2-D Ultrasound Probe Complete Guidance by Visual Servoing Using Image Moments

Rafik Mebarki, Alexandre Krupa, and François Chaumette

Abstract—This paper presents a visual-servoing method that is based on 2-D ultrasound (US) images. The main goal is to guide a robot actuating a 2-D US probe in order to reach a desired cross-section image of an object of interest. The method we propose allows the control of both in-plane and out-of-plane probe motions. Its feedback visual features are combinations of moments extracted from the observed image. The exact analytical form of the interaction matrix that relates the image-moments time variation to the probe velocity is developed, and six independent visual features are proposed to control the six degrees of freedom of the robot. In order to endow the system with the capability of automatically interacting with objects of unknown shape, a model-free visual servoing is developed. For that, we propose an efficient online estimation method to identify the parameters involved in the interaction matrix. Results obtained in both simulations and experiments validate the methods presented in this paper and show their robustness to different errors and perturbations, especially those inherent to the noisy US images.

Index Terms—Medical robotics, model-free servoing, modeling, ultrasound (US) imaging, visual servoing.

I. INTRODUCTION

IMAGE-BASED guidance is a promising approach to performing a wide range of applications. Especially, in medicine, different imaging modalities have been used to assist either surgical or diagnosis interventions. Among these modalities, ultrasound (US) imaging presents relevant advantages of noninvasiveness, safety, and portability. In particular, conventional 2-D US imaging affords noticeably more advantages, i.e., its real-time streaming with high pixel resolution, its widespread in medical centers, and its low cost. In this paper, we present a visual-servoing method to fully and automatically position a 2-D US probe actuated by a medical robot in order to reach a desired cross-section image of an object of interest. The method we propose makes direct use of the US images that are provided by the probe in the servo control scheme. Potential applications are numerous. For instance, in pathology analysis, it can be used to accurately position the 2-D US probe, in order to obtain a 2-D cross-sectional image having a maximum similarity with one previously obtained with the same or other imaging modalities, like magnetic resonance imaging (MRI) and computed tomography scan (CT-SCAN). Also, during a biopsy or a radio-frequency ablation, it could assist the surgeon for needle insertion by positioning the probe on an appropriate soft-tissue cross-section image. However, up until now, few works have dealt with the direct use of US images in visual servoing.

The first work in this area has been presented in [1]. The robotic task was to automatically center the section of the aorta artery in the US image, while an operator was telemanipulating the robot. Visual servoing was thus limited to control only the three degrees of freedom (DOFs) of the in-plane motions of the US probe. An US-image-based visual-servoing method to position a needle for percutaneous cholecystostomy has been proposed in [2]. The needle was mechanically constrained to lie in the observation plane of an eye-to-hand 2-D US probe, and only two in-plane motions were controlled by visual servoing. In fact, the ability to control out-of-plane motions directly from the observed 2-D US images is a real challenge. The main problem is related to the manner by which a 2-D US probe interacts with its environment. Indeed, such a sensor provides full information in its observation plane but none outside of it. Another alternative consists of using 3-D US imaging system. In [3], a motionless 3-D US probe allows guiding of a laparoscopic surgical instrument actuated by a robot arm. The robotic task was to position the instrument tip at a 3-D target location. The proposed approach is a position-based technique that requires an estimation of the instrument pose. Currently, 3-D US imaging systems, however, suffer from low pixel resolution, they are time-consuming, present significant cost, and, furthermore, provide only limited spatial perception. Therefore, in the work presented in this paper, we focus solely on the use of the 2-D US modality.

Recently, few investigations have dealt with the issue of controlling the out-of-plane motions from the observed 2-D US images. A 2-D US-image-based servoing of a robotized laparoscopic surgical instrument that aimed at intracardiac surgery has been presented in [4] and [5]. In those works, a static 2-D US probe observed forceps connected to the tip of the instrument. The intersection of the US probe beam with the forceps results in two image points that were selected as the visual features in the servo scheme, in order to control the 4 DOFs of that instrument. The robotic task was to automatically position the forceps in such a manner that they intersect the US beam at desired image-points positions. However, those servoing methods deal with images of instruments with known geometry,
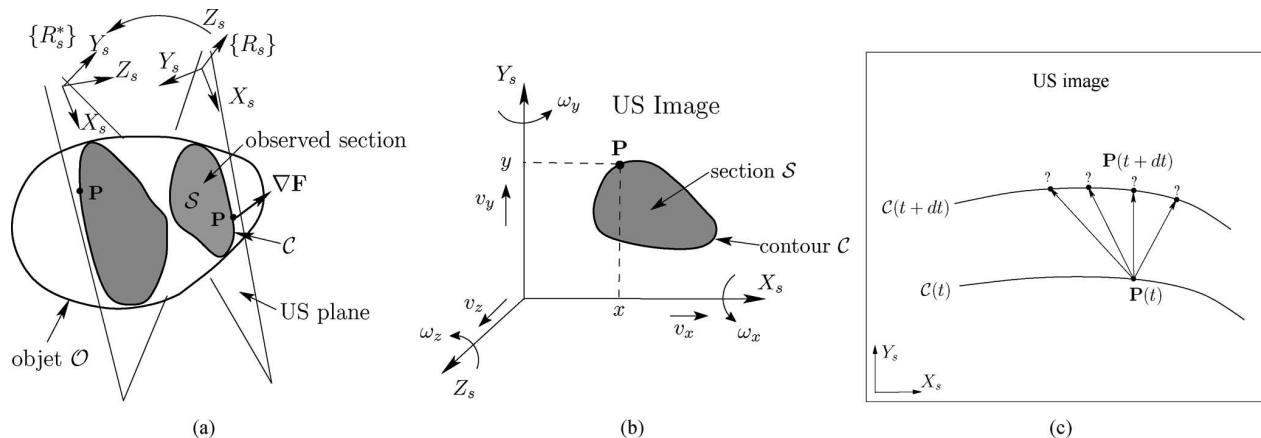
Fig. 1.   Geometrical interpretation of the interaction of the US probe plane with the observed object. (a) Global representation. (b) Observed cross section $\mathcal{S}$ by a 2-D US probe, whose frame three axes (X, Y, and Z) and corresponding velocities are represented. (c) Evolution of an image point **P** due to out-of-plane motion.

namely 3-D straight lines. Recently, visual servoing to control a 2-D US probe from image measurements obtained on soft tissues has been presented in [6]. That approach makes direct use of the speckle correlation contained in the US B-scan images, in order to estimate the soft-tissue displacements that have to be compensated. However, that method developed solely for US images is devoted for compensation and cannot reach a desired image starting from one that is totally different.

In this paper, we present a visual servoing method, based on image moments, that allows the control of both the in-plane and out-of-plane motions of a 2-D US probe. The feedback visual features are combinations of moments extracted from the observed 2-D US image provided by the probe. Using image moments seems to be a judicious direction. Indeed, image moments have the advantage of being general, and those of low order have intuitive and geometric meaning, since they are directly related to the area, the center of gravity, and the orientation of the object of interest in the image. Furthermore, image moments do not necessitate a point-to-point matching in the image but a global segmentation of the object. They are also robust to image noise since they are computed using an integration step (more precisely a summation on discretized image), which filters local errors in the segmentation of the object or in the extraction of its contours. That robustness is of great interest in US modality because of the very noisy images of the latter. Image moments have been widely used in computer vision, especially in pattern-recognition applications [10]–[12]. They have been introduced in visual servoing using cameras. For that, the interaction matrix that relates the image-moments time variation to the camera velocity has been modeled in [13]. However, the modeling in the case of optical systems quite differs from 2-D US one. Indeed, optical systems are generally modeled by a perspective or a spherical projection from the 3-D world to the 2-D image. For 2-D US probes, all the information is available in the cross section but none at all outside. This makes the modeling and control of out-of-plane motions particularly difficult. New techniques have thus to be developed. This was attempted in our previous work [7], where the interaction matrix relating the image moments was approximated. Moreover, only five visual features had been proposed to control the system, while at

least six independent visual features are required to control the 6 DOFs. Furthermore, in that previous work, the 2-D US probe was considered to interact with an ellipsoidal object, whose 3-D parameters were assumed to be coarsely known. In fact, the interaction for out-of-plane motions strongly depends on the 3-D shape of the observed object. This hindered visual servoing using 2-D US images. All those shortcomings are addressed in the present paper. First, we develop the exact analytical form of the interaction matrix that relates the image-moments time variation to the probe velocity. Second, we propose six independent visual features to control the 6 DOFs of the system. Finally, we endow the system with the capability of interacting with objects of unknown shape and location, thanks to a model-free visual-servoing method we develop. To do so, an efficient online estimation method of the parameters that are involved in the interaction matrix is proposed.

The remainder of the paper is organized as follows. In Section II, we derive the exact analytical form of the interaction matrix that relates the image-moments time variation to the probe velocity. In Section III, we check this general result on simple shapes like spheres. The online estimation method of the parameters that are involved in the interaction matrix is presented in Section IV. The visual-servoing control law is briefly derived in Section V. Finally, results obtained in both simulations and real experiments are presented and discussed in Section VI.

II. MODELING

The robotic task consists in automatically positioning an US probe held by a medical robot in order to view a desired cross section of a given soft-tissue object. The choice of appropriate visual features and the determination of the interaction matrix relating their time variation to the probe velocity is a fundamental step to designing the visual-servoing control scheme.

Let $\mathcal{O}$ be the object of interest and $\mathcal{S}$ the intersection of $\mathcal{O}$ with the US probe plane [see Fig. 1(a) and (b)]. The image moment m_{ij} of order $(i + j)$ is defined by

$$m_{ij} = \iint_{\mathcal{S}} f(x, y)\, dx\, dy \tag{1}$$

where $f(x, y) = x^i\, y^j$, and (x, y) represent US image-point coordinates. Note that we do not consider the intensity level in the definition of moments, which means that an image processing algorithm is first applied to segment the object of interest in the image or, equivalently, to extract its contour (the algorithm we have used in practice is briefly described at the beginning of Section VI). The shape of section $\mathcal{S}$ and its configuration in the image are thus the only information used to design the visual features.

The objective here is to determine the analytical form of the time variation $\dot{m}_{ij}$ of moment m_{ij} as function of the probe velocity $\mathbf{v} = (\boldsymbol{v}, \boldsymbol{\omega})$ such that

$$\dot{m}_{ij} = \mathbf{L}_{m_{ij}} \mathbf{v} \tag{2}$$

where $\boldsymbol{v} = (v_x, v_y, v_z)$ and $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ represent the translational and the rotational velocity components, respectively, along and around the X_s, Y_s, and Z_s axes of the cartesian frame $\{R_s\}$, attached to the US probe [see Fig. 1(b)]. The two axes (X_s, Y_s) lie within the image plane, while axis Z_s is orthogonal to the latter. $\mathbf{L}_{m_{ij}}$ is the interaction matrix related to m_{ij} and is denoted by

$$\mathbf{L}_{m_{ij}} = [\, m_{vx}\ \ m_{vy}\ \ m_{vz}\ \ m_{\omega x}\ \ m_{\omega y}\ \ m_{\omega z}\,]. \tag{3}$$

One can intuitively note that the probe in-plane motions (v_x, v_y, ω_z) do not modify the shape of section $\mathcal{S}$, but only its position and orientation in the image [see Fig. 1(a)]. As for the out-of-plane motions $(v_z, \omega_x, \omega_y)$, they also induce variations of the shape, as soon as the object is not a cylinder. We now enter in the complete derivations.

The time variation of moments as a function of the image-point velocity is given by [13]

$$\dot{m}_{ij} = \iint_{\mathcal{S}} \left[\frac{\partial f}{\partial x}\, \dot{x} + \frac{\partial f}{\partial y}\, \dot{y} + f(x,y) \left(\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} \right) \right] dx\, dy \tag{4}$$

that can be written as follows:

$$\dot{m}_{ij} = \iint_{\mathcal{S}} \left[\frac{\partial}{\partial x}\, (\dot{x}\, f(x,y)) + \frac{\partial}{\partial y}\, (\dot{y}\, f(x,y)) \right] dx\, dy \tag{5}$$

where $(\dot{x}, \dot{y})$ is the velocity of an image point (x, y) belonging to section $\mathcal{S}$. In order to determine the relation giving $\dot{m}_{ij}$ as a function of $\mathbf{v}$, the image-point velocity $(\dot{x},\ \dot{y})$ needs to be expressed as function of $\mathbf{v}$, which is the subject of the following part.

A. US Image-Point Velocity Modeling

Let $\mathbf{P}$ be a point of the contour $\mathcal{C}$ of image section $\mathcal{S}$ [see Fig. 1(a) and (b)]. The expression of point $\mathbf{P}$ in the US probe plane is

$${}^s\mathbf{P} = {}^s\mathbf{R}_o{}^o\mathbf{P} + {}^s\mathbf{t}_o \tag{6}$$

where ${}^s\mathbf{P} = (x, y, 0)$ and ${}^o\mathbf{P} = ({}^ox, {}^oy, {}^oz)$ are the coordinates of point $\mathbf{P}$ in the US probe frame $\{R_s\}$ and in the object frame $\{R_o\}$, respectively. The former represents the image coordinates of $\mathbf{P}$. ${}^s\mathbf{R}_o$ is the rotation matrix defining the orientation of the object frame $\{R_o\}$ with respect to probe frame $\{R_s\}$. ${}^s\mathbf{t}_o = (t_x, t_y, t_z)$ is the translation defining the origin of $\{R_o\}$ with respect to $\{R_s\}$.

The time variation of ${}^s\mathbf{P}$ according to the relationship (6) is as follows:

$${}^s\dot{\mathbf{P}} = {}^s\dot{\mathbf{R}}_o{}^o\mathbf{P} + {}^s\mathbf{R}_o{}^o\dot{\mathbf{P}} + {}^s\dot{\mathbf{t}}_o. \tag{7}$$

We use the classical kinematic relationship that states

$$\begin{cases} {}^s\dot{\mathbf{R}}_o = - [\boldsymbol{\omega}]_\times\, {}^s\mathbf{R}_o \\ {}^s\dot{\mathbf{t}}_o = -\boldsymbol{v} + [{}^s\mathbf{t}_o]_\times\, \boldsymbol{\omega} \end{cases} \tag{8}$$

where $[\mathbf{a}]_\times$ denotes the skew-symmetric matrix associated to vector $\mathbf{a}$. Thus, replacing (8) in (7), we obtain

$${}^s\dot{\mathbf{P}} = -\boldsymbol{v} + [{}^s\mathbf{P}]_\times\, \boldsymbol{\omega} + {}^s\mathbf{R}_o{}^o\dot{\mathbf{P}}. \tag{9}$$

Since $\mathbf{P}$ always appears in the image, its velocity expressed in the probe frame is ${}^s\dot{\mathbf{P}} = (\dot{x}, \dot{y}, 0)$. The point $\mathbf{P}$ results from the intersection of the US probe planar beam with the object surface. The only condition that $\mathbf{P}$ must satisfy is that it has to remain on $\mathcal{C}$ during the probe motions. Consequently, in the 3-D space, $\mathbf{P}$ is a moving point that slides on the object surface with a velocity ${}^o\dot{\mathbf{P}} = ({}^o\dot{x}, {}^o\dot{y}, {}^o\dot{z})$ in such a way that this point remains in the US probe beam. Note that when only in-plane motions occur, ${}^o\dot{\mathbf{P}}$ can be set to zero, which has most sense, since the observed section is the same in that case. Therefore, ${}^o\dot{\mathbf{P}}$ is only generated by the out-of-plane motions. Thus, the relationship (9), which represents three constraints, has five unknowns (the two we are looking for in ${}^s\dot{\mathbf{P}}$ and three in ${}^o\dot{\mathbf{P}}$). In order to solve this system, two supplementary constraints have to be established. The first constraint corresponds to the sliding of $\mathbf{P}$ on the object surface. We will show that it can be expressed so that ${}^o\dot{\mathbf{P}}$ belongs to the plane tangent to that surface. In other words, this first constraint represents the fact that the image motion of any contour point $\mathbf{P}(t)$ has to belong to the contour $\mathcal{C}(t + dt)$ [see Fig. 1(c)]. It is clear from Fig. 1(c) that there is an infinity of possibilities for any point $\mathbf{P}(t)$ to be located at a point $\mathbf{P}(t + dt)$ on $\mathcal{C}(t + dt)$. The second constraint, as we will see next, just consists in selecting a direction for the image-point velocity. More precisely, it consists in choosing a direction for ${}^o\dot{\mathbf{P}}$ on the plane tangent to the object surface. Let us note that this way to proceed is valid to determine the variation of the image moments, since this variation is obtained by the integration of the image-point velocity all around contour $\mathcal{C}$. In other words, choosing a different direction would modify the result for the image-point velocity but would not change the result for the variation of the image moments, which is what we want to achieve. This shows the relevance of image moments. We now determine the equations related to these constraints described above.

Let $\mathcal{OS}$ be the set of the 3-D points that lie on the object surface. Any 3-D point $\mathbf{P}$ that belongs to $\mathcal{OS}$ satisfies an equation of the form $F({}^ox, {}^oy, {}^oz) = 0$ that describes the object surface. The fact that any point of $\mathcal{OS}$ always remains on $\mathcal{OS}$ can be expressed by

$$\dot{F}({}^ox, {}^oy, {}^oz) = 0 \quad \forall \mathbf{P} \in \mathcal{OS}. \tag{10}$$

Assuming that the object is rigid, we obtain

$$\dot{F}({}^ox, {}^oy, {}^oz) = \frac{\partial F}{\partial {}^ox}{}^o\dot{x} + \frac{\partial F}{\partial {}^oy}{}^o\dot{y} + \frac{\partial F}{\partial {}^oz}{}^o\dot{z}$$

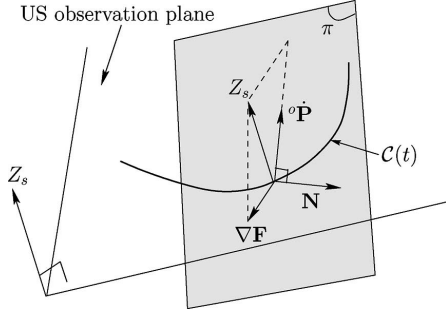$$= {}^o\nabla\mathbf{F}^{\top\,o}\dot{\mathbf{P}} \tag{11}$$

Fig. 2. Point velocity in the 3-D space. $^o\dot{\mathbf{P}}$ and $\mathbf{N}$ lie on π.

where $^o\nabla\mathbf{F}$ is the gradient of F expressed in the object frame $\{R_o\}$. It represents the normal vector to the object surface at point $\mathbf{P}$, as depicted in Figs. 1(a) and 2. The constraint that point $\mathbf{P}$ slides on the object surface is then

$$^o\nabla\mathbf{F}^\top \, ^o\dot{\mathbf{P}} = 0, \tag{12}$$

which ensures that vector $^o\dot{\mathbf{P}}$ lies on the plane tangent to the object surface at $\mathbf{P}$. This plane is denoted π in the following (see Fig. 2). Finally, we now determine the second constraint. As explained earlier, velocity $^o\dot{\mathbf{P}}$ is only due to the out-of-plane motions. This means that the most tangible direction of $^o\dot{\mathbf{P}}$ is orthogonal to in-plane motions, namely, the direction of the probe Z_s-axis. Since Z_s does not necessarily lie within π, wherein $^o\dot{\mathbf{P}}$ is lying, we consider the projection of Z_s on π as the direction of $^o\dot{\mathbf{P}}$. To conclude, $^o\dot{\mathbf{P}}$ has to be orthogonal to the vector $^o\mathbf{N}$ lying on π defined by (see Fig. 2)

$$^o\mathbf{N} = \,^o\mathbf{Z}_s \times \,^o\nabla\mathbf{F} \tag{13}$$

such that $^o\mathbf{Z}_s$ is the expression of $\mathbf{Z}_s$ in the object frame $\{R_o\}$. It is defined by $^o\mathbf{Z}_s = \,^s\mathbf{R}_o^\top \,^s\mathbf{Z}_s$. The second constraint that defines the orientation of $^o\dot{\mathbf{P}}$ can thus be written as follows:

$$^o\mathbf{N}^\top \, ^o\dot{\mathbf{P}} = 0. \tag{14}$$

Going back to the relationship (9), it can be written as follows:

$$^s\mathbf{R}_o^\top \,^s\dot{\mathbf{P}} = -\,^s\mathbf{R}_o^\top \, \boldsymbol{v} + \,^s\mathbf{R}_o^\top \, [^s\mathbf{P}]_\times \, \boldsymbol{\omega} + \,^o\dot{\mathbf{P}}. \tag{15}$$

Multiplying (15) once by $^o\nabla\mathbf{F}^\top$ and then by $^o\mathbf{N}^\top$ and taking into account the constraints (12) and (14) yields

$$\begin{cases} ^o\nabla\mathbf{F}^\top{}^s\mathbf{R}_o^\top{}^s\dot{\mathbf{P}} = -\,^o\nabla\mathbf{F}^\top{}^s\mathbf{R}_o^\top \, \boldsymbol{v} + \,^o\nabla\mathbf{F}^\top{}^s\mathbf{R}_o^\top \, [^s\mathbf{P}]_\times \, \boldsymbol{\omega} \\ ^o\mathbf{N}^\top{}^s\mathbf{R}_o^\top{}^s\dot{\mathbf{P}} = -\,^o\mathbf{N}^\top{}^s\mathbf{R}_o^\top \, \boldsymbol{v} + \,^o\mathbf{N}^\top{}^s\mathbf{R}_o^\top \, [^s\mathbf{P}]_\times \, \boldsymbol{\omega}. \end{cases} \tag{16}$$

Since, we have

$$\begin{cases} ^s\nabla\mathbf{F} = \,^s\mathbf{R}_o{}^o\nabla\mathbf{F} \\ ^s\mathbf{N} = \,^s\mathbf{R}_o{}^o\mathbf{N} = \,^s\mathbf{Z}_s \times \,^s\nabla\mathbf{F} \end{cases} \tag{17}$$

the relationships (16) become

$$\begin{cases} ^s\nabla\mathbf{F}^\top{}^s\dot{\mathbf{P}} = -\,^s\nabla\mathbf{F}^\top \, \boldsymbol{v} + \,^s\nabla\mathbf{F}^\top \, [^s\mathbf{P}]_\times \, \boldsymbol{\omega} \\ ^s\mathbf{N}^\top{}^s\dot{\mathbf{P}} = -\,^s\mathbf{N}^\top \, \boldsymbol{v} + \,^s\mathbf{N}^\top \, [^s\mathbf{P}]_\times \, \boldsymbol{\omega}. \end{cases} \tag{18}$$

The aforementioned system of two scalar equations has two unknowns $\dot{x}$ and $\dot{y}$, which yields to the following unique solution:

$$\begin{cases} \dot{x} = -v_x - K_x \, v_z - y \, K_x \, \omega_x + x \, K_x \, \omega_y + y \, \omega_z \\ \dot{y} = -v_y - K_y \, v_z - y \, K_y \, \omega_x + x \, K_y \, \omega_y - x \, \omega_z \end{cases} \tag{19}$$

with

$$\begin{cases} K_x = \dfrac{f_x \, f_z}{(f_x^2 + f_y^2)} \\[2mm] K_y = \dfrac{f_y \, f_z}{(f_x^2 + f_y^2)} \end{cases} \tag{20}$$

such that $^s\nabla\mathbf{F} = (f_x, f_y, f_z)$. From (19) and (20), we can conclude that the image-point velocity depends only on the image-point position, as for the in-plane motions (v_x, v_y, ω_z), and also depends on the normal vector to the object surface at that point as for the out-of-plane motions $(v_z, \omega_x, \omega_y)$.

B. Image-Moments Time-Variation Modeling

Using the previous modeling of an image-point velocity, the analytical form of the image-moment time variation $\dot{m}_{ij}$ can be developed.

The image points for which the velocity was modeled in the previous section belong to contour $\mathcal{C}$ of $\mathcal{S}$. The image-moment time variation $\dot{m}_{ij}$ given by (5) thus has to be expressed as function of these points and their velocities. This is done by formulating $\dot{m}_{ij}$ on contour $\mathcal{C}$, thanks to the Green's theorem [8], that states

$$\oint_\mathcal{C} F_x \, dx + \oint_\mathcal{C} F_y \, dy = \iint_\mathcal{S} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dx \, dy. \tag{21}$$

Therefore, by taking $F_x = -\dot{y} \, f(x,y)$ and $F_y = \dot{x} \, f(x,y)$ in (5), $\dot{m}_{ij}$ is reformulated as follows:

$$\dot{m}_{ij} = - \oint_\mathcal{C} [f(x,y) \, \dot{y}] \, dx + \oint_\mathcal{C} [f(x,y) \, \dot{x}] \, dy. \tag{22}$$

The image moments can also be expressed on contour $\mathcal{C}$ instead on image section $\mathcal{S}$ by using again the Green's theorem. By setting $F_x = -1/(j+1) x^i \, y^{j+1}$ and $F_y = 0$, we have

$$m_{ij} = \frac{-1}{j+1} \oint_\mathcal{C} x^i \, y^{j+1} \, dx \tag{23}$$

and by setting $F_x = 0$ and $F_y = 1/(i+1) \, x^{i+1} \, y^j$, we have

$$m_{ij} = \frac{1}{i+1} \oint_\mathcal{C} x^{i+1} \, y^j \, dy. \tag{24}$$

Replacing now, in (22), the expressions of the image-points velocity $(\dot{x}, \dot{y})$ with respect to $\mathbf{v}$, which are given by the relationship (19), and then using (23) and (24), we finally obtain the elements of $\mathbf{L}_{m_{ij}}$ defined in (3) as follows:

$$\begin{cases} m_{vx} = -i \, m_{i-1,j} \\ m_{vy} = -j \, m_{i,j-1} \\ m_{vz} = \,^x m_{ij} - \,^y m_{ij} \\ m_{\omega x} = \,^x m_{i,j+1} - \,^y m_{i,j+1} \\ m_{\omega y} = -\,^x m_{i+1,j} + \,^y m_{i+1,j} \\ m_{\omega z} = i \, m_{i-1,j+1} - j \, m_{i+1,j-1} \end{cases} \tag{25}$$

where

$$\begin{cases} ^x m_{ij} = \oint_\mathcal{C} x^i \, y^j \, K_y \, dx \\[3mm] ^y m_{ij} = \oint_\mathcal{C} x^i \, y^j \, K_x \, dy. \end{cases} \tag{26}$$

Similar to the image-point velocity, we can note that the terms corresponding to the in-plane motions (v_x, v_y, ω_z) only depend on the measurements in the image, while the terms corresponding to the out-of-plane motions $(v_z, \omega_x, \omega_y)$ also require the knowledge of the normal vector to the object surface for each point of the observed contour.

III. INTERPRETATION FOR SIMPLE SHAPES

In this section, we analytically verify the validity of the general modeling step on spheres. The case of cylindrical objects is analyzed in [17].

A. Image-Point Velocity

The 3-D points lying on the object surface satisfy the following relationship:

$$F(^ox, {}^oy, {}^oz) = \left(\frac{^ox}{R}\right)^2 + \left(\frac{^oy}{R}\right)^2 + \left(\frac{^oz}{R}\right)^2 - 1 = 0 \quad (27)$$

where R is the radius of the sphere. The gradient vector $^o\nabla\mathbf{F}$ is thus obtained by $^o\nabla\mathbf{F} = 2/R^2 \ (^ox, {}^oy, {}^oz)^\top = 2/R^2 \ ^o\mathbf{P}$.

The point $^o\mathbf{P}$ is expressed as function of its coordinates in the US image using (6)

$$^o\mathbf{P} = {}^s\mathbf{R}_o^\top \ (^s\mathbf{P} - {}^s\mathbf{t}_o). \quad (28)$$

Substituting (28) into the expression of $^o\nabla\mathbf{F}$, which was given earlier, we obtain the normal vector as function of the image-point coordinates

$$^o\nabla\mathbf{F} = \frac{2}{R^2} {}^s\mathbf{R}_o^\top \ (^s\mathbf{P} - {}^s\mathbf{t}_o) \quad (29)$$

that we express in the probe frame

$$^s\nabla\mathbf{F} = \frac{2}{R^2} {}^s\mathbf{R}_o {}^s\mathbf{R}_o^\top \ (^s\mathbf{P} - {}^s\mathbf{t}_o)$$

$$= \frac{2}{R^2} \ (^s\mathbf{P} - {}^s\mathbf{t}_o). \quad (30)$$

Remembering the expression of $^s\mathbf{P}$ and $^s\mathbf{t}_o$ given in Section II-A, we obtain

$$^s\nabla\mathbf{F} = \frac{2}{R^2} \ (x - t_x, y - t_y, -t_z)^\top. \quad (31)$$

The coefficients K_x and K_y, which are involved in the image-point velocity (19), are calculated according to the relation (20) as follows:

$$\begin{cases} K_x = \dfrac{-t_z \ (x - t_x)}{(x - t_x)^2 + (y - t_y)^2} \\ K_y = \dfrac{-t_z \ (y - t_y)}{(x - t_x)^2 + (y - t_y)^2}. \end{cases} \quad (32)$$

We can note that the US image-point velocity does not depend on the rotation matrix $^s\mathbf{R}_o$ between the object frame and the probe frame. This can be explained by the fact that a sphere has no orientation in the 3-D space.

We now write the coefficients K_x and K_y in a more compact form. The constraint (27) is formulated as follows:

$$F(^ox, {}^oy, {}^oz) = \frac{1}{R^2} {}^o\mathbf{P}^\top {}^o\mathbf{P} - 1 = 0. \quad (33)$$

Replacing $^o\mathbf{P}$ given by (28) in (33), we have

$$(^s\mathbf{P} - {}^s\mathbf{t}_o)^\top \ (^s\mathbf{P} - {}^s\mathbf{t}_o) - R^2 = 0. \quad (34)$$

Then, remembering the expressions of $^s\mathbf{P}$ and $^s\mathbf{t}_o$ given in Section II-A yields

$$(x - t_x)^2 + (y - t_y)^2 + t_z^2 - R^2 = 0 \quad (35)$$

that represents the equation of a circle of center (t_x, t_y) and of radius $\rho = \sqrt{R^2 - t_z^2}$. Thus, the area of the image section is $a = \pi\rho^2 = \pi \ (R^2 - t_z^2)$. Replacing (35) in (32), the coefficients K_x and K_y are finally obtained as follows:

$$\begin{cases} K_x = \dfrac{-\pi \, t_z \, (x - t_x)}{a} \\ K_y = \dfrac{-\pi \, t_z \, (y - t_y)}{a}. \end{cases} \quad (36)$$

B. Interaction Matrix

The terms $^x m_{ij}$ and $^y m_{ij}$ involved in (25) are calculated by substituting (36) in (26), where the moments m_{ij}, $m_{i,j-1}$, and $m_{i-1,j}$ are identified according to (23) and (24). We obtain

$$\begin{cases} ^x m_{ij} = \dfrac{\pi t_z \ [(j+1)m_{ij} - j \, t_y \, m_{i,j-1}]}{a} \\ ^y m_{ij} = \dfrac{\pi t_z \ [-(i+1)m_{ij} + i \, t_x \, m_{i-1,j}]}{a} \end{cases} \quad (37)$$

which yields

$$^x m_{ij} - {}^y m_{ij} = \frac{\pi \, t z}{a} \ [k \, m_{ij} - i t_x \, m_{i-1,j} - j t_y \, m_{i,j-1}] \quad (38)$$

with $k = i + j + 2$. We can note that the image-moment time variation, in the case of a sphere-shaped object, depends only on the image moments and the position of the sphere center with respect to $\{R_s\}$.

Since the intersection of a plane with a sphere is a circle [given by (35)], we can define only three independent features from the image. The simplest choice is obviously the area a of the circle $\mathcal{S}$, and the coordinates (x_g, y_g) of its center of gravity. They are defined in terms of the image moments as follows:

$$\begin{cases} a = m_{00} \\ x_g = \dfrac{m_{10}}{m_{00}} \\ y_g = \dfrac{m_{01}}{m_{00}}. \end{cases} \quad (39)$$

The interaction matrices of these features are derived by replacing (38) in (25). We obtain after some developments

$$\mathbf{L}_a = 2\,\pi\,t_z [0 \quad 0 \quad 1 \quad y_g \quad -x_g \quad 0]$$

$$\mathbf{L}_{x_g} = [-1 \quad 0 \quad 0 \quad 0 \quad -t_z \quad y_g]$$

$$\mathbf{L}_{y_g} = [0 \quad -1 \quad 0 \quad t_z \quad 0 \quad -x_g]. \quad (40)$$

As expected, the area a does not vary with the in-plane motions. Also, the coordinates (x_g, y_g) of the center of gravity do not change in response to translational motion in the direction of the probe z-axis. Finally, we can note that when $t_z = 0$, all the coefficients, of the aforementioned interaction matrices, which are involved in the probe out-of-plane motions $(v_z, \omega_x, \omega_y)$, are equal to zero. This can be explained by the fact that, when
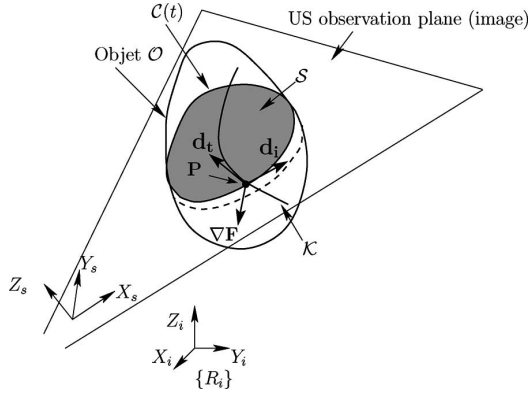
Fig. 3. Surface normals: planar curved lines.

Fig. 4. Object cross-section contour 3-D evolution. The angle γ denotes the polar orientation of the point in counter-clockwise sense. It is defined using as origin the center of gravity of the object and the orientation with respect to X_s-axis of the image plane.

$t_z = 0$, the US probe plane passes through the sphere center. For example, the image section area a is maximal at that configuration and then decreases as soon as the plane moves away from that pose. This means that the derivative of a with respect to the US probe pose ${}^s\mathbf{h}_o$ (${}^s\mathbf{h}_o \in SE_3$) is equal to zero at that configuration (i.e., $\partial a / \partial\, {}^s\mathbf{h}_o = 0$), and then, since $\dot{a} = \partial a / \partial\, {}^s\mathbf{h}_o \cdot {}^s\dot{\mathbf{h}}_o$, we have $\dot{a} = 0$.

IV. NORMAL VECTOR ONLINE ESTIMATION

As shown in Section II-B, the interaction matrix relating the image moments requires the knowledge of the normal vector to the object surface. This normal vector could be derived if a preoperative 3-D model of the object was available. This would also necessitate a difficult calibration step to localize the object frame in the sensor frame. Moreover, that derivation would be possible under the assumption that the object is motionless. To overcome this limitation, we propose in this section an efficient online method that uses the successive 2-D US images to estimate the normal vector.

Let $\mathbf{d_i}$ be the tangent vector to cross-section image contour $\mathcal{C}$ at point $\mathbf{P}$ such that it belongs to that observed image (see Fig. 3). Let $\mathbf{d_t}$ be another tangent vector to the object surface also at $\mathbf{P}$. This vector, in contrast to $\mathbf{d_i}$, does not belong to the image plane. Therefore, from these two vectors, we can express the normal vector $\nabla\mathbf{F}$ in the sensor frame $\{R_s(t)\}$ by

$$^s\nabla\mathbf{F} = {}^s\mathbf{d_i} \times {}^s\mathbf{d_t}. \tag{41}$$

Since ${}^s\mathbf{d_i}$ lies in the US probe observation plane and is moreover expressed in frame $\{R_s\}$, it can directly be measured from the observed image, which is not the case for ${}^s\mathbf{d_t}$. Thus, we only need to estimate ${}^s\mathbf{d_t}$ in order to obtain ${}^s\nabla\mathbf{F}$. We propose to use successive US images to estimate this vector. The principle is to estimate, for each point extracted from the contour $\mathcal{C}$, a 3-D curved line, which is denoted by $\mathcal{K}$, that fits a set of successive points extracted from previous successive images at the same image polar orientation (see Fig. 4). To illustrate the principle, consider the two cross-section image contours $\mathcal{C}(t)$ observed at time t and $\mathcal{C}(t + dt)$ observed at time $t + dt$ after an out-of-plane motion of the US probe (see Fig. 4). The points $\mathbf{P}(t)$ and $\mathbf{P}(t + dt)$ extracted from $\mathcal{C}(t)$ and $\mathcal{C}(t + dt)$, respectively, at the same polar orientation γ lie on the object surface, and
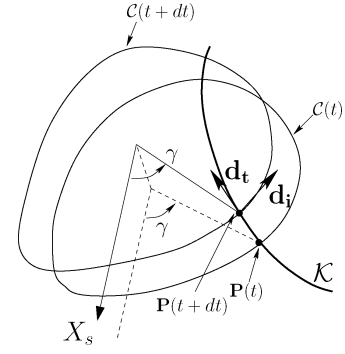
consequently, the curve $\mathcal{K}$ that passes through these points is tangent to the object surface. The direction of $\mathcal{K}$, at $\mathbf{P}(t)$, is nothing but the tangent vector $\mathbf{d_t}$ we want to estimate. Therefore having a set of points at the same polar orientation that have been extracted from successive US images, the objective is to estimate $\mathcal{K}$ that best fits these points. Using a curve allows the consideration of the curvature of the object, which was not the case in [19], where such points have been fitted with a 3-D straight line.

We propose to represent the curve by an analytical model of a second order as follows:

$$\begin{cases} {}^ix = \eta_2\, {}^iz^2 + \eta_1\, {}^iz + \eta_0 \\ {}^iy = \tau_2\, {}^iz^2 + \tau_1\, {}^iz + \tau_0 \end{cases} \tag{42}$$

where $\eta_{i|_{i=0\dots2}}$ and $\tau_{j|_{j=0\dots2}}$ are 3-D parameters to be estimated, and ${}^i\mathbf{P} = ({}^ix, {}^iy, {}^iz)$ are the coordinates of point $\mathbf{P}$ in the initial probe frame $\{R_i\}$. These coordinates are obtained after expressing the image coordinates in frame $\{R_i\}$ by using the robot odometry. One should note that the aforementioned model relates a planar curve. This has the advantage to make the estimation more robust to different perturbations due to the noisy images and the system calibration errors. Estimating the curve $\mathcal{K}$ consists in estimating the vector parameters $\Theta = (\eta_2, \tau_2, \eta_1, \tau_1, \eta_0, \tau_0)$. The system (42) is written as follows:

$$\mathbf{Y} = \mathbf{\Phi}^\top\Theta \tag{43}$$

$$\mathbf{Y} = \begin{bmatrix} {}^ix \\ {}^iy \end{bmatrix}, \quad \text{and} \quad \mathbf{\Phi}^\top = \begin{bmatrix} {}^iz^2 & 0 & {}^iz & 0 & 1 & 0 \\ 0 & {}^iz^2 & 0 & {}^iz & 0 & 1 \end{bmatrix}. \tag{44}$$

We propose to perform the estimation by means of a recursive least squares with stabilization algorithm [18]. It consists of minimizing the following quadratic sum of residual errors:

$$J(\hat{\Theta}_{[t]}) = \sum_{i=t_0}^{t} \beta^{(i-t_0)} (\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\Theta}_{[i]})^\top (\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\Theta}_{[i]})$$

$$\tag{45}$$

where $0 < \beta \le 1$ is a forgetting factor, which is used to assign a weight $\beta^{(i-t_0)}$ to the different estimation errors, in order to take the current measure more into account than the previous ones. The estimate $\hat{\Theta}$ is obtained by nullifying the gradient of $J(\hat{\Theta})$

and is thus given by the following recursive relationship:

$$\hat{\Theta}_{[t]} = \hat{\Theta}_{[t-1]} + \mathbf{F}_{[t]} \, \boldsymbol{\Phi}_{[t]} \left(\mathbf{Y}_{[t]} - \boldsymbol{\Phi}_{[t]}^{\top} \, \hat{\Theta}_{[t-1]} \right) \qquad (46)$$

where $\mathbf{F}_{[t]}$ is a covariance 6×6 matrix. It is defined by the recursive expression

$$\mathbf{F}_{[t]}^{-1} = \beta \, \mathbf{F}_{[t-1]}^{-1} + \boldsymbol{\Phi}_{[t]} \, \boldsymbol{\Phi}_{[t]}^{\top} + (1 - \beta) \, \beta_0 \, \mathbf{I}_6 \qquad (47)$$

where $\mathbf{I}_6$ is the 6×6 identity matrix. The initial parameters are set to $\mathbf{F}_{[t_0]} = f_0 \, \mathbf{I}_6$, with $0 < f_0 \le 1/\beta_0$, and $\hat{\Theta}_{[t_0]} = \Theta_0$. The objective of the stabilization term $(1 - \beta) \, \beta_0 \, \mathbf{I}_6$ is to prevent the matrix $\mathbf{F}_{[t]}^{-1}$ to become ill-conditioned when there is not enough excitation in the input signal $\mathbf{Y}$, which occurs when there is no out-of-plane motion.

The tangent vector ${}^i\mathbf{d_t}$ to the curve $\mathcal{K}$ can then be derived by

$${}^i\mathbf{d_t} = \begin{bmatrix} \dfrac{\partial \, {}^i x}{\partial \, {}^i z} & \dfrac{\partial \, {}^i y}{\partial \, {}^i z} & 1 \end{bmatrix} . \qquad (48)$$

Applying (48) on (42), we get the formula of ${}^i\mathbf{d_t}$ as follows:

$${}^i\mathbf{d_t} = \begin{bmatrix} 2 \, \eta_2 \, {}^i z + \eta_1 \\ 2 \, \tau_2 \, {}^i z + \tau_1 \\ 1 \end{bmatrix} . \qquad (49)$$

Finally, the normal vector ${}^s\nabla\mathbf{F}$ is obtained by taking back the relationship (41) after expressing ${}^i\mathbf{d_t}$ in the probe frame by ${}^s\mathbf{d_t} = {}^b\mathbf{R}_s^{\top} \, {}^b\mathbf{R}_i \, {}^i\mathbf{d_t}$, once the parameters are estimated. ${}^b\mathbf{R}_s$ and ${}^b\mathbf{R}_i$ are the rotation matrices defining the orientation of frames $\{R_s\}$ and $\{R_i\}$, respectively, with respect to the robot base frame $\{R_b\}$. They are obtained using the robot odometry.

V. VISUAL SERVOING

We now present the visual-servoing scheme. The visual features are selected as combinations of the image moments such that $\mathbf{s} = \mathbf{s}\,(m_{ij})$. To control the 6 DOFs of the probe, we have to select at least six independent visual features. Our selection is as follows:

$$\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, \phi_1, \phi_2) \qquad (50)$$

x_g, y_g, and a have already been introduced in Section III-B and are given by the relationship (39), α is the angle of the main orientation of the object in the image, and ϕ_1 and ϕ_2 are moments invariant to the image scale, translation, and rotation [14]. They are given by

$$\begin{cases} \alpha = \dfrac{1}{2} \arctan \left(\dfrac{2 \, \mu_{11}}{\mu_{20} + \mu_{02}} \right) \\[2mm] \phi_1 = \dfrac{I_1}{I_2} \\[2mm] \phi_2 = \dfrac{I_3}{I_4} \end{cases} \qquad (51)$$

where μ_{ij} is the central image moment of order $i + j$ defined by

$$\mu_{ij} = \iint_{\mathcal{S}} (x - x_g)^i \, (y - y_g)^j \, dx \, dy, \qquad (52)$$

and where $I_1 = \mu_{11}^2 - \mu_{20} \, \mu_{02}$, $I_2 = 4 \, \mu_{11}^2 + (\mu_{20} - \mu_{02})^2$, $I_3 = (\mu_{30} - 3 \, \mu_{12})^2 + (3 \, \mu_{21} - \mu_{03})^2$, and $I_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$. We select $\sqrt{a}$ instead of a in the visual-features

vector since x_g, y_g, and $\sqrt{a}$ have the same unit. The last three elements of $\mathbf{s}$ are selected to control the probe out-of-plane motions. Indeed, all these features are invariant to in-plane motions, which allows the system to be partly decoupled. (ϕ_1, ϕ_2) have been chosen, since they are invariant to scale and are thus not sensitive to the size variation of the cross section. Consequently, (ϕ_1, ϕ_2) are expected to be decoupled with the area a. Furthermore, ϕ_1 and ϕ_2 are also expected to be independent from each other, since the former is calculated from moments of order 2 and the latter from moments of order 3.

The time variation of the visual-features vector as function of the probe velocity is written using (25) and (26) as follows:

$$\dot{\mathbf{s}} = \mathbf{L_s} \, \mathbf{v} \qquad (53)$$

where

$$\mathbf{L_s} = \begin{bmatrix} -1 & 0 & x_{g_{vz}} & x_{g_{\omega x}} & x_{g_{\omega y}} & y_g \\ 0 & -1 & y_{g_{vz}} & y_{g_{\omega x}} & y_{g_{\omega y}} & -x_g \\ 0 & 0 & \alpha_{vz} & \alpha_{\omega x} & \alpha_{\omega y} & -1 \\ 0 & 0 & \dfrac{a_{vz}}{2\sqrt{a}} & \dfrac{a_{\omega x}}{2\sqrt{a}} & \dfrac{a_{\omega y}}{2\sqrt{a}} & 0 \\ 0 & 0 & \phi_{1vz} & \phi_{1\omega x} & \phi_{1\omega y} & 0 \\ 0 & 0 & \phi_{2vz} & \phi_{2\omega x} & \phi_{2\omega y} & 0 \end{bmatrix} . \qquad (54)$$

The expressions of some elements involved in (54) are not detailed here because of their tedious form. On one hand, we can check that $\sqrt{a}$, ϕ_1, and ϕ_2 are invariant to the in-plane motions (v_x, v_y, ω_z). On the other hand, the remaining features x_g, y_g, and α present a good decoupling property for the in-plane motions owing to the triangular part they form.

Finally, a very classical control law is used [15]

$$\mathbf{v_c} = -\lambda \, \hat{\mathbf{L}}_\mathbf{s}^{-1} \, (\mathbf{s} - \mathbf{s}^*) \qquad (55)$$

where $\mathbf{v_c}$ is the US probe velocity sent to the low-level robot controller, λ is a positive control gain, $\mathbf{s}^*$ is the desired visual-features vector, and $\hat{\mathbf{L}}_\mathbf{s}^{-1}$ is the inverse of the estimated interaction matrix $\mathbf{L_s}$.

The control scheme (55) is known to be locally asymptotically stable when a correct estimation $\hat{\mathbf{L}}_\mathbf{s}$ of $\mathbf{L_s}$ is used (i.e., as soon as $\mathbf{L_s} \, \hat{\mathbf{L}}_\mathbf{s}^{-1} > 0$) [15].

Some of the experiments later presented have been conducted with less than six features in the visual vector $\mathbf{s}$. In these cases, instead of using $\hat{\mathbf{L}}_\mathbf{s}^{-1}$ in (55), we use the pseudoinverse $\hat{\mathbf{L}}_\mathbf{s}^{+}$ given by

$$\hat{\mathbf{L}}_\mathbf{s}^{+} = \hat{\mathbf{L}}_\mathbf{s}^{\top} \left(\hat{\mathbf{L}}_\mathbf{s} \, \hat{\mathbf{L}}_\mathbf{s}^{\top} \right)^{-1} . \qquad (56)$$

VI. RESULTS

The methods presented earlier have been implemented in C++ language under Linux operating system. The computations are performed on a PC computer equipped with a 3-GHz processor. The image processing used in the experiments presented in Section VI-C–E is described in [20]. In few words, it consists in extracting and tracking in real time the contour of the object of interest in the US image using a snake approach, and a polar description to model the contour.
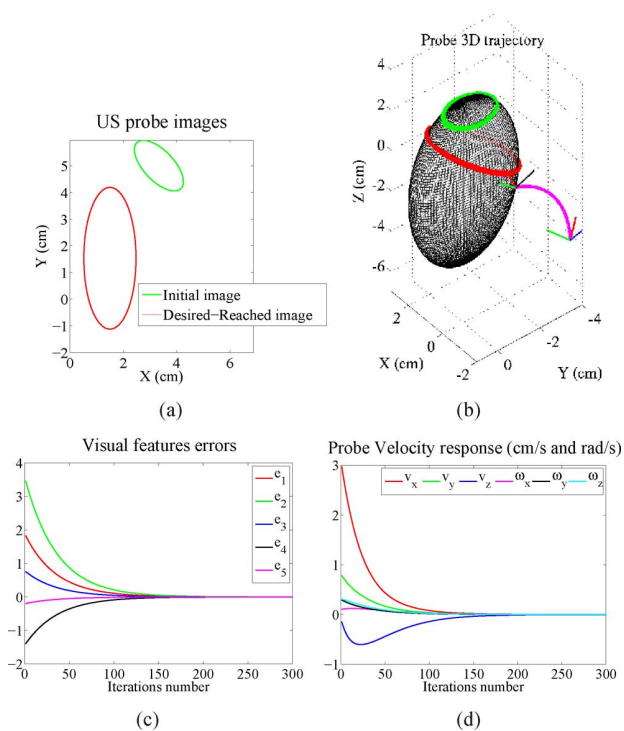
Fig. 5. Simulation results from an ellipsoidal object. (a) US images, where the initial cross-section image is contoured with green, and the desired-reached ones are contoured with red. (b) Interaction between the virtual 2-D US probe and the object, where the initial cross section is contoured with green, and the reached one is contoured with red. The probe 3-D trajectory is in magenta, where the initial and the reached probe frames are each one represented by the three axes (X, Y, and Z), which are, respectively, depicted with three (red, green, and blue) lines. (c) Visual-features errors time response (cm, cm, rad, cm, unit). (d) Probe velocity sent to the virtual robot controller.

A. Simulation Results From an Ellipsoidal Object

In a first part, we designed a simulator in C++ language, where the interaction of a 2-D US probe with a perfect ellipsoid-shaped object is fully mathematically modeled. For this simulation, we assume the exact knowledge of the object 3-D parameters and its location. This allows us to first validate the theoretical developments presented in Section II. Indeed, in this case, the interaction matrix $\mathbf{L}_{m_{ij}}$, which is derived in Section II-B, is expected to be exact, since all its parameters can be computed from the mathematical model. The image points of the object contour are also computed directly from the mathematical model. The half-length values of the ellipsoidal object main axes are $(a_1, a_2, a_3) = (1, 2.5, 4)$ cm. Since the intersection of an ellipsoid with a plane is an ellipse, only five independent visual features can be defined. In this case, the visual-features vector we choose is $\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, \phi_1)$. The control gain λ is set to 0.7. The corresponding simulation results are shown in Fig. 5. We can see that the five visual features errors converge exponentially at the same time to zero and that the reached cross-section image corresponds to the desired one [see Fig. 5(a) and (c)]. Furthermore, a correct and smooth motion has been performed by the probe, as can be seen in Fig. 5(b) and (d). These results validate the theoretical developments presented in Section II.
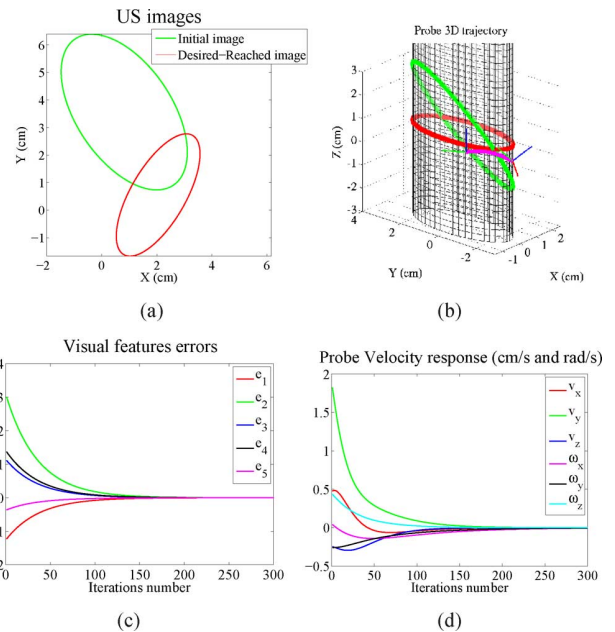

Fig. 6. Simulation results from a cylindrical object. (a) US images, where the initial cross-section image is contoured with green, and the desired-reached ones are contoured with red. (b) Interaction between the virtual 2-D US probe and the object, where the initial cross section is contoured with green, and the reached one is contoured with red. (c) Visual-features errors time response (cm, cm, rad, cm, unit). (d) Probe velocity sent to the virtual robot controller.

B. Simulation Results From a Cylindrical Object

We also tested the method in the case a 2-D US probe interacts with a cylinder-shaped object. Similar to the previous simulation, the object 3-D parameters and their locations are assumed to be exactly known, due to a mathematical model we developed. This simulation allows us to validate the generality of the developed method in the sense that it can deal with different shapes. The half-length values of the object main axes are $(a_1, a_2) = (1, 2.5)$ cm. Since the intersection of a plane with this cylinder $(a_1 \neq a_2)$ is an ellipse, we can again define only five independent visual features. We select similarly $\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, \phi_1)$. As before, the control gain λ is set to 0.7. The corresponding simulation results are shown in Fig. 6. As expected, we can see that the five visual-features errors converge exponentially at the same time to zero and the reached cross-section image corresponds to the desired one [see Fig. 6(c) and (a)]. Also, a correct and smooth motion has been performed by the probe, as can be seen in Fig. 6(d) and (b).

After validating the theoretical developments, we now test the capability of the system to deal with objects of unknown shape thanks to the online estimation method presented in Section IV. This is done in the following experiments that consist of applying the model-free servoing method on objects without any prior information either about their shape, their 3-D parameters, or their location in the 3-D space.

C. Simulation Results From a Virtual Binary Object

We now present the case a 2-D US probe interacting with an object, whose shape does not present symmetries. This will
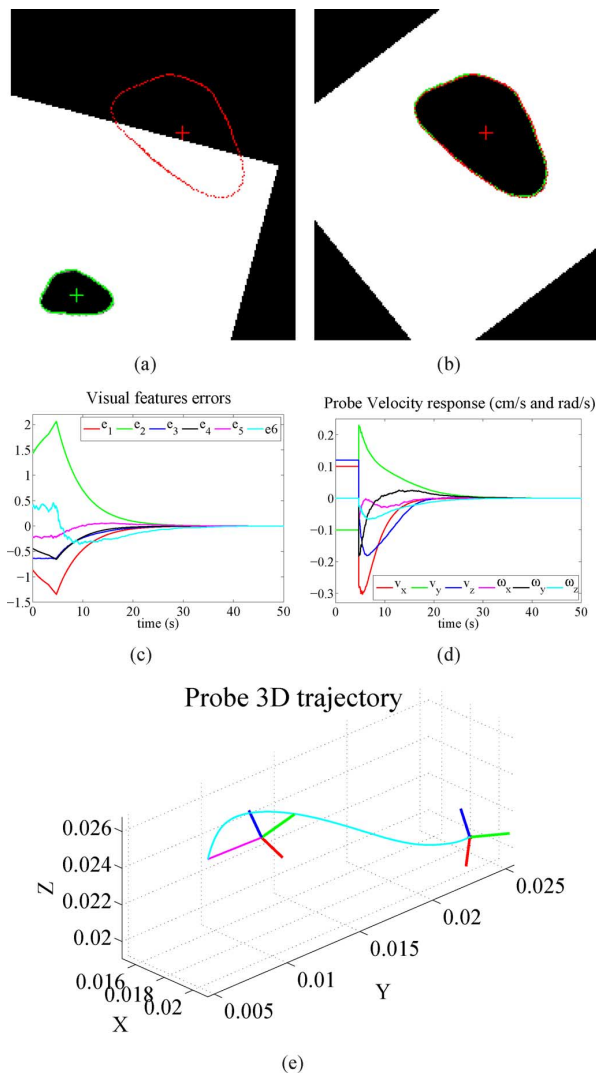
Fig. 7. Simulation results from a virtual binary object. (a) Initial cross section (contoured with green), just before launching visual servoing, to reach the target one (contoured with red). (b) Desired cross section is reached after visual servoing. (c) Visual-features errors time response (cm, cm, rad, cm, unit, 10×unit). (d) Probe velocity sent to the robot controller. (e) 3-D trajectory (m, m, m) performed by the probe (the trajectory corresponding to the moving away with constant velocity is ploted with magenta, while that obtained by visual servoing is ploted with cyan) that retrieves the pose, where the target image has been captured.

allow us to not only reach a desired image but also to correctly position the probe with respect to the object by controlling the 6 DOFs of the system. The visual-features vector **s** is now given by (50). We consider for this simulation a virtual object represented by a binary volume constructed from 100 binary cross-section images. A simulator, which has been employed in [6], is used to perform the interaction of a virtual 2-D US probe with the volume. It allows to position and move the 2-D virtual probe and provides the corresponding 2-D US image. This simulator has been built from the Visualization Toolkit (VTK) software system [21]. VTK is used to render the 3-D view of the US volume and generate the 2-D cross-section image [see Fig. 7(a) and (b)] observed by the virtual 2-D probe, by means of a cubic interpolation. No information about the object 3-D model or its location in the 3-D space is provided nor is it used in this

Fig. 8. Experimental setup consisting of a 6-DOF medical robot arm (right), a 2-D US probe transducer, and a water-filled tank.

trial. The simulation consists first in learning a desired cross-section image target, positioning the probe at a different pose, moving away from that pose by applying a constant probe velocity during 100 iterations, and then, applying visual servoing in order to retrieve the desired image. While moving away with constant probe velocity, a nonrecursive least-squares algorithm of 60 images window is applied in order to obtain an initial estimate Θ_0. The control gain λ is set to 0.2, and the parameters involved in the recursive algorithm to estimate the normal vector are $\beta = 0.8$, $f_0 = 1e6$, and $\beta_0 = 1/(20 f_0)$. The corresponding simulation results[1] are shown in Fig. 7. We can see that the visual-features errors converge roughly exponentially to zero and that the reached cross-section image corresponds to the desired one [see Fig. 7(c) and (b)]; this is despite the large difference from the initial image [see Fig. 7(a)]. The pose reached by the probe corresponds to the one where the desired image was captured [see Fig. 7(e)]. Moreover, correct and smooth behavior has been performed by the probe, as can be seen in Fig. 7(d) and (e). This result validates the model-free method developed in this paper, as well as the relevance of the selection of the six visual features to control the 6 DOFs of the system.

D. Experimental Results From a Spherical Object

We first present experimental results, where we consider the simple case of a 2-D US probe interacting with a spherical object. We use a 6-DOFs medical robot arm similar to the Hippocrate system [22] that actuates a 2-5 MHz 2-D broadband US transducer (see Fig. 8). The PC grabs the US images with a cadence of 25 frames/s to compute the control velocity that is sent to the robot at the same frequency rate. Since the system interacts with a sphere, we can select only three independent visual features to control the system. We choose $\mathbf{s} = (x_g, y_g, \sqrt{a})$ (see Section III). The experiment consists in first learning a desired cross-section image target, moving away from it by applying a constant probe velocity and then applying visual servoing in order to reach that desired image. The control gain λ is set to 0.1. The parameters involved in the recursive algorithm to estimate the normal vector, which are presented in Section IV, are $\beta = 0.8$, $f_0 = 1e6$, and $\beta_0 = 1/(20 f_0)$. During the moving away with constant probe velocity, a nonrecursive least-squares

---
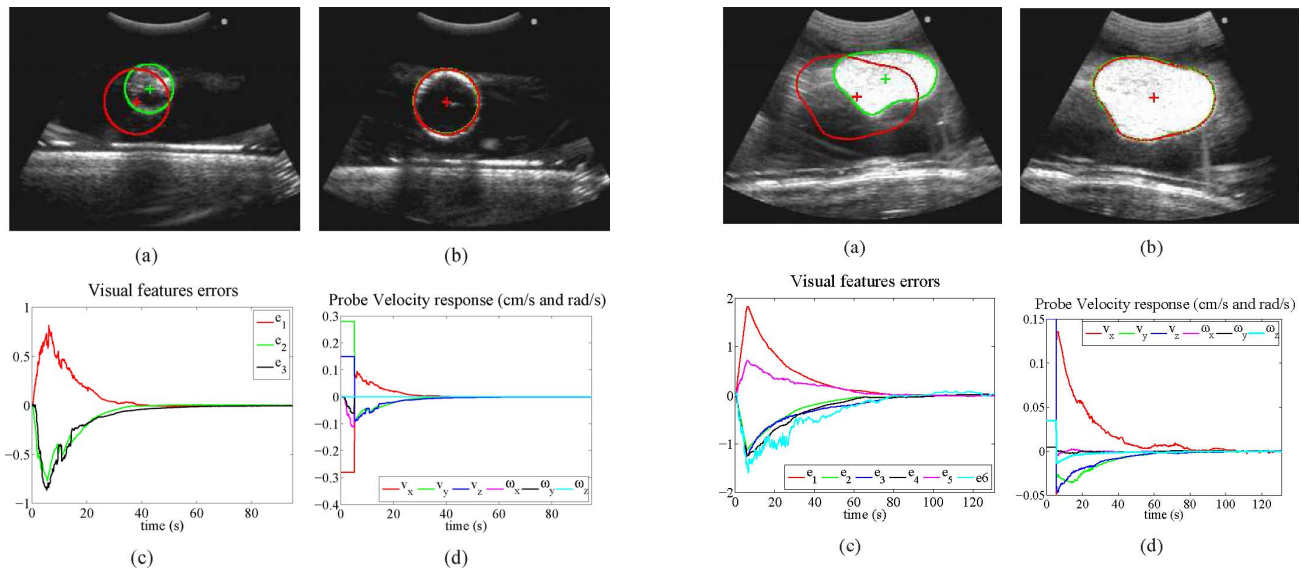
[1] A video accompanies the paper.

Fig. 9. Experimental results from a sphere. (a) Initial cross-section image (contoured with green), just before launching visual servoing, to reach the target one (contoured with red). (b) Desired cross-section image is reached after visual servoing. (c) Visual-features errors time response (cm, cm, cm). (d) Probe velocity sent to the robot controller.

algorithm of 60 images window is applied in order to obtain an initial estimate Θ_0. The corresponding experimental results[1] are shown in Fig. 9. We can see that the three visual-features errors converge exponentially to zero and the reached cross-section image corresponds to the desired one [see Fig. 9(c) and (b)]. Moreover, the probe has performed a correct behavior, as can be seen in Fig. 9(d). This result gives a first experimental validation of the model-free servoing method proposed in this paper.

E. Experimental Results From a Soft-Tissue Object

Finally, we test the method during experiments on an asymmetric object in such a way we can use six visual features. The vector **s** is now given by (50). We use the same medical robot and 2-D broadband US transducer described earlier (see Fig. 8). The object considered is made by gelatin immersed in a water-filled tank in such a way to mimic a real soft tissue. The experiment consists in first learning a desired cross-section image, moving away from it by applying constant probe velocity during 5.5 s and then applying the visual servoing developed in this paper in order to retrieve the desired image. The control gain λ is set to 0.05, and the parameters involved in the recursive algorithm to estimate the normal vector are as usual $\beta = 0.8$, $f_0 = $ 1e6, and $\beta_0 = 1/(20f_0)$. During the moving away with constant probe velocity, a nonrecursive least-squares algorithm of 60 images window is performed in order to obtain an initial estimate Θ_0. The corresponding experimental results[1] are shown in Fig. 10. We can see that the six visual-features errors converge exponentially to zero, and the reached cross section corresponds to the desired one [see Fig. 10(c) and (b)]. As expected, the US probe automatically comes back very near to the pose, where the desired cross-section image was captured [see Fig. 10(e)]. The pose errors are $(0.4, 0.6, -0.2)$ mm and $(0.05, -0.7, -0.8)°$ for the position and the θu rotation, respectively. Moreover, de-
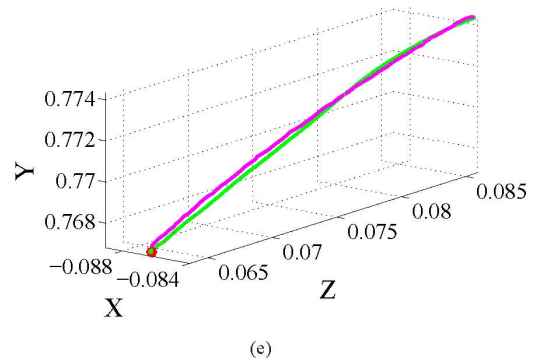

Fig. 10. Experimental results from gelatin object. (a) Initial cross section (contoured with green), just before launching visual servoing, to reach the target one (contoured with red). (b) Desired cross section is reached after visual servoing. (c) Visual-features errors time response (cm, cm, deg/10, cm, unit, 10×unit). (d) Probe velocity sent to the robot controller. (e) 3-D trajectory (m, m, m) performed by the US probe (the trajectory corresponding to the motion moving away is in magenta, and that obtained during visual servoing is in green) that retrieves the pose (red stared point), where the desired cross-section image was captured.

spite the different perturbations mainly generated by the very noisy images and system calibration errors, the robot performed a smooth motion, as can be seen in Fig. 10(d) and (e).

Experimental results obtained using a lamb kidney immersed in a water-filled tank are described in [19] and [17].

VII. CONCLUSION

The contribution of this paper is a new visual-servoing method from 2-D US images by using image moments. The exact analytical form of the interaction matrix that relates the image-moments time variation to the probe velocity has been developed. Six independent visual features have been proposed to control the 6 DOFs of the system, thus allowing an accurate positioning of the 2-D US probe with respect to an observed object. For this, we made the assumption that the observed object is not symmetric. If it is not the case, the probe may not be correctly positioned with respect to the observed object. This is due to the fact that an infinity of probe positions may correspond to a same desired US image. The servoing system has been

endowed with the capability of automatically interacting with objects of unknown shapes without any prior knowledge of their 3-D parameters nor their 3-D location, by developing a model-free visual-servoing method. For that, we proposed an efficient online estimation technique of the 3-D parameters involved in the servo scheme. The results obtained in both simulations and experiments have shown the validity of the developed method and its robustness with respect to the noisy images. The method we proposed is general in the sense that it can be applied to different imaging modalities that, like US, provide full information in their observation plane, as for instance, MRI and CT-SCAN. The presented model-free visual servoing method is however currently devoted for motionless objects. Considering moving objects can be technically addressed by using a high sampling-rate frequency in the online estimation algorithm and the visual servoing in such a way that they become insensitive to these motions. Nevertheless, if the object moves with a high velocity, the online estimation algorithm may fail. That is the reason why it will be necessary to theoretically improve the model-free servoing method in the future to take into account such motion.

REFERENCES

[1] 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.

[2] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata, "An ulrasound-driven needle insertion robot for percutaneous cholecystostomy," *Phys. Med. Biol.*, vol. 49, no. 3, pp. 441–445, 2004.

[3] P. M. Novotny, J. A. Stoll, P. E. Dupont, and R. D. Howe, "Real-time visual servoing of a robot using three-dimensional ultrasound," in *Proc. IEEE Int. Conf. Robot. Autom.*, Roma, Italy, May 2007, pp. 2655–2660.

[4] M. A. Vitrani, H. Mitterhofer, N. Bonnet, and G. Morel, "Robust ultrasound-based visual servoing for beating heart intracardiac surgery," in *Proc. IEEE Int. Conf. Robot. Autom.*, Roma, Italy, Apr. 2007, pp. 3021–3027.

[5] M. Sauvée, P. Poignet, and E. Dombre, "Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control," *Int. J. Robot. Res.*, vol. 27, no. 1, pp. 25–40, Jan. 2008.

[6] A. Krupa, G. Fichtinger, and G. D. Hager, "Real-time motion stabilization with B-mode ultrasound using image speckle information and visual servoing," *Int. J. Robot. Res.*, vol. 28, no. 10, pp. 1334–1354, 2009.

[7] R. Mebarki, A. Krupa, and F. Chaumette, "Image moments-based ultrasound visual servoing," in *Proc. IEEE Int. Conf. Robot. Autom.*, Pasadena, CA, May 2008, pp. 113–119.

[8] J. Stewart, *Calculus*, 2nd ed. Pacific Grove, CA: Brooks/Cole, 1991.

[9] A. G. Mamistvalov, "N-dimensional invariants and conceptual mathematical theory of recognition n-dimentional solids," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 819–831, Aug. 1998.

[10] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, pp. 179–187, 1962.

[11] R. Mukundan and K. R. Ramakrishman, *Moment Functions in Image Analysis: Theory and Application*. Singapore: World Scientific, 1998.

[12] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *Comput. Vis., Graph., Image Process. Conf.*, vol. 54, pp. 438–460, Sep. 1992.

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

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

[15] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Autom.*, vol. 8, no. 6, pp. 313–326, Jun. 1992.

[16] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[17] R. Mebarki, "Automatic guidance of robotized 2D ultrasound probes with visual servoing based on image moments," Ph.D. dissertation, IRISA, Rennes, Mar. 2010.

[18] G. Kreisselmeier, "Stabilized least-squares type adaptive identifiers," *IEEE Trans. Automat. Control*, vol. 35, no. 3, pp. 306–309, Mar. 1990.

[19] R. Mebarki, A. Krupa, and F. Chaumette, "Modeling and 3D local estimation for in-plane and out-of-plane motion guidance by 2D ultrasound visual servoing," in *Proc. IEEE Int. Conf. Robot. Autom.*, Kobe, Japan, May 2009, pp. 1206–1212.

[20] C. Collewet, "Polar snakes: A fast and robust parametric active contour model," presented at IEEE Int. Conf. Image Process., Cairo, Egypt, Nov. 2009.

[21] W. Schroeder, K. Martin, and B. Lorensen *The Visualization Toolkit: An Object Oriented Approach to 3D Graphics*, 4th ed., Kitware, Dec. 1, 2006, ISBN: 193093419X.

[22] F. Pierrot, E. Dombre, E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Gariepy, and J. Megnien, "Hippocrate: A safe robot arm for medical applications with force feedback," *Med. Image Anal.*, vol. 3, no. 3, pp. 285–300, 1999.

Rafik Mebarki was born in Algeria in April 1983. He graduated with the Engineer degree in automatic control from the National Polytechnic School, Algiers, Algeria, in 2005. He received the M.S. degree in automatic systems from Paul Sabatier University, Toulouse, France, in 2006. He is currently working toward the Ph.D. degree with the Lagadic team at IRISA/INRIA, Rennes-Bretagne Atlantique, Rennes, France.

His current research interests include robotics and visual servoing, especially servoing from ultrasound images.

Mr. Mebarki was a Finalist for the Best Vision Paper Award at the 2008 IEEE International Conference on Robotics and Automation and the 2008 Medical Image Computing and Computer-Assisted Intervention Young Scientist Awards in the Robotics and Interventions category.

Alexandre Krupa 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. His Ph.D. research work was carried out with the eAVR team (Control Vision and Robotics) with the Laboratoire des Sciences de l'Image de l'Informatique et de la Télédétection, 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, 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. 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 medical images.

François Chaumette received the M.Sc. degree from École Nationale Supérieure de Mécanique, Nantes, France, in 1987, and the Ph.D. degree in computer science from the University of Rennes, Rennes, France, in 1990.

Since 1990, he has been with INRIA, Rennes, where he is currently "Directeur de Recherches" and Head of the Lagadic group. He is currently on the Editorial Board of the *International Journal of Robotics Research*. His research interests include robotics and computer vision, especially visual servoing and active perception.

Dr. Chaumette was an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS from 2001 to 2005. He was the recipient of the AFCET/CNRS Prize for the best French thesis in automatic control in 1991 and the 2002 King-Sun Fu Memorial Best IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION Paper Award.