x_t + \lambda_c (\theta_c - \theta_t) y_t \\ -\frac{\lambda_f e_f}{k} \\ \lambda_t z_t + \lambda_c (\phi_c - \phi_t) y_t \\ \lambda_c (\phi_c - \phi_t) \\ 0 \\ \lambda_c (\theta_t - \theta_c) \end{pmatrix} \quad (45)$$

Note that this is only a particular solution, to which one can add any element of the null space  $\ker(\mathbf{L}_f, \mathbf{L}_t, \mathbf{L}_c^*)$ , i.e., any scalar multiple of  $(-z_t \ 0 \ x_t \ 0 \ 1 \ 0)^\top$ .

### F. Region Tracking by Virtual Servoing

In Section III-E, we have assumed, that some algorithm provided the position of a target in the ultrasound image. We now expose the specific method used in this work. Considering the application, we are interested in a tracking algorithm that is fast and robust, rather than in a precise delineation of a target's contour. Therefore, we define a region of interest  $\mathcal{R}$  in the image, which we wish to maintain visible. Following the approach recently proposed in [32], we perform a *virtual servoing* of the region of interest. That is, we control the position of  $\mathcal{R}$  within the ultrasound volume, as one would do with a physical sensor. Using an intensity-based visual servoing approach, the visual feature is defined as

$$\mathbf{s}_{\text{track}} = (I(\mathbf{p}))_{\mathbf{p} \in G(\mathcal{R})}, \quad (46)$$

where  $G(\mathcal{R})$  is a regular grid over  $\mathcal{R}$ . According to [32], the interaction matrix for a component  $I(\mathbf{p})$  can be written

$$\mathbf{L}_{I(\mathbf{p})} = \begin{bmatrix} \nabla I_x & \nabla I_y & \nabla I_z \\ y\nabla I_z - z\nabla I_y & -x\nabla I_z + z\nabla I_x & x\nabla I_y - y\nabla I_x \end{bmatrix}, \quad (47)$$

where  $(x, y, z)$  are the physical coordinates corresponding to the point  $\mathbf{p}$ , expressed in a frame attached to the region of interest, and  $\nabla I_x, \nabla I_y, \nabla I_z$  are the image gradients along the directions  $x, y,$  and  $z$  at this point. The full interaction matrix  $\mathbf{L}_{\text{strack}}$  corresponding to the feature  $\mathbf{s}_{\text{track}}$  is obtained by stacking the interaction matrices for all points in  $G(\mathcal{R})$ . Then, the position of the region of interest is optimized by running a virtual servo loop with control law

$$\mathbf{v}_{\mathcal{R}} = -\lambda_{\text{track}} \mathbf{L}_{\text{strack}}^\dagger (\mathbf{s}_{\text{track}} - \mathbf{s}_{\text{track}}^*), \quad (48)$$

where  $\lambda_{\text{track}} \in \mathbb{R}^+$  is a control gain,  $\mathbf{L}_{\text{strack}}^\dagger$  is the Moore-Penrose pseudoinverse of  $\mathbf{L}_{\text{strack}}$ , and  $\mathbf{s}_{\text{track}}^*$  is the initial value of  $\mathbf{s}_{\text{track}}$ .

### G. Implementation Details

With a 2D ultrasound probe, B-mode frames can typically be acquired with a frame rate of about 30 Hz. This is comparable to the frame rate of a video camera, and it is sufficient to implement a visual control loop. With a wobbler probe, however, the acquisition of a volume takes about 1 s. Defining the visual control loop with this period leads to a poorly reactive system.

Instead, we propose to update the visual control loop continuously during the volume acquisition. Let us consider, for instance, the computation of the angular confidence moment  $\theta_c$  defined in (27). It involves the summation, over the volume, of the quantity  $\theta C(r, \theta, \phi, t) |\det(J)|$ . The confidence map  $C$  can be computed frame-by-frame, rather than from the complete volume. Then, storing in a buffer the current representation  $C_{\text{buffer}}$  of the volumetric confidence map, a new confidence frame  $C_k$  leads to the following update of  $\theta_c$ , for each  $i, j$ :

$$\theta_c \leftarrow \theta_c + \theta |\det(J(i, j, k))| (C_k(i, j) - C_{\text{buffer}}(i, j, k)), \quad (49)$$

whereupon the  $k$ -th frame of  $C_{\text{buffer}}$  is replaced with  $C_k$ .

## IV. RESULTS

### A. Experimental Setup

1) *Equipment*: For the experiments, we used a KUKA LWR iiwa R800 (KUKA Roboter GmbH, Germany), which has internal torque sensors in each joint. A 4DC7-3/40 convex ultrasound probe (Ultrasonix Medical Corporation, Canada) was rigidly attached to the robot's end-effector. For experiments on phantom, we used a CIRS Multi-Purpose Multi-Tissue Ultrasound Phantom (model 040GSE, Computerized Imaging Reference Systems, Inc (CIRS), USA). The ultrasound data were grabbed from the ultrasound scanner via Ethernet and processed on a workstation.

2) *Implementation*: The image processing and control algorithms were implemented in C++, using the Visual Servoing Platform (ViSP) library [33]. The application, with graphical user interface (GUI) and thread management, was implemented using the Qt library. The GUI contained widgets dedicated to the teleoperation of the robot by the user. Communication with the robot's controller was achieved through the Robot Operating System (ROS) software platform [34].

In order to achieve a reactive closed-loop control, we use multi-threading. The following components operate in parallel:

- The *grabber* thread is responsible for grabbing the prescan ultrasound images. When a new prescan image is available, the grabber sends it to the main thread. Thus, the visual servoing control law is updated at the same rate as the data acquisition rate.
- The *main* thread is in charge of user interaction, display, computation of the visual features and control law. It also takes care of the management of the other threads.
- The *converter* thread takes care of the conversion of ultrasound images and confidence map from prescan to postscan format.
- The *tracker* thread is responsible for target tracking.
- The *confidence* thread performs the estimation of the confidence map.
- The *control* thread implements the control of the robot. It translates the desired Cartesian control velocity into joint velocities, and it is in charge of force control. In order to have a responsive force control, the joint velocities are updated every 5 ms.

3) *Control Parameters*: For each control task, there is a gain parameter tunable by the user. These gains ( $\lambda_f, \lambda_t, \lambda_c$ ) govern the convergence rate and the stability of the system. In our experiments, satisfying gain values were found experimentally. Stable behavior at convergence without disturbance was obtained for  $\frac{\lambda_f}{k} \leq 2.0 \times 10^{-3} \text{ m} \cdot \text{N}^{-1} \cdot \text{s}^{-1}$ ,  $\lambda_t \leq 1.0 \text{ s}^{-1}$  and  $\lambda_c \leq 1.0 \text{ s}^{-1}$ . Note that these are the maximum gain values that we tested experimentally. The acceptable range of gain values also depends on the robot's properties and the low-level controller. If a model of the robot (inner loop dynamics and input/output delays) is available, it is possible to derive an upper bound on the gain values to ensure local stability. The interested reader can refer, e.g., to [35]. For security reasons, we limited the control velocities to  $5 \text{ mm} \cdot \text{s}^{-1}$  in translation, and  $5^\circ \cdot \text{s}^{-1}$  in rotation.

## B. Convergence Evaluation

We start with an experimental evaluation of the convergence of the proposed control laws in a static environment.

1) *Global Confidence-Driven Control*: First, let us consider the global confidence-driven control framework presented in Section III-D. We recall that the aim of this control framework is to globally optimize the image quality, by adjusting the orientation of the probe. No target tracking was performed in this experiment. In [1], we presented similar results for the in-plane control only. We reproduced the experiments with the complete control law for a 3D control of the probe. For this experiment, the probe was placed at an angle from the normal of the surface, barely in contact [Fig. 5(a)]. Then, the control law was activated, and we let the system converge in a static environment. The system systematically converged to such a position that the probe was roughly orthogonal to the phantom's surface [Fig. 5(b)]. With the confidence control gain  $\lambda_c = 0.8 \text{ s}^{-1}$ , the mean absolute final feature error were  $0.12^\circ$  for  $\theta_c$  and  $0.23^\circ$  for  $\phi_c$ . The evolution of the confidence feature errors and the contact force during convergence is represented in Fig. 6. For comparison, the angle between two consecutive scan lines is  $0.61^\circ$ , and the angle between two consecutive frames is  $1.46^\circ$ . As a result, the accuracy is of the same order of magnitude as the image resolution. Therefore, our system successfully regulates the angular features to their desired value with high accuracy, and the robotic system has the expected behavior.

2) *Target-Specific Confidence Control*: Let us now study the convergence of the target-specific confidence-driven control, which was presented in Section III-E. For these experiments, the probe was placed at an angle from the normal to the surface, such that the phantom contents was partly visible in the image [Fig. 7(a)]. Then, a region of interest was manually selected in the live ultrasound view, and the control law (45) was activated. The region of interest was tracked using the intensity-based approach described in Section III-F. The system converges as expected to such a position that the region of interest is centered in the field of view and the probe is roughly orthogonal to the surface normal [Fig. 7(b)].

In order to study the accuracy at convergence for different gain settings, we performed 5 series of 5 experiments each, with different values for  $\frac{\lambda_f}{k}$ ,  $\lambda_t$ , and  $\lambda_c$ . In each series, labelled A, B, C, D, E in Table I, the control gains were kept constant across experiments. The initial configuration for the experiments was such that the initial contact force ranged from 3.7 N to 4.6 N, the initial target feature error ranged from 4.3 cm to 5.6 cm, and the initial angular confidence feature error ranged from  $1.0^\circ$  to  $7.8^\circ$ . We report the mean and standard deviation of the absolute feature errors at convergence for each series in Table I.

We observe that the force control gain impacts the convergence accuracy of the contact force feature. On the other hand, the control gains for the target centering and confidence control tasks did not have any significant impact on the convergence accuracy of the associated features.

The evolution of the features during one of the experiments of series E is presented in Fig. 8. In this example, the

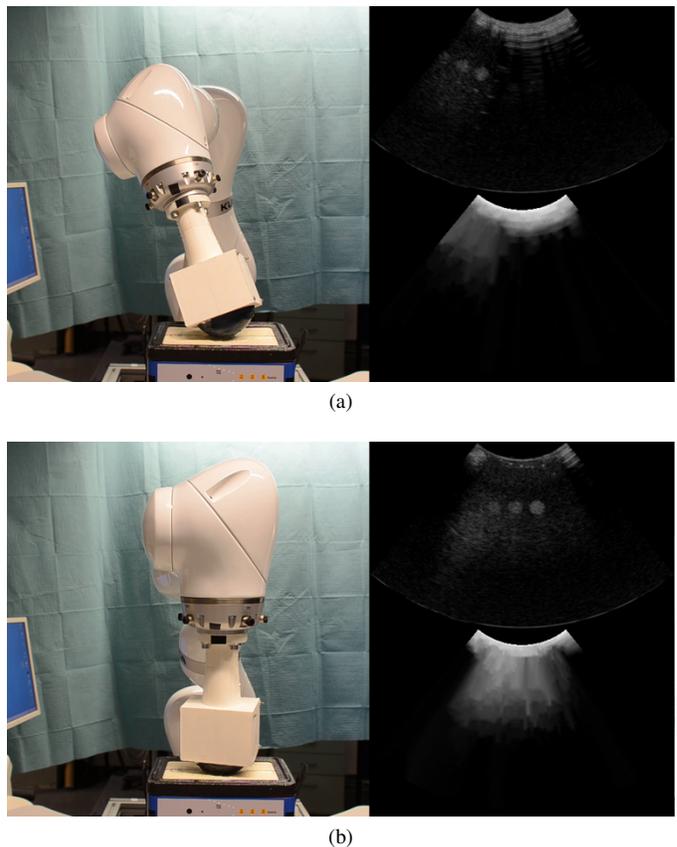


Fig. 5. 3D global confidence-driven control. (a) Initial configuration – the probe is not fully in contact with the phantom. (b) The system has converged, and the image quality is higher.

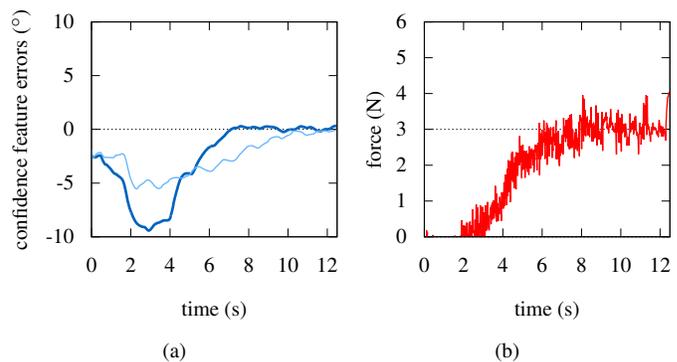


Fig. 6. Convergence of the global confidence-driven control in a static environment. (a) Confidence feature errors  $e_\theta$  (—) and  $e_\phi$  (—) in degrees. (b) Contact force  $s_f$  in Newtons.

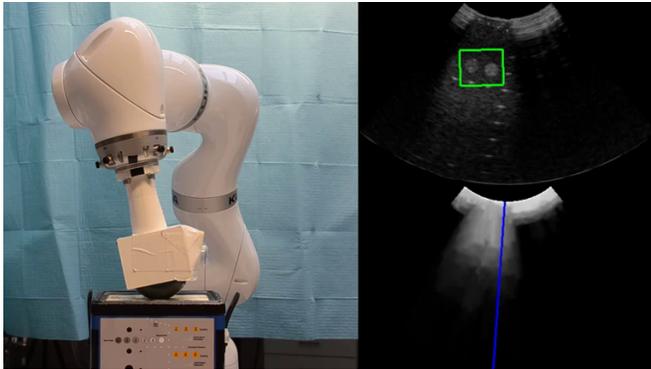
mean final feature errors were 0.19 N for the force control task,  $130 \mu\text{m}$  for the target centering task, and  $1.0^\circ$  for the confidence control task. Compared to the image resolution, these results indicate that our system can reach high accuracy simultaneously for target centering and confidence control.

## C. Reaction to Disturbances

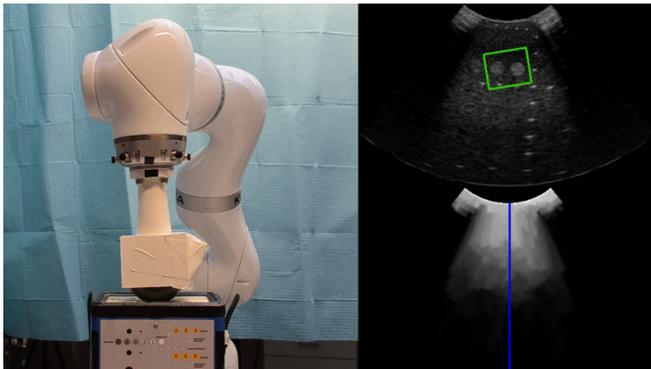
We now present the results of experiments aimed at testing the reaction of the system to disturbances such as shadowing or motion.

TABLE I  
ABSOLUTE FEATURE ERRORS AT CONVERGENCE FOR THE TARGET-SPECIFIC CONFIDENCE-DRIVEN CONTROL WITH DIFFERENT GAIN SETTINGS

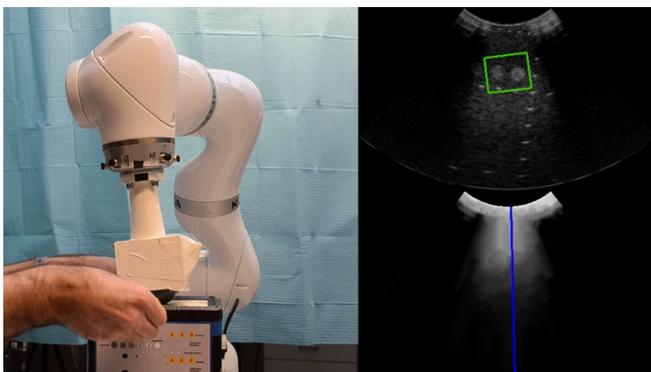
| series | $\frac{\lambda_f}{k}$ ( $\text{mm} \cdot \text{N}^{-1} \cdot \text{s}^{-1}$ ) | $\lambda_r$ ( $\text{s}^{-1}$ ) | $\lambda_c$ ( $\text{s}^{-1}$ ) | $ e_f $ (mN) | $ e_t $ ( $\mu\text{m}$ ) | $ e_c $ ( $^\circ$ ) |
|--------|---|---------------------------------|---------------------------------|--------------|---------------------------|----------------------|
| A      | 1.0   | 0.2                             | 0.3                             | $56 \pm 39$  | $60 \pm 25$               | $0.31 \pm 0.10$      |
| B      | 2.0   | 0.4                             | 0.4                             | $253 \pm 43$ | $67 \pm 30$               | $0.28 \pm 0.17$      |
| C      | 2.0   | 0.6                             | 0.6                             | $256 \pm 52$ | $74 \pm 51$               | $0.21 \pm 0.06$      |
| D      | 2.0   | 0.8                             | 0.8                             | $223 \pm 42$ | $67 \pm 44$               | $0.25 \pm 0.15$      |
| E      | 2.0   | 1.0                             | 1.0                             | $207 \pm 64$ | $80 \pm 38$               | $0.29 \pm 0.40$      |



(a)



(b)



(c)

Fig. 7. Target-specific confidence-driven control. (a) Initial configuration. (b) The system has converged. The target is centered, and the image quality is higher. (c) The system adapts to a manually generated shadow, while keeping the target centered.

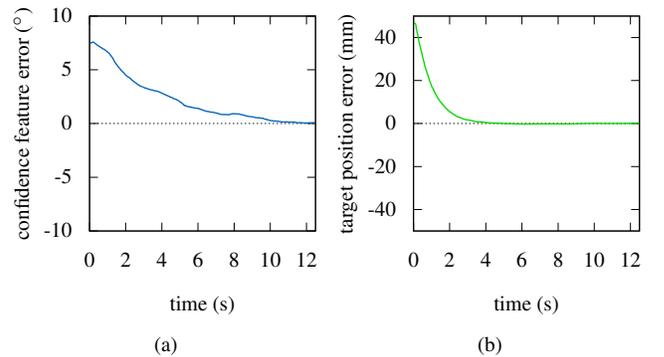


Fig. 8. Convergence of the target-specific confidence-driven control in a static environment. (a) Confidence feature error  $\theta_c$  in degrees. (b) Target centering error  $e_t$  in millimeters.

1) *Global Confidence-Driven Control*: We first describe an experiment illustrating the behavior of the global confidence-driven controller in the presence of shadows. To this end, a band of paper was manually inserted between the probe and the surface of the phantom, thus generating a strong shadow in the ultrasound image (Fig. 9). The band of paper was inserted and withdrawn on different sides of the probe. As a result, our system automatically reoriented the probe in order to avoid the shadowed area. We refer the reader to the supplementary video for a demonstration. The evolution of the confidence features is represented in Fig. 10, where the 4 gray areas correspond to the 4 shadow generations:

- At  $t = 21$  s, the band of paper was inserted on the right<sup>1</sup> [Fig. 5(c)], resulting in a negative shift of  $\theta_c$ , which triggered an in-plane rotation of the probe. Similarly, when the band of paper was withdrawn at  $t = 27$  s, we observed a positive shift of  $\theta_c$ , and the probe moved back to its original position. During this perturbation, the largest absolute deviation of the features was  $3.12^\circ$  for  $\theta_c$  and  $0.84^\circ$  for  $\phi_c$ .
- From  $t = 33$  s to  $t = 37$  s, the band of paper was inserted at the rear of the probe, inducing a perturbation of  $\phi_c$ .
- From  $t = 45$  s to  $t = 50$  s, the band of paper was inserted on the left of the probe, inducing a perturbation of  $\theta_c$ .
- From  $t = 57$  s to  $t = 61$  s, the band of paper was inserted at the front of the probe, inducing a perturbation of  $\phi_c$ .

The results of these 4 perturbations are summarized in Table II. We observe that the absolute deviation is significantly higher for the feature which corresponds to the direction of

<sup>1</sup>The terms *right*, *left*, *rear* and *front* refer to the viewpoint of Fig. 5.



Fig. 9. 3D global confidence-driven control – A shadow is generated using a paper band, which is inserted between the probe and the phantom by the experimenter. The system reacts automatically by reorienting the probe.

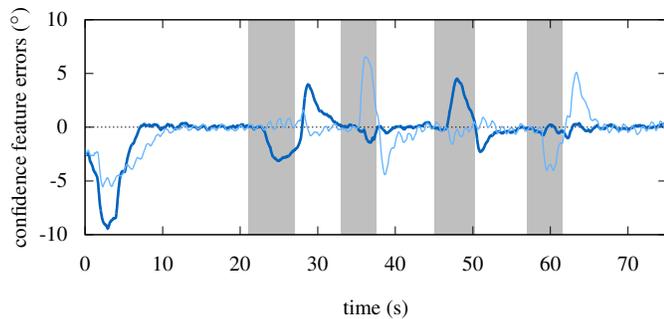


Fig. 10. Evolution of the confidence feature errors  $\theta_c$  (—) and  $\phi_c$  (---), in degrees, during a shadow avoidance experiment under global confidence-driven control.

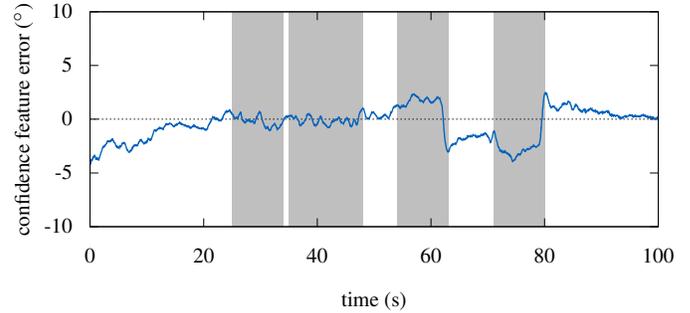
the perturbation (highlighted in bold). Thus, this experiment validates the discriminative power of the chosen features.

2) *Target-Specific Confidence Control*: Let us now consider the target-specific confidence-driven control. In order to test the tracking capability of the system, we disturbed its equilibrium by manually moving the phantom. In addition, we used the same method as previously, with a paper band, to generate shadows in the ultrasound image. We refer the reader to the video supplied as supplementary material for a demonstration of the experiment. The evolution of the confidence and target features is represented in Fig. 11, where the gray areas correspond to the periods of disturbance:

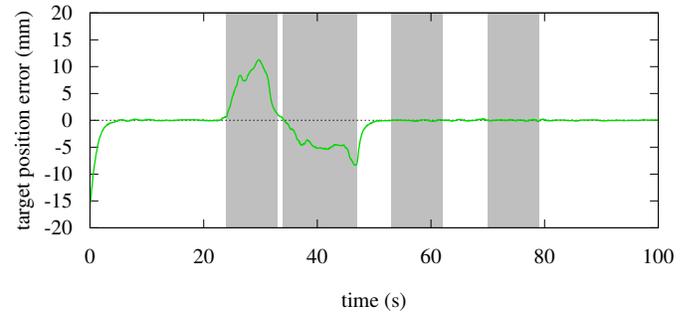
- **Motion**: The phantom was manually moved [Fig. 7(c)] forward ( $24\text{ s} < t < 33\text{ s}$ ) and backward ( $34\text{ s} < t < 47\text{ s}$ ). The robot automatically compensated by moving the

TABLE II  
MAXIMUM ABSOLUTE DEVIATION OF THE ANGULAR CONFIDENCE FEATURES DURING 4 DIFFERENT PERTURBATIONS OF THE SYSTEM. VALUES IN BOLD CORRESPOND TO THE MAIN DIRECTION OF THE PERTURBATION

| direction | start | end  | max $ e_\theta $ | max $ e_\phi $ |
|-----------|-------|------|------------------|----------------|
| right     | 21 s  | 27 s | <b>3.12°</b>     | 0.84°          |
| rear      | 33 s  | 37 s | 1.42°            | <b>6.55°</b>   |
| left      | 45 s  | 50 s | <b>4.51°</b>     | 1.60°          |
| front     | 57 s  | 61 s | 0.74°            | <b>4.02°</b>   |



(a)



(b)

Fig. 11. Evolution of the features under a target-specific confidence-driven control. The gray parts correspond to the periods during which a perturbation was applied to the system (motion to the right, motion to the left, shadow on the right and shadow on the left). (a) Confidence feature error  $\theta_c$  in degrees. (b) Target centering error  $e_t$  in millimeters.

TABLE III  
MAXIMUM ABSOLUTE FEATURE ERRORS DURING 4 DIFFERENT PERTURBATIONS OF THE SYSTEM. VALUES IN BOLD CORRESPOND TO THE FEATURE WHICH WAS SPECIFICALLY EXCITED BY THE PERTURBATION

| perturbation | start | end  | max $ e_c $  | max $ e_t $     | max $ e_f $ |
|--------------|-------|------|--------------|-----------------|-------------|
| motion right | 24 s  | 33 s | 1.04°        | <b>11.26 mm</b> | 1.47 N      |
| motion left  | 34 s  | 47 s | 0.77°        | <b>8.34 mm</b>  | 0.95 N      |
| shadow left  | 53 s  | 62 s | <b>2.35°</b> | 0.18 mm         | 0.95 N      |
| shadow right | 70 s  | 79 s | <b>3.89°</b> | 0.18 mm         | 0.86 N      |

probe along, while keeping an optimal orientation. The maximal target tracking error was 11.25 mm, and the maximal confidence feature error was 1.04°.

- **Shadowing** ( $53\text{ s} < t < 62\text{ s}$ ) and ( $70\text{ s} < t < 79\text{ s}$ ): Similar to Section IV-C1, a band of paper was manually inserted between the probe and the phantom [Fig. 7(d)] in order to generate an important shadow. The robot automatically compensated by changing the orientation of the probe, while keeping the region of interest centered. The maximal confidence feature error was 3.89°, and the maximal tracking error was 0.18 mm.

The maximal feature errors of the 3 tasks for each perturbation are reported in Table III. We observe a good decoupling of the different tasks. In particular, the effect of shadow compensation on the target positioning task is negligible (0.18 mm).

#### D. Tracking Quality Evaluation

We propose an investigation of the impact of our confidence-driven control scheme on the tracking quality. To this end, we use the normalized correlation between the intensities  $\mathbf{s}_{\text{track}}$  in the tracked region of interest and those of the initial template,  $\mathbf{s}_{\text{track}}^*$ . Let us recall the expression of the normalized correlation:

$$\rho(\mathbf{s}_{\text{track}}, \mathbf{s}_{\text{track}}^*) = \sum_i \frac{(\mathbf{s}_{\text{track}}(i) - \bar{\mathbf{s}}_{\text{track}})(\mathbf{s}_{\text{track}}^*(i) - \bar{\mathbf{s}}_{\text{track}}^*)}{\sigma(\mathbf{s}_{\text{track}})\sigma(\mathbf{s}_{\text{track}}^*)}, \quad (50)$$

where  $\bar{\mathbf{s}}$  and  $\sigma(\mathbf{s})$  are the mean and standard deviation of  $\mathbf{s}$ , respectively. The normalized correlation  $\rho$  takes values in  $[-1, 1]$ , where 1 denotes a perfect correlation, 0 denotes a total decorrelation, and  $-1$  denotes a perfect anti-correlation. Note that the normalized correlation can account both for tracking errors and for a decrease in the signal quality. Therefore, it is an interesting tool to measure the tracking quality.

To assess the impact of confidence control on the tracking quality, we compared the normalized correlation obtained after convergence in a series of experiments, with and without confidence control. In these experiments, we tested different initial configurations, by defining the target as a  $60 \times 60$  pixels region of interest at 5 mm, 10 mm, 25 mm, and 50 mm from the central scan line. The probe was placed close to the rib cage, so that important shadows could be generated during the experiments. In order to ensure a fair comparison between the results obtained with target positioning only and with combined target positioning / confidence control, the same initial configuration was used in both cases. For each configuration, we run the experiment 10 times.

The results of these experiments are summarized in Fig. 12 as a box-and-whisker diagram representing the minimum, first quartile, median, third quartile, and maximum normalized correlation obtained for each series of experiments. In some cases, the tracking algorithm lost the target due to shadowing and could not recover. This happened 15 times out of 40 for the target positioning without confidence control, and 2 times out of 40 with our confidence control. In such a case, we considered the normalized correlation with the initial template to be zero, so that the occurrence of failure cases is clearly visible in Fig. 12. Note that for target positioning without confidence control with the target initialized at an offset of 25 mm, all trials failed.

These results show an improvement of the final normalized correlation when using confidence control for target positioning, compared to target positioning alone. The number of tracking failures was also significantly lower with confidence control. This suggests that confidence control can indeed improve target tracking.

In order to illustrate the benefit provided by confidence control during target tracking in a non-static environment, we performed another experimental comparison where disturbances (motion, shadowing) were applied to the phantom. In both experiments, the target was initialized manually in a region with high confidence. Results are displayed in Fig. 13, where the gray areas correspond to external disturbances (see the corresponding video sequence). We observe that the normalized correlation with confidence control was high and

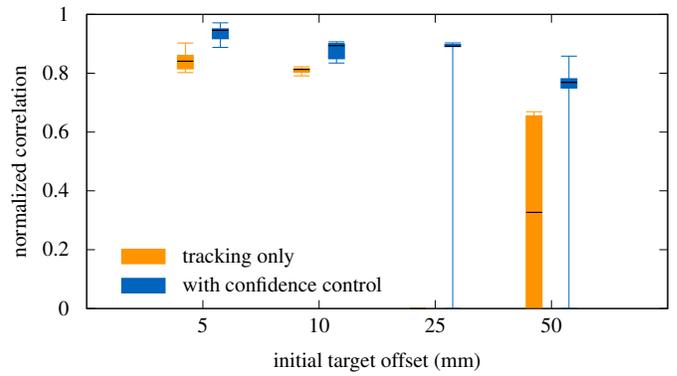
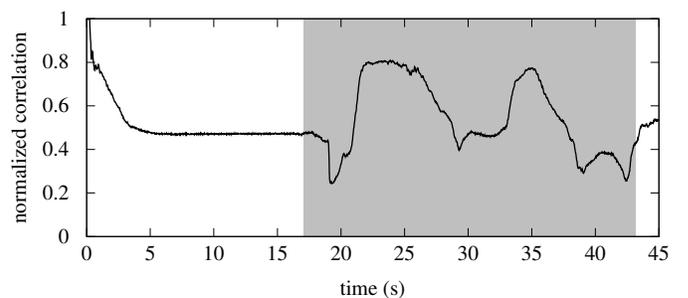
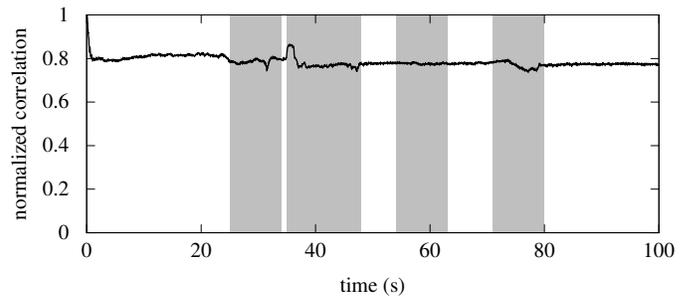


Fig. 12. Normalized correlation between the initial region of interest and the tracked target at convergence, for different initial positions of the target, without and with confidence control. Ten trials were performed for each target configuration. The boxes indicate the first and third quartiles, the whiskers indicate the minima and maxima, and the black lines indicate the median values obtained for each series of trials.



(a) Without confidence control



(b) With confidence control

Fig. 13. Evolution of the normalized correlation between the tracked target and the initial template, (a) without and (b) with confidence control.

stable, at  $0.78 \pm 0.02$ . On the other hand, without confidence control, the normalized correlation took significantly lower values and varied greatly ( $0.53 \pm 0.15$ ) when the phantom was moved. This is because nothing prevented the target region from being shadowed.

To better visualize the evolution of the normalized correlation, we have colored the target region according to the value of  $\rho$  in the video. The color of the rectangle is set, in RGB format, as  $(1 - \max(\rho, 0), \max(\rho, 0), 0)$ , so that the rectangle is red when  $\rho \leq 0$ , green when  $\rho = 1$ , and its color varies continuously in between.

### E. Experiments on Human

We now present the results of experiments performed on a human volunteer. The purpose was to validate the proposed control framework in real conditions. In particular, the system now has to cope with breathing and tissue motion.

For these experiments, we used the force controller provided by the KUKA software in order to guarantee the safety of the procedure. The desired contact force was set to 2 N, which was sufficient to obtain a satisfying image quality. The confidence control gain was set to  $\lambda_c = 0.8 \text{ s}^{-1}$ .

Thereafter, we present the results of two experiments: one with normal breathing, where we use our system for teleoperation, and one with heavy breathing, in order to test the robustness of the system to large motion.

1) *Normal breathing*: The volunteer was asked to breath normally. The breathing period was approximately 3 s, with an average vertical amplitude of 4 mm. The probe was placed vertically above the abdomen [Fig. 14(a)]. Force control was first activated, followed by the global confidence-driven control law. After initial convergence, the feature errors were  $0.58^\circ \pm 0.31^\circ$  for  $\theta_c$  (in-plane orientation), and  $-0.33^\circ \pm 0.02^\circ$  for  $\phi_c$  (out-of-plane orientation). Then, the operator moved the probe towards the right side of the patient. Consequently, the system automatically tilted the probe in order to follow the curvature of the body surface. As a result, the visibility of the tissues was maintained across the field of view of the probe [Fig. 14(b)]. Finally, the operator moved the probe back towards the center of the abdomen, and the robot tilted back in the opposite direction [Fig. 14(c)]. During the teleoperation, the maximum confidence feature errors were  $3.0^\circ$  for  $\theta_c$  and  $2.6^\circ$  for  $\phi_c$ . The evolution of the confidence feature errors during the experiment is represented in Fig. 15. These results show that our confidence-driven controller is able to adapt the orientation of the probe in order to maintain a good acoustic coupling. The maximum confidence feature error ( $3^\circ$ ) was small compared to the field of view of the ultrasound probe and the rotation applied to the probe.

2) *Heavy breathing*: In a second experiment, the volunteer was asked to breath heavily. The purpose of this experiment was to assess the system's capability to adapt the orientation of the ultrasound probe when large motions occur. The probe was placed vertically above the abdomen. Again, force control was activated first, followed by confidence control. The patient had a breathing period of approximately 6 s, with an amplitude of 4 cm along the elevation axis (Fig. 16). Breathing also impacted the optimal scanning direction, so that the system automatically re-oriented the probe, following the breathing pattern. We report in Fig. 17 the confidence feature errors during the experiment. Thanks to confidence control, the confidence feature errors were kept below  $3.2^\circ$  for  $\theta_c$ , and  $1.6^\circ$  for  $\phi_c$ . Therefore, we did not observe any significant increase in the confidence feature error, compared to the experiment performed under normal breathing. This shows that our system is able to compensate for large motions.

### V. CONCLUSION

We have presented a framework for controlling the pose of an ultrasound probe based on an image quality signal. We have

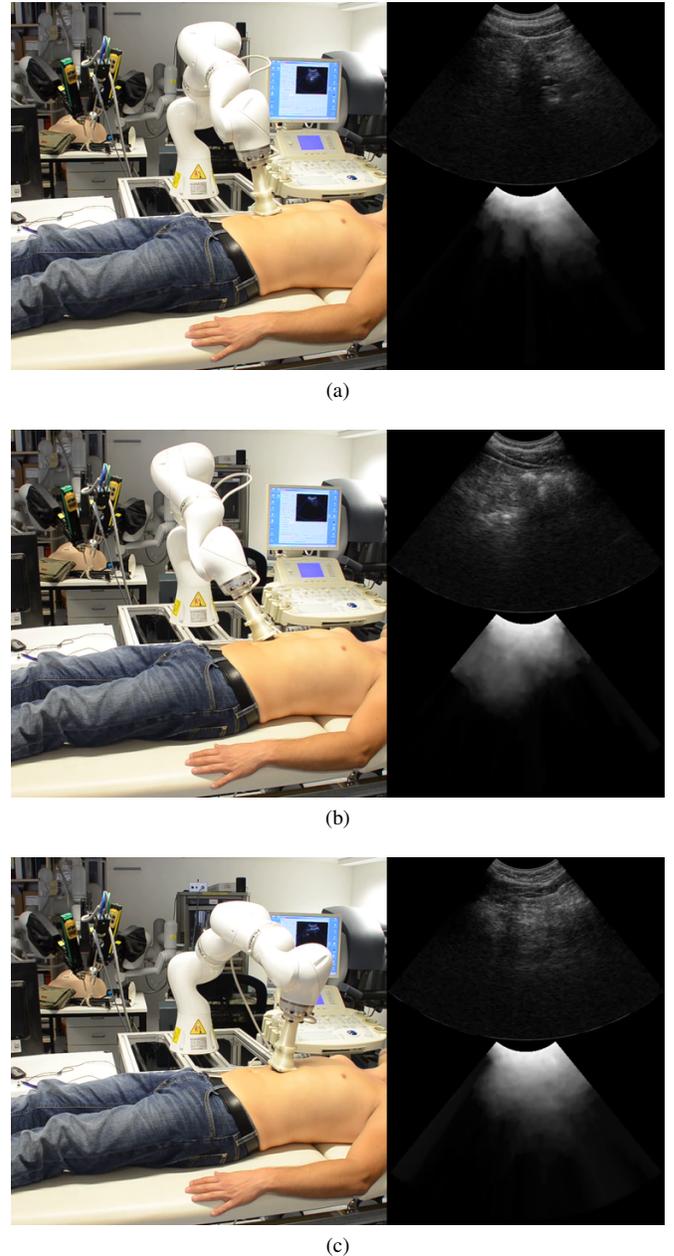


Fig. 14. Confidence-optimized tele-echography on a human volunteer – frontal abdominal scan – normal breathing. (a) Initial position. (b) The operator moves the probe towards the right side of the patient ( $t = 21$  s). (c) The operator moves the probe back to the center of the torso ( $t = 51$  s).

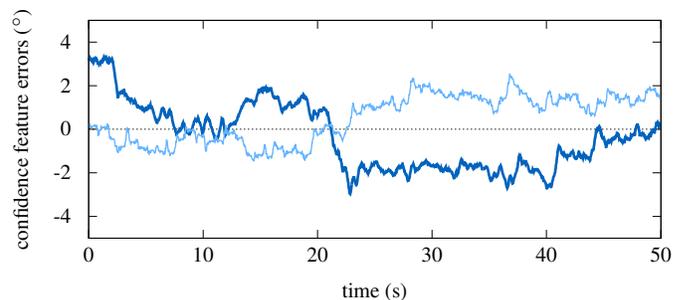


Fig. 15. Evolution of the confidence feature errors  $\theta_c$  (—) and  $\phi_c$  (—), in degrees, during a tele-echography on a human volunteer – frontal abdominal scan – normal breathing.

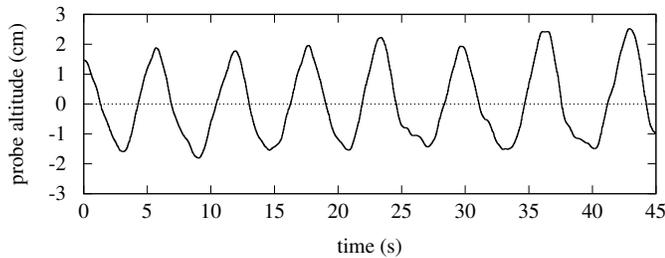


Fig. 16. Position of the probe along the  $z$ -axis (altitude) in the reference frame during heavy breathing.

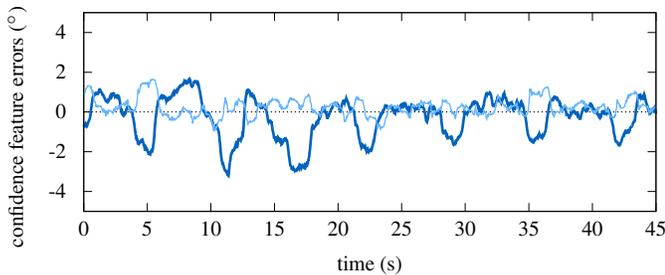


Fig. 17. Evolution of the confidence feature errors  $\theta_c$  (—) and  $\phi_c$  (---), in degrees, during an ultrasound acquisition on a human volunteer – frontal abdominal scan – heavy breathing.

shown how this confidence-driven control can be combined with force control and an image control to design (i) a quality-optimized tele-echography system or (ii) an autonomous tracking system which maintains the visibility of an anatomy of interest. The results obtained on phantom show that:

- Confidence-driven control can provide a robotized ultrasound system with the ability to avoid drop-out artifacts and important shadowing.
- Confidence control can be combined with force and position control in a decoupled manner.
- Using confidence control during the tracking of a region of interest improves the tracking quality.

Moreover, the results of experiments performed on a human volunteer indicate that our control framework is applicable in real conditions, and is robust to perturbation motion.

The confidence-driven control is designed as a subtask which aims at improving other functional tasks. The use of the redundancy framework emphasizes the modular nature of the different tasks, and, while we have presented in this paper some specific implementations, our confidence-driven control could be useful in other applications as well. For instance, the target-specific confidence control could easily be adapted to optimize the imaging quality along a path, in order to perform an autonomous scan of a larger region of interest.

Another direction for future research is the definition of the confidence map. In this work, we used the random walk model introduced in [22]. While this approach provides a good indication of ultrasound signal quality, it only partially models the complex properties of ultrasound propagation. More elaborate models, using for instance an analysis of the radio frequency signal, could be used instead, without

modifying the control part of our framework.

Finally, the results of experiments performed on a volunteer are encouraging for a potential future clinical study.

## APPENDIX

We show how the time derivative of the feature  $\theta_c$  defined in (22) was obtained. Under the assumption  $\dot{C}_\Omega = 0$ , the time derivative of this feature is

$$\dot{\theta}_c = \frac{1}{C_\Omega} \iint_{(r,\theta) \in \Omega} \theta \dot{C}(r, \theta, t) r dr d\theta, \quad (51)$$

which can be written, using (21),

$$\dot{\theta}_c = \frac{1}{C_\Omega} \iint_{(r,\theta) \in \Omega} \theta \omega_z \frac{\partial C}{\partial \theta}(r, \theta, t) r dr d\theta \quad (52)$$

$$= \frac{\omega_z}{C_\Omega} \int_r r \left( \int_\theta \theta \frac{\partial C}{\partial \theta}(r, \theta, t) d\theta \right) dr. \quad (53)$$

Integrating by parts the inner integral, we have:

$$\int_\theta \theta \frac{\partial C}{\partial \theta}(r, \theta, t) r d\theta = [\theta C(r, \theta, t)]_{\theta_{\min}}^{\theta_{\max}} - \int_\theta C(r, \theta, t) d\theta, \quad (54)$$

where the term  $[\theta C(r, \theta, t)]_{\theta_{\min}}^{\theta_{\max}}$ , corresponding to the image borders, can be neglected. Indeed, the confidence at the limits of the field of view is typically lower than inside the field of view, and  $\theta_{\max} - \theta_{\min} < \pi$ , so that  $[\theta C(r, \theta, t)]_{\theta_{\min}}^{\theta_{\max}} \ll \int_{\theta_{\min}}^{\theta_{\max}} C(r, \theta, t) d\theta$ . With this simplification, we obtain

$$\dot{\theta}_c = \frac{\omega_z}{C_\Omega} \int_r r \left( \int_\theta -C(r, \theta, t) d\theta \right) dr = -\omega_z. \quad (55)$$

## REFERENCES

- [1] P. Chatelain, A. Krupa, and N. Navab, "Optimization of ultrasound image quality via visual servoing," in *2015 IEEE Int. Conf. Robotics and Automation (ICRA)*, 2015, pp. 5997–6002.
- [2] P. Chatelain, A. Krupa, and N. Navab, "Confidence-driven control of an ultrasound probe: Target-specific acoustic window optimization," in *2016 IEEE Int. Conf. Robotics and Automation (ICRA)*, 2016, pp. 3441–3446.
- [3] H. E. Vanderpool, E. A. Friis, B. S. Smith, and K. L. Harms, "Prevalence of carpal tunnel syndrome and other work-related musculoskeletal problems in cardiac sonographers," *J. Occupat. Environ. Med.*, vol. 35, no. 6, pp. 604–610, Jun. 1993.
- [4] N. Magnavita, L. Bevilacqua, P. Mirk, A. Fileni, and N. Castellino, "Work-related musculoskeletal complaints in sonologists," *J. Occupat. Environ. Med.*, vol. 41, no. 11, pp. 981–988, Jun. 1999.
- [5] S. Boudet, J. Gariepy, and S. Mansour, "An integrated robotics and medical control device to quantify atheromatous plaques: experiments on the arteries of a patient," in *Proc. 1997 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, vol. 3, 1997, pp. 1533–1538.
- [6] F. Pierrot *et al.*, "Hippocrate: a safe robot arm for medical applications with force feedback," *Med. Image Anal.*, vol. 3, no. 3, pp. 285–300, Sep. 1999.
- [7] S. E. Salcudean, G. Bell, S. Bachmann, W. H. Zhu, P. Abolmaesumi, and P. D. Lawrence, "Robot-assisted diagnostic ultrasound – design and feasibility experiments," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*, C. Taylor and A. Colchester, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 1062–1071.
- [8] D. De Cunha *et al.*, "The MIDSTEP system for ultrasound guided remote telesurgery," in *Proc. 20th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, 1998, vol. 3, 1998, pp. 1266–1269.
- [9] K. Masuda, E. Kimura, N. Tateishi, and K. Ishihara, "Three dimensional motion mechanism of ultrasound probe and its application for tele-echography system," in *Proc. 2001 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, vol. 2, 2001, pp. 1112–1116.
- [10] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Trans. Robot. Autom.*, vol. 18, no. 1, pp. 11–23, Feb. 2002.

- [11] A. Vilchis, J. Troccaz, P. Cinquin, K. Masuda, and F. Pellissier, "A new robot architecture for tele-echography," *IEEE Trans. Robot. Autom.*, vol. 19, no. 5, pp. 922–926, Oct. 2003.
- [12] C. Delgorgue *et al.*, "A tele-operated mobile ultrasound scanner using a light-weight robot," *IEEE Trans. Inf. Technol. Biomed.*, vol. 9, no. 1, pp. 50–58, Mar. 2005.
- [13] R. Mebarki, A. Krupa, and F. Chaumette, "2-D ultrasound probe complete guidance by visual servoing using image moments," *IEEE Trans. Robot.*, vol. 26, no. 2, pp. 296–306, Apr. 2010.
- [14] A. Krupa, G. Fichtinger, and G. D. Hager, "Real-time tissue tracking with B-mode ultrasound using speckle and visual servoing," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, N. Ayache, S. Ourselin, and G. D. Hager, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 1–8.
- [15] R. Nakadate, J. Solis, A. Takanishi, E. Minagawa, M. Sugawara, and K. Niki, "Out-of-plane visual servoing method for tracking the carotid artery with a robot-assisted ultrasound diagnostic system," in *2011 IEEE Int. Conf. Robotics and Automation (ICRA)*, 2011, pp. 5267–5272.
- [16] C. Nadeau and A. Krupa, "Intensity-based ultrasound visual servoing: modeling and validation with 2D and 3D probes," *IEEE Trans. Robot.*, vol. 29, no. 4, pp. 1003–1015, Aug. 2013.
- [17] A. Krupa, D. Folio, C. Novales, P. Vieyres, and T. Li, "Robotized tele-echography: an assisting visibility tool to support expert diagnostic," *IEEE Syst. J.*, vol. 10, no. 3, pp. 974–983, Sep. 2016.
- [18] C. P. Loizou, C. S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaidis, "Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering," *Med. Biol. Eng. Comput.*, vol. 44, no. 5, pp. 414–426, Apr. 2006.
- [19] I. Kuhlemann, "Force and image adaptive strategies for robotised placement of 4D ultrasound probes," M.S. thesis, Inst. Robot. Cogn. Syst., Univ. Lübeck, Lübeck, Germany, 2013.
- [20] N. El-Zehiry, M. Yan, S. Good, T. Fang, S. K. Zhou, and L. Grady, "Learning the manifold of quality ultrasound acquisition," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 122–130.
- [21] P. Hellier, P. Coupé, X. Morandi, and D. L. Collins, "An automatic geometrical and statistical method to detect acoustic shadows in intra-operative ultrasound brain images," *Med. Image Anal.*, vol. 14, no. 2, pp. 195–204, Apr. 2010.
- [22] 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, Aug. 2012.
- [23] J. Kutarnia and P. Pedersen, "A markov random field approach to group-wise registration/mosaicing with application to ultrasound," *Med. Image Anal.*, vol. 24, no. 1, pp. 106–124, Aug. 2015.
- [24] C. Schulte zu Berge, D. Declara, C. Hennersperger, M. Baust, and N. Navab, "Real-time uncertainty visualization for b-mode ultrasound," in *2015 IEEE Scientific Visualization Conf. (SciVis)*, 2015, pp. 33–40.
- [25] B. Siciliano and J.-J. E. Slotine, "A general framework for managing multiple tasks in highly redundant robotic systems," in *5th Int. Conf. Advanced Robotics, 1991. 'Robots in Unstructured Environments', 91 ICAR*, 1991, pp. 1211–1216.
- [26] D. Lee and A. Krupa, "Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4D ultrasound," in *2011 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2011, pp. 2831–2836.
- [27] (2017, Jun.) MELODY, a remote, robotic ultrasound solution. [Online]. Available: <http://www.adechotech.com/products>
- [28] M. Gilbertson and B. Anthony, "Force and position control system for freehand ultrasound," *IEEE Trans. Robot.*, vol. 31, no. 4, pp. 835–849, Aug. 2015.
- [29] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 3, pp. 313–326, Jun. 1992.
- [30] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.
- [31] F. Chaumette and S. Hutchinson, "Visual servo control, part I. Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, Dec. 2006.
- [32] C. Nadeau, H. Ren, A. Krupa, and P. Dupont, "Intensity-based visual servoing for instrument and tissue tracking in 3D ultrasound volumes," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 367–371, Jan. 2015.
- [33] E. Marchand, F. Spindler, and F. Chaumette, "ViSP for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robot. Autom. Mag.*, vol. 12, no. 4, pp. 40–52, Dec. 2005.
- [34] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," presented at the ICRA Workshop Open Source Software in Robotics, 2009.
- [35] M. Bjerkgeng, P. Falco, C. Natale, and K. Y. Pettersen, "Stability analysis of a hierarchical architecture for discrete-time sensor-based control of robotic systems," *IEEE Trans. Robot.*, vol. 30, no. 3, pp. 745–753, Jun. 2014.



**Pierre Chatelain** (S'15–M'16) received the B.S. degree in computer science and the B.S. degree in mathematics from the University of Rennes 1, Rennes, France, in 2010, the M.S. degree in mathematics from Ecole Normale Supérieure de Cachan, Cachan, France, in 2012, and the Ph.D. degree in signal processing and telecommunications from the University of Rennes 1 and the Technical University of Munich, Munich, Germany, in 2016.

From 2012 to 2013, he was a Research Assistant with the Chair of Computer-Aided Medical Procedures, Technical University of Munich, and also with the Institute of Robotics and Mechatronics, German Aerospace Center, Germany. He is currently a Postdoctoral Research Assistant at the University of Oxford, Oxford, UK. His research interests include medical robotics, medical image analysis, and machine learning.



**Alexandre Krupa** (S'00–M'04) received the M.S. and Ph.D. degrees in control systems and signal processing from the National Polytechnic Institute of Lorraine, Nancy, France, in 1999 and 2003, respectively, and the "Habilitation à Diriger des Recherches" in signal processing from the University of Rennes, Rennes, France, in 2012.

He carried out his Ph.D. research work with the Control Vision and Robotics team (eAVR), Laboratoire des Sciences de l'Image, de l'Informatique et de la Télé-détection, Strasbourg, France. From 2002 to 2004, he was an Assistant Associate Professor for undergraduate student lectures in electronics, control, and computer programming with Strasbourg University, Strasbourg, France. Since 2004, he has been a Research Scientist with Inria Rennes – Bretagne Atlantique, Rennes, France, where he is currently a member of the Lagadic group. In 2006, he was a Postdoctoral Associate with the Computer Integrated Surgical Systems and Technology Engineering Research Center, Johns Hopkins University, Baltimore, MD, USA. His current research interests include medical robotics, computer-assisted systems in the medical and surgical fields, and, most specifically, the control of medical robots by visual servoing using ultrasound images.



**Nassir Navab** is a Professor and Director of the Laboratory for Computer Aided Medical Procedures, Technical University of Munich and Johns Hopkins University, where he also has secondary faculty appointments at both affiliated Medical Schools. He received the Ph.D. degree from Inria and University of Paris XI, Paris, France, and enjoyed two years of post-doctoral fellowship at MIT Media Laboratory before joining Siemens Corporate Research (SCR) in 1994. At SCR, he was a distinguished member and received the Siemens Inventor of the Year Award in

2001. He received the SMIT Society Technology award in 2012 and the '10 years lasting impact award' of IEEE ISMAR in 2015. In 2012, he was elected as a Fellow of the MICCAI Society. He is a member of the board of directors of the MICCAI Society and the Steering committee of the IEEE Symposium on Mixed and Augmented Reality (ISMAR). He is the author of hundreds of peer reviewed scientific papers, more than twenty awarded papers, 46 U.S. and more than 50 International Patents. His current research interests include medical augmented reality, computer-aided surgery, medical robotics, surgical data science, and medical image analysis.