Modeling and 3D local estimation for in-plane and out-of-plane motion guidance by 2D ultrasound-based visual servoing

Rafik Mebarki, Alexandre Krupa and François Chaumette

Abstract— **This paper presents a new model-free visual servoing that is able to servo a robotized 2D ultrasound probe that interacts with a soft tissue object. It makes direct use of the B-mode ultrasound images in order to reach a desired one. This approach does not require the 3D model of the object nor its location in the 3D space. The visual features are based on image moments. The exact analytical form of the interaction matrix relating the image moments variation to the probe velocity is modelled. To perform model-free servoing, the approach combines the image points coordinates with the probe pose to estimate efficiently 3D parameters required in the control law. The approach is validated with simulation and experimental results showing its robustness to different errors and perturbations.**

Index Terms— **Visual servoing, ultrasound, model-free control, image moments, medical robotics.**

I. INTRODUCTION

Image-guided interventions are subject to an increasing number of investigations in order to be used in medicine in a near future thanks to the relevant and valuable assistance they are able to bring. Among the different imaging modalities, ultrasound (US) systems present inherent advantages of non-invasiveness, safety and portability. Especially, 2D US systems, beyond of being relatively non-expensive, provide images with high resolution at high speed that bring them a great ability to be used in real-time applications. In this paper, we propose a model-free visual servoing method to automatically position a 2D US probe actuated by a medical robot in order to reach a desired image. Potential applications are numerous. For instance, it can be employed in pathology analysis to automatically and accurately position the US probe in order to obtain a 2D cross-section image of a tumor having a maximum similarity with a previous one derived from a pre-operative 3D imaging system that was performed with the same (US) or other imaging modality (MRI, CT-SCAN). Also, during a biopsy or a radio frequency ablation procedure, reaching and maintaining an appropriate cross-section image may help the surgeon.

A visual servo control system of a robot arm actuating a 2D US probe for medical diagnostic has already been presented in [1]. However, only the in-plane motions of the probe have been considered to center the section of an artery in the image. Thus only three of the six degrees of freedom (DOFs) of the robot were controlled by visual servoing while its remaining axes being telemanipulated by an operator. An US-guided needle insertion robot for

The authors are with IRISA, INRIA Rennes-Bretagne Atlantique, Campus de Beaulieu, 35700 Rennes cedex, France {Rafik.Mebarki, Alexandre.Krupa, Francois.Chaumette}@irisa.fr

percutaneous cholecystostomy therapy has been presented in [2]. However, visual servoing was limited to control only the 2 DOFs of the needle driven part to automatically insert the needle tip. Therefore, the methods mentioned above can not perform tasks requiring the control of the probe out-of-plane motions. This is due to the fact that 2D US probes provide full information only in their observation plane but none at all outside.

In [3] [4], 3D US systems have been used to track a surgical instrument and position it to reach a target. However, 3D US systems present currently noticeable drawbacks such as a limited spatial perception, low voxel resolution leading to low quality imaging, and a high cost. Furthermore, performing real-time tasks with 3D US systems requires specialized powerful processors cards, which increase the global system cost. The servo of a surgical instrument actuated by a 4 DOFs robot arm by using an eye-to-hand motionless 2D US probe has been reported in [5] [6]. However, that approach is restrained to an interaction with an instrument forceps but not with real soft tissue objects. In [7], a visual servoing method to automatically position a robotized 2D US probe interacting with an egg-shaped object at a desired cross-section image has been presented. The visual features used are the coefficients of a third order polynomial fitting the current edge in the image. However, that method was dedicated only for egg-shaped objects. Moreover, the visual features used have no physical signification and are sensitive to modeling errors and features extraction perturbations. A visual servoing approach to automatically compensate soft tissue motions has been presented in [9]. That approach made use of the speckle correlation contained in successive B-mode US images. However, the developed servoing method is devoted for tracking and can not be used to reach a desired B-scan image starting from one totally different.

In [10], we described an US visual servoing method to control automatically a robotized 2D US probe interacting with an object of interest in order to reach a desired cross-section B-scan image. We used image moments in the control law. However this first method encounters limitations since it was designed for an ellipsoid-shaped object whose 3D parameters and pose were assumed to be coarsely known. The work presented in this paper is built upon this previous work and improves it in order to address its shortcomings. The first contribution of this paper, presented in Section II, is the determination of the exact analytical form of the interaction matrix related to any moment of a segmented US image, while only an approximation was derived in [10]. The second contribution, presented in Section III, deals with an

on-line estimation method for all the parameters involved in the interaction matrix. It allows the system to handle objects whose 3D shape is a priori unknown. The visual servoing control law is briefly derived in Section IV. Results obtained in both simulations and experiments are presented in Section V, demonstrating the validity of the approach.

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 design 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 and Fig. 2). 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. The objective 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 respectively the translational and the rotational velocity components and $\mathbf{L}_{m_{ij}}$ is the interaction matrix related to m_{ij} denoted by:

$$\mathbf{L}_{m_{ij}} = \begin{bmatrix} m_{vx} & m_{vy} & m_{vz} & m_{\omega x} & m_{\omega y} & m_{\omega z} \end{bmatrix} \tag{3}$$

The time variation of moments as function of the image points velocity is given by [10]:

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

that can be written:

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

where $(\dot{x}, \dot{y})$ is the velocity of an image point (x,y) belonging to the section $\mathcal{S}$. In order to determine the relation giving $\dot{m}_{ij}$ as 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 the image section $\mathcal{S}$ (see Fig. 1 and Fig. 2). The expression of the 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 the point $\mathbf{P}$ in the US probe frame $\{R_s\}$
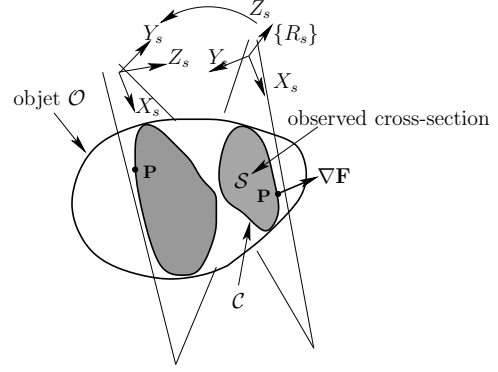

Fig. 1. Interaction between the 2D ultrasound probe and the concerned object

and in the object frame $\{R_o\}$ respectively. For the former it represents the image coordinates of the point $\mathbf{P}$. $^s\mathbf{R}_o$ is the rotation matrix defining the orientation of the object frame $\{R_o\}$ with respect to the probe frame $\{R_s\}$. $^s\mathbf{t}_o$ 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:

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

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 the vector $\mathbf{a}$. Thus, replacing (8) in (7), we obtain:

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

The point $\mathbf{P}$ results from the intersection of the US probe planar beam with the object. Consequently, in the 3D 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. Since $\mathbf{P}$ remains in the probe plane, its velocity expressed in the probe image frame is $^s\dot{\mathbf{P}} = (\dot{x}, \dot{y}, 0)$. Therefore, the relationship (9), that represents three constraints, has five unknowns ($^s\dot{\mathbf{P}}$ and $^o\dot{\mathbf{P}}$). In order to solve this system, two other constraints are needed.

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

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

Assuming that the object is rigid, we obtain:

$$\begin{aligned} \dot{\mathbf{F}}(^ox, \, ^oy, \, ^oz) & = & \frac{\partial \mathbf{F}}{\partial^ox} \, ^o\dot{x} + \frac{\partial \mathbf{F}}{\partial^oy} \, ^o\dot{y} + \frac{\partial \mathbf{F}}{\partial^oz} \, ^o\dot{z} = 0 \\ & \Leftrightarrow & ^o\nabla\mathbf{F}^\top \, ^o\dot{\mathbf{P}} = 0 \end{aligned} \tag{11}$$

where $^o\nabla\mathbf{F}$ is the gradient of $\mathbf{F}$. It represents the normal vector to the object surface at the point $\mathbf{P}$, as depicted in Fig. 1.

The second constraint concerns the direction of $^o\dot{\mathbf{P}}$. Indeed, $\nabla\mathbf{F}$ defines a plane, denoted by π on Fig. 3, and the constraint (11) states only that the vector $^o\dot{\mathbf{P}}$ lies on that plane. We need thus to define in which direction $^o\dot{\mathbf{P}}$ lies on π in such a way to match its corresponding point $^o\mathbf{P}(t+dt)$ on the contour $\mathcal{C}(t+dt)$. Since $^o\mathbf{P}(t+dt)$ is not physically the same as $^o\mathbf{P}(t)$, there are many possibilities of matching. However, since $^o\dot{\mathbf{P}}$ is involved by the probe out-of-plane motions, we consider $^o\dot{\mathbf{P}}$ in the direction of Z_s. Since Z_s does not lie on π, its projection on π is considered as the direction of $^o\dot{\mathbf{P}}$. To conclude, $^o\dot{\mathbf{P}}$ has to be orthogonal to the vector $^o\mathbf{N}$ (see Fig. 3), lying on π, defined by:

$$^o\mathbf{N} = {}^o\mathbf{Z}_s \times {}^o\nabla\mathbf{F} \tag{12}$$

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 can thus be written:
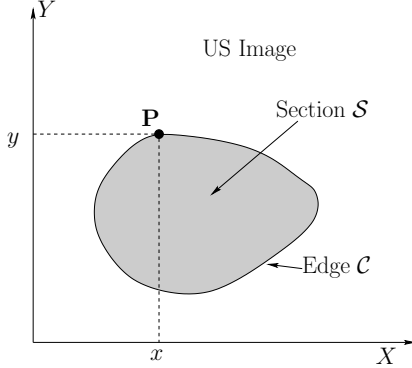
$$^o\mathbf{N}^\top {}^o\dot{\mathbf{P}} = 0 \tag{13}$$

Fig. 2.   US Image

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

$$^s\mathbf{R}_o^\top {}^s\dot{\mathbf{P}} = -{}^s\mathbf{R}_o^\top \boldsymbol{v} + {}^s\mathbf{R}_o^\top [{}^s\mathbf{P}]_\times \boldsymbol{\omega} + {}^o\dot{\mathbf{P}} \tag{14}$$

Multiplying (14) once by $^o\nabla\mathbf{F}^\top$ and then by $^o\mathbf{N}^\top$ and taking into account the constraints (11) and (13), yields to:

$$\begin{cases} ^o\nabla\mathbf{F}^\top {}^s\mathbf{R}_o^\top {}^s\dot{\mathbf{P}} = -{}^o\nabla\mathbf{F}^\top {}^s\mathbf{R}_o^\top \boldsymbol{v} + {}^o\nabla\mathbf{F}^\top {}^s\mathbf{R}_o^\top [{}^s\mathbf{P}]_\times \boldsymbol{\omega} \\ ^o\mathbf{N}^\top {}^s\mathbf{R}_o^\top {}^s\dot{\mathbf{P}} = -{}^o\mathbf{N}^\top {}^s\mathbf{R}_o^\top \boldsymbol{v} + {}^o\mathbf{N}^\top {}^s\mathbf{R}_o^\top [{}^s\mathbf{P}]_\times \boldsymbol{\omega} \end{cases} \tag{15}$$

Since we have:

$$\begin{cases} ^s\nabla\mathbf{F} = {}^s\mathbf{R}_o {}^o\nabla\mathbf{F} \\ ^s\mathbf{N} = {}^s\mathbf{R}_o {}^o\mathbf{N} = {}^s\mathbf{Z}_s \times {}^s\nabla\mathbf{F} \end{cases} \tag{16}$$

The relationships (15) become:

$$\begin{cases} ^s\nabla\mathbf{F}^\top {}^s\dot{\mathbf{P}} = -{}^s\nabla\mathbf{F}^\top \boldsymbol{v} + {}^s\nabla\mathbf{F}^\top [{}^s\mathbf{P}]_\times \boldsymbol{\omega} \\ ^s\mathbf{N}^\top {}^s\dot{\mathbf{P}} = -{}^s\mathbf{N}^\top \boldsymbol{v} + {}^s\mathbf{N}^\top [{}^s\mathbf{P}]_\times \boldsymbol{\omega} \end{cases} \tag{17}$$

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

$$\begin{cases} \dot{x} = -v_x - K_x v_z - y K_x \omega_x + x K_x \omega_y + y \omega_z \\ \dot{y} = -v_y - K_y v_z - y K_y \omega_x + x K_y \omega_y - x \omega_z \end{cases} \tag{18}$$

with:

$$\begin{cases} K_x = f_x f_z / (f_x^2 + f_y^2) \\ K_y = f_y f_z / (f_x^2 + f_y^2) \end{cases} \tag{19}$$

such that $^s\nabla\mathbf{F} = (f_x, f_y, f_z)$.

From (18) and (19) we can conclude that the image point velocity depends only of the image point position as for the in-plane motions (v_x, v_y, ω_z) and depends also of the normal vector to the object surface at that point as for the out-of-plane motions $(v_z, \omega_x, \omega_y)$.
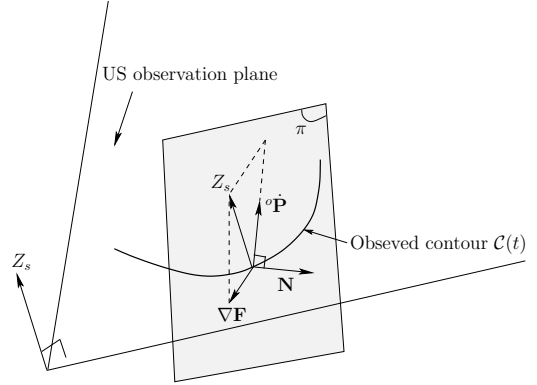
Fig. 3.   3D point velocity vector orientation

B. Image moments time variation modeling

Using the previous modeling of an image point velocity, the image moment time variation, $\dot{m}_{ij}$ can be developed. The image points for which the velocity was modelled belong to the countour $\mathcal{C}$ of $\mathcal{S}$. The image moment variation $\dot{m}_{ij}$ given by (5) has thus to be expressed as function of these points and their velocities. That is done by formulating $\dot{m}_{ij}$ on the contour $\mathcal{C}$ thanks to the Green's theorem:

$$\oint_\mathcal{C} F_x \, dx + \oint_\mathcal{C} F_y \, dy = \iint_\mathcal{S} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dx \, dy \tag{20}$$

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:

$$\dot{m}_{ij} = -\oint_\mathcal{C} [f(x, y) \dot{y}] \, dx + \oint_\mathcal{C} [f(x, y) \dot{x}] \, dy \tag{21}$$

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

$$m_{ij} = \frac{-1}{j+1} \oint_\mathcal{C} x^i y^{j+1} \, dx \tag{22}$$

and by setting $F_x = 0$ and $F_y = \frac{1}{i+1} x^{i+1} y^j$, we also have:

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

Replacing now (18) in (21), and then using (22) and (23), we finally obtain the elements of $\mathbf{L}_{m_{ij}}$ defined in (3):

$$\left\{ \begin{array}{rcl} m_{vx} & = & -i\,m_{i-1,j} \\ m_{vy} & = & -j\,m_{i,j-1} \\ m_{vz} & = & {}^x m_{ij} \; - \; {}^y m_{ij} \\ m_{\omega x} & = & {}^x m_{i,j+1} \; - \; {}^y m_{i,j+1} \\ m_{\omega y} & = & -{}^x m_{i+1,j} \; + \; {}^y m_{i+1,j} \\ m_{\omega z} & = & i\,m_{i-1,j+1} - j\,m_{i+1,j-1} \end{array} \right. \tag{24}$$

where:

$$\left\{ \begin{array}{rcl} {}^x m_{ij} & = & \oint_{\mathcal{C}} x^i\,y^j\,K_y\,dx \\ {}^y m_{ij} & = & \oint_{\mathcal{C}} x^i\,y^j\,K_x\,dy \end{array} \right. \tag{25}$$

Similarly to the image point velocity, we can note that the terms corresponding to the in-plane motions only depend on the measurements in the image, while the terms corresponding to the out-of-plane motions require the knowledge of the normal of the object for each point of the contour. This normal vector could be derived if a pre-operative 3D model of the object was available. That would also necessitates a difficult calibration step to localize the object in the sensor frame. To overcome this limitation, we propose in the next section a method using the successive 2D US images to estimate the normal vector.

III. NORMAL VECTOR ON-LINE ESTIMATION

Let $\mathbf{d_i}$ be the tangent vector to the cross-section image contour $\mathcal{C}$ at point $\mathbf{P}$ such that it belongs to that observed image (see Fig. 4). Let $\mathbf{d_t}$ be an other tangent vector to the object surface also at $\mathbf{P}$ defined such that, expressed in the sensor frame $\{R_s(t)\}$, we have:

$$ {}^s\nabla\mathbf{F} = {}^s\mathbf{d_i} \times {}^s\mathbf{d_t} \tag{26}$$

Fig. 4. Surface Normals

Since ${}^s\mathbf{d_i}$ can directly be measured from the current image, we just need to estimate ${}^s\mathbf{d_t}$ to obtain ${}^s\nabla\mathbf{F}$.
The vector ${}^s\mathbf{d_t}$ lies out of the US image plane. Consequently, it can not be obtained only from the observed US image. We propose in what follows an approach to estimate ${}^s\mathbf{d_t}$ from successive images. The principle is to estimate, for each point extracted from the contour $\mathcal{C}$, the 3D straight line that fits a set of successive points extracted from previous successive

images at the same image polar positions (see Fig. 4 and Fig. 5). 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. 5). The points $\mathbf{P}(t)$ and $\mathbf{P}(t + dt)$, extracted respectively from $\mathcal{C}(t)$ and $\mathcal{C}(t + dt)$ at the same polar position, lie on the object surface, and thus, a straight line $\mathcal{D}$ that passes through these points is tangent to the object surface. The direction of $\mathcal{D}$ is nothing but the tangent vector $\mathbf{d_t}$ we want to estimate. In theory, two successive extracted points are enough to estimate that direction. More will be used in practice to improve the robustness of the estimation. However, since the object surface may be curved, different weights will be associated to the points to take into account this curvature. The complete method is detailed bellow.

The straight line $\mathcal{D}$ is defined by its direction ${}^i\mathbf{d_t} = (d_x,\, d_y,\, d_z)$ and a point ${}^i\mathbf{P}_0$ lying on it. They are expressed in the initial probe frame $\{R_i\}$. The point $\mathbf{P}_0$ is defined such that each point ${}^i\mathbf{P} = ({}^ix,\, {}^iy,\, {}^iz)$, expressed in $\{R_i\}$, lying on $\mathcal{D}$ satisfies the relation:

$$\left({}^i\mathbf{P_0} - {}^i\mathbf{P}\right) \times {}^i\mathbf{d} = 0 \tag{27}$$

Using a set of points $\mathbf{P}$ from the successive US images, the direction ${}^i\mathbf{d_t}$ can be estimated once the points are expressed in $\{R_i\}$ using the robot odometry. From (27) we derive the following minimal relationship expressing the constraint that the point ${}^i\mathbf{P}$ lies on $\mathcal{D}$:

Fig. 5. Soft tissue cross-section 3D evolution

$$\left\{ \begin{array}{rcl} {}^ix & = & \eta_1\,{}^iz + \eta_0 \\ {}^iy & = & \tau_1\,{}^iz + \tau_0 \end{array} \right. \tag{28}$$

where $\eta_1 = d_x/d_z$ and $\tau_1 = d_y/d_z$. η_0 and τ_0 are function of ${}^i\mathbf{P}_0$ and ${}^i\mathbf{d}$. Thus $\theta = (\eta_1,\, \tau_1,\, \eta_0,\, \tau_0\,)$ is the vector parameters to estimate. The system (28) is written as follows:

$$\mathbf{Y} = \mathbf{\Phi}^\top \theta \tag{29}$$

with

$$\mathbf{Y} = ({}^ix,\, {}^iy) \text{ and } \mathbf{\Phi}^\top = \left(\begin{array}{cccc} {}^iz & 0 & 1 & 0 \\ 0 & {}^iz & 0 & 1 \end{array} \right) \tag{30}$$

The goal consists in computing the estimate $\hat{\theta}_{[t]}$ of θ at time t. For that, we propose to use a stabilized recursive

least-squares algorithm [11]. It consists in minimizing the following quadratic sum of the estimation errors [12]:

$$J(\hat{\theta}_{[t]}) = \sum_{i=t_0}^{t} \lambda_\theta^{(i-t_0)} \left(\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\theta}_{[i]}\right)^\top \left(\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\theta}_{[i]}\right)$$

(31)

where $0 < \lambda_\theta \leq 1$ is a weight assigned to the different estimation errors in order to give more importance to the current measure than to the previous ones. The estimate $\hat{\theta}$ is given by:

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

(32)

where $\mathbf{F}_{[t]}$ is a covariance matrix defined by:

$$\mathbf{F}_{[t]}^{-1} = \lambda_\theta \, \mathbf{F}_{[t-1]}^{-1} + \mathbf{\Phi}_{[t]} \, \mathbf{\Phi}_{[t]}^\top + (1 - \lambda_\theta) \, \beta_0 \, \mathbf{I}_4$$

(33)

where $\mathbf{I}_4$ is a 4×4 identity matrix. The initial parameters are set to $\mathbf{F}_{t_0} = f_0 \, \mathbf{I}_4$ with $0 < f_0 \leq 1/\beta_0$ and $\hat{\theta}_{t_0} = \theta_0$. The objective of the stabilization term $(1 - \lambda_\theta) \, \beta_0 \, \mathbf{I}_4$ is to prevent the matrix $\mathbf{F}_t^{-1}$ becoming singular when there is not enough excitation in the input signal $\mathbf{Y}$, which occurs when there is no out-of-plane motion.

We then obtain after normalization ${}^i\mathbf{d_t} = (\hat{\eta}_1, \hat{\tau}_1, 1)/\|(\hat{\eta}_1, \hat{\tau}_1, 1)\|$ and ${}^s\mathbf{d_t} = {}^s\mathbf{R}_i \, {}^i\mathbf{d_t}$, which allows estimating $\nabla\mathbf{F}$ using (26). In practice, we use around 400 image points to characterize the contour $\mathcal{C}$.

IV. VISUAL SERVOING

The visual features are selected as combinations of the image moments such that $\mathbf{s} = \mathbf{s}(m_{ij})$. they are given by:

$$\mathbf{s} = (x_g, \ y_g, \ \alpha, \ \sqrt{a}, \ l_1)$$

(34)

where x_g and y_g are the gravity center coordinates of the cross-section $\mathcal{S}$ in the image, α its main orientation angle, a its area and l_1 its main axis. More precisely, we have:

$$\begin{cases} x_g &= m_{10}/m_{00} \\ y_g &= m_{01}/m_{00} \\ a &= m_{00} \\ \alpha &= \frac{1}{2}\arctan\left(\frac{2\,\mu_{11}}{\mu_{20}+\mu_{02}}\right) \\ l_1^2 &= \frac{2}{a}\left(\mu_{02}+\mu_{20}+\sqrt{(\mu_{20}-\mu_{02})^2+4\mu_{11}^2}\right) \end{cases}$$

(35)

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

The time variation of the visual features vector as function of the probe velocity is written using (24) and (25) as follows:

$$\dot{\mathbf{s}} = \mathbf{L_s} \, \mathbf{v}$$

(36)

where:

$$\mathbf{L_s} = \begin{bmatrix} -1 & 0 & x_{g_{vz}} & x_{g_{\omega x}} & x_{g_{\omega y}} & y_g \\ 0 & -1 & y_{g_{vz}} & y_{g_{\omega x}} & y_{g_{\omega y}} & -x_g \\ 0 & 0 & \alpha_{vz} & \alpha_{\omega x} & \alpha_{\omega y} & -1 \\ 0 & 0 & \frac{a_{vz}}{2\sqrt{a}} & \frac{a_{\omega x}}{2\sqrt{a}} & \frac{a_{\omega y}}{2\sqrt{a}} & 0 \\ 0 & 0 & l_{1_{vz}} & l_{1_{\omega x}} & l_{1_{\omega y}} & 0 \end{bmatrix}$$
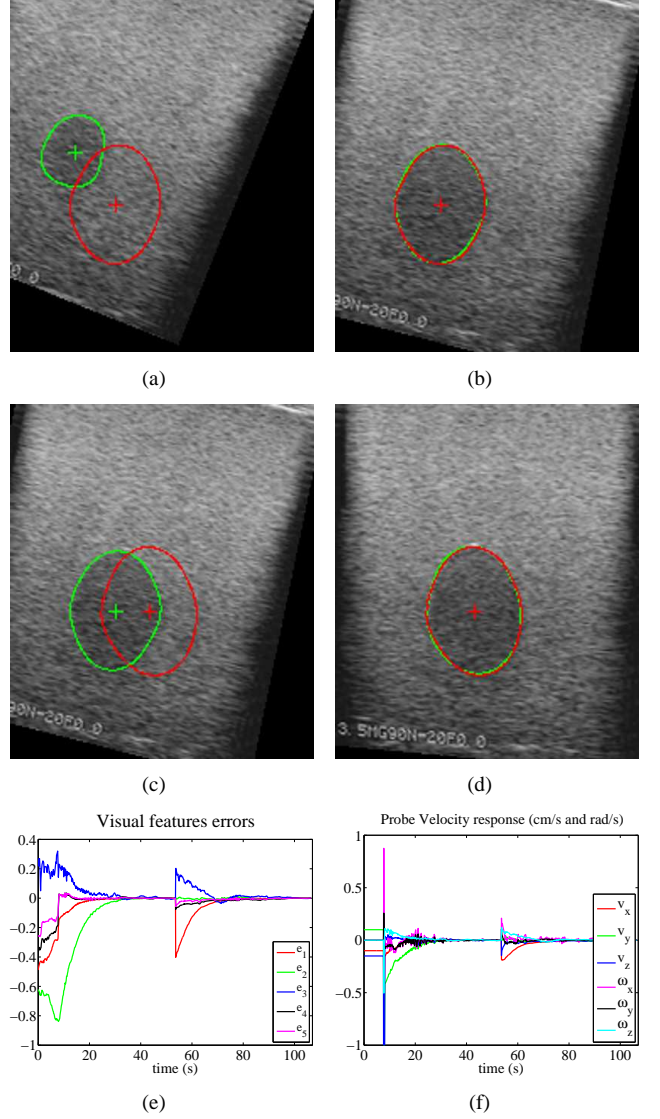
(37)

Fig. 6. Results from the ultrasound simulator (the current contour is in green and the desired one in red): (a) Initial and first target image - (b) First target reached after visual servoing - (c) A second target image is sent to the virtual robot - (d) The second target image is reached after visual servoing - (e) Visual error time response - (f) Control velocity applied to the probe

The expressions of some elements involved in (37) are not detailed here for a lack of place. On the other hand, we can notice that the 4th and the 5th component of $\mathbf{s}$ ($\sqrt{a}$ and l_1) are invariant to the probe in-plane motions (v_x, v_y, ω_z) and therefore are expected to be well adapted for the out-of-plane motions of the probe (v_z, ω_x, ω_y). On the other hand, the remaining features x_g, y_g and α present a good decoupling for the in-plane motions owing to the triangular part they form. We finally note that the 6th feature remains to be found to control all the 6 DOFs.

A very classical control law is used [8]:

$$\mathbf{v_c} = -\hat{\mathbf{L}}_\mathbf{s}^+ \, \mathbf{K} \, (\mathbf{s} - \mathbf{s}^*)$$

(38)

where $\mathbf{v_c}$ is the US probe instantaneous velocity sent to the low-level robot controller, $\mathbf{K}$ is a positive gain matrix, $\mathbf{s}^*$
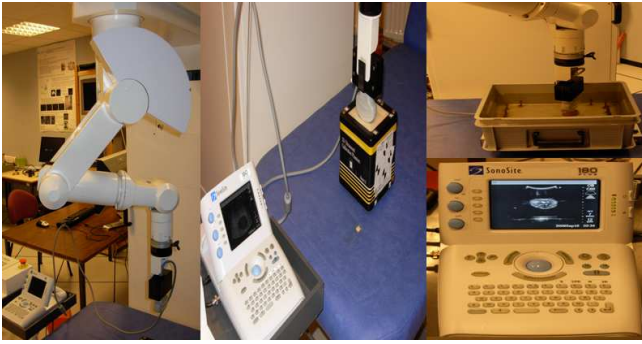
Fig. 7. Experimental setup: (left) medical robot equipped with an US probe transducer - (middle) interaction with an US phantom - (right) interaction with a lamb kidney immersed in a water-filled tank

is the desired visual features vector, and $\hat{\mathbf{L}}_{\mathbf{s}}^{+}$ is the pseudo inverse of the estimated interaction matrix.

V. RESULTS

Image processing and control law have been implemented in C++ under Linux operating system. The computations are performed on a PC computer equipped with a 3GHz Dual Core Xeon Intel Processor. The image processing used to extract the US image contour in real-time is described in [13].

In both simulations and experiments, the robotic task consists to reach a first target image and, once this first task has been achieved, a second target image is sent to the visual control system after a certain time span. This allows to validate the ability to re-estimate the normal vector when there has been no out-of-plane motions during that span and also, to analyze the robot behavior during the second task since the estimated normal vector is supposed to have reached the real one. Before launching the servoing to reach the first image target, the robot is moved in open-loop for 10 iterations. This allows providing an initial value to the normal vector. The second task does not need this open-loop motion since an initial value is available from the previous task. At the starting time t_0, the initial parameters vector θ_0 is randomly set to (0.2, 0.2, -0.2, -0.2) for all the points of the contour. The initial contour is extracted after a user had roughly indicated it by hand-clicking.

A. Simulation results on a realistic ultrasound volume

We have first used a simulator allowing the interaction of a virtual robotized 2D US probe with a realistic volume of an egg-shaped object. This volume was reconstructed from 100 real B-scan images. This simulator provides the observed US image (see Fig. 6).

The control gain matrix is set to $\mathbf{K} = 0.7 \cdot diag(1, 1, 1, 1/1.4, 1)$. The control gain concerning the area is decreased in order to avoid abrupt robot motion that occur just after launching the servoing. Since at that time the normal vector is not yet correctly estimated. The stabilized least-squares algorithm parameters are set to $\lambda_\theta = 0.8$, $f_0 =$ 1e8, and $\beta_0 = 1/(20.f_0)$. The results[1] are depicted in Fig. 6.
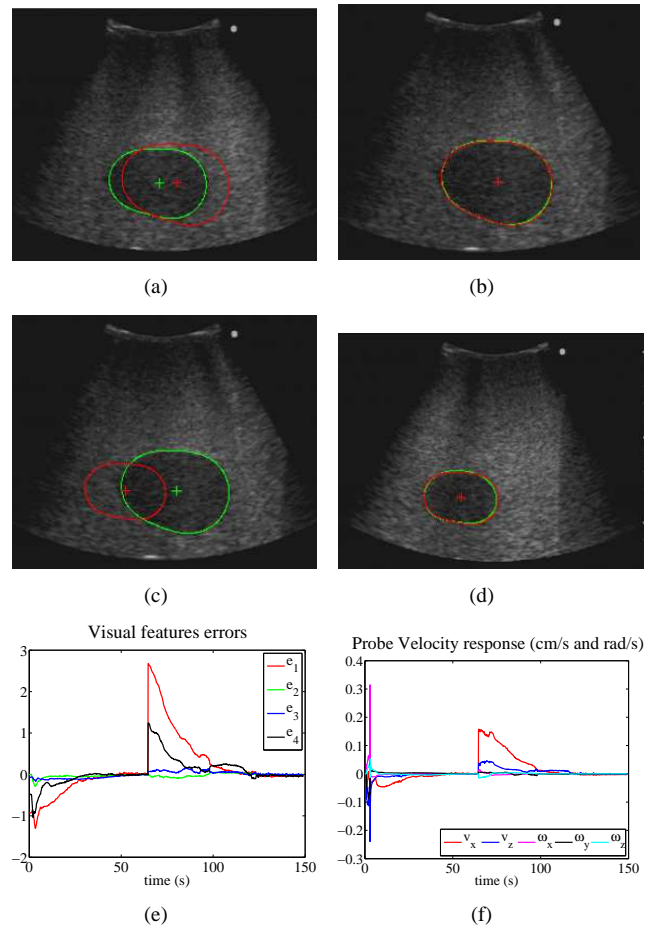
[1]A short video accompanies the paper

Fig. 8. Experimental results with the ultrasound phantom (the current contour is in green and the desired one in red): (a) Initial and first target image - (b) First target reached after visual servoing - (c) A second target image is sent to the robot - (d) The second target image is reached after visual servoing - (e) Visual error time response (cm, cm, rad, cm)- (f) Control velocity applied to the probe

The visual features errors converge roughly exponentially to zero. This validates the method on realistic US images that are very noisy and thus the robustness to different perturbations like those generated by the contour detection.

B. Experimental results on an ultrasound phantom

We have also tested the method on an US phantom using a 6 DOFs medical robot similar to the Hippocrate system [14] and a 5-2 MHz broadband ultrasound transducer (see Fig 7). In this case, the US transducer is in contact with the phantom, therefore, the probe velocity v_y is constrained by force control to keep the transducer on the phantom surface with a 2N force. In this experiment, we removed l_1 from the visual features vector since we noticed its coupling with the area a, due probably to the shape of the observed object in the phantom. The selected visual features are then $\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a})$. The control gain matrix is set to $\mathbf{K} = 0.07 \cdot diag(1, 1, 1, 1/1.4)$. The stabilized least-squares algorithm parameters have been set to $\lambda_\theta = 0.95$, $f_0 = $ 1e8, and $\beta_0 = 1/(20.f_0)$. The experimental results[1] are depicted in Fig. 8. The visual features errors converge to
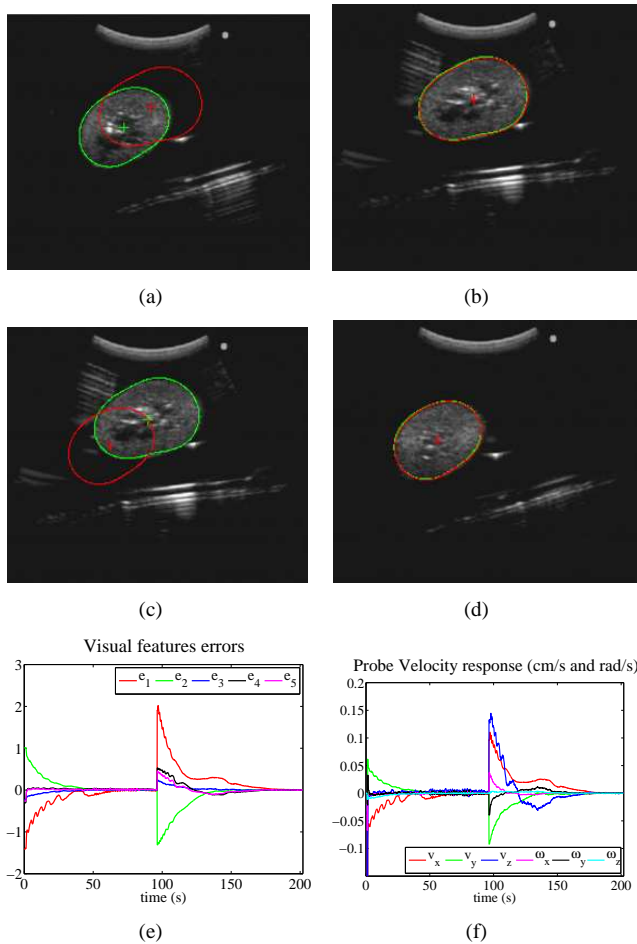
Fig. 9. Experimental results with lamb kidney (the current contour is in green and the desired one in red): (a) Initial and first target image - (b) First target reached after visual servoing - (c) A second target image is sent to the robot - (d) The second target image is reached after visual servoing - (e) Visual error time response (cm, cm, rad, cm, cm)- (f) Control velocity applied to the probe

zero and the reached image is the desired one. This validates experimentally the method on a phantom where the images are very noisy and where the experimental conditions are more close to real ones.

C. Ex-vivo experimental results

Finally, we have tested the method in ex-vivo experiments on a motionless real lamb kidney immersed in a water-filled tank (see Fig. 7). The selected visual features are now $\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, l_1)$. The control gain matrix is set to $\mathbf{K} = 0.07 \cdot diag(1, 1, 1, 1/1.6, 1)$. The stabilized least-squares algorithm parameters have been set to $\lambda_\theta = 0.8$, $f_0 = 1e5$, and $\beta_0 = 1/(20.f_0)$. The experimental results[1] are depicted in Fig. 9. Once again, the visual features errors converge to zero and the reached image is the desired one. These results validate the method on motionless real soft tissue.

VI. CONCLUSIONS AND FUTURE WORKS

This paper presented a new model-free visual servoing method based on 2D US images. We addressed shortcomings that were present in our previous work [10]. Moreover, we

have improved and broadened the approach. First, in this work, we have derived the exact analytical form of the interaction matrix related to the image moments. Then, we have endowed the servo system to deal with objects of unknown shape by developing a model-free servoing. For that, the normal vector involved in the interaction matrix has been estimated by a 3D on-line method. Successful results have been obtained in both simulations, experiments on an US phantom, and ex-vivo experiments on a lamb kidney. These results validate the method and its robustness to different perturbations especially those generated by the US image that, inherently, is very noisy. Thanks to the added stabilization term in the on-line estimation, the normal vector identification was achieved even when no out-of-plane motions occurred. This servoing method is generic since it can be applied to different kinds of imaging systems that provide, like US, full information in their observation plane, such as MRI and CT-SCAN. Challenges that remain to be tackled concern the cases of moving and deformable objects.

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. on Robotics*, vol. 18, no. 1, pp. 11-23, february 2002.

[2] J. Hong, T. Dohi, M. Hashizume, K. Konishi, N. Hata, An ulrasound-driven needle insertion robot for percutaneous cholecystostomy, *Physics in Medecine and Biology*, 49(3), pp. 441-445, 2004.

[3] J. Stoll, P. Novotny, R. Howe, P. Dupont, Real-time 3D ultrasound-based servoing for a surgical instrument, *in IEEE Int. Conf. on Robotics and Automation*, Orlando, Florida, May 2006.

[4] P.M. Novotny, J.A. Stoll, P. Dupont, R.D. Howe. Real-time visual servoing of a robot using three-dimensional ultrasound, *in IEEE Int. Conf. on Robotics and Automation*, pp. 2655-2660, Roma, Italy, May 2007.

[5] M. A. Vitrani, H. Mitterhofer, N Bonnet, G. Morel. Robust ultrasound-based visual servoing for beating heart intracardiac surgery, *in IEEE Int. Conf. on Robotics and Automation*, pp. 3021-3027, Roma, Italy, April 2007.

[6] M. Sauvée, P. Poignet, E. Dombre, Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control, *International Journal of Robotics Research*, vol. 27, no 1, January 2008.

[7] W. Bachta and A. Krupa. Towards ultrasound image-based visual servoing. *In Proc. IEEE Int. Conf. on Robotics and Automation,* pp. 4112-4117, Orlando, Florida, May 2006.

[8] B. Espiau, F. Chaumette and P. Rives. A new approach to visual servoing in robotics. *In IEEE Trans. on Robotics*, vol. 8, pp. 313-326, June 1992.

[9] A. Krupa, G. Fichtinger and G. D. Hager. Full motion tracking in ultrasound using image speckle information and visual servoing. *In IEEE Int. Conf. on Robotics and Automation ICRA'07*, p. 2458-2464, Roma, Italy, May 2007.

[10] R. Mebarki, A. Krupa and F. Chaumette. Image moments-based ultrasound visual servoing. *In IEEE Int. Conf. on Robotics and Automation, ICRA'08,* pp. 113-119, Pasadena, CA, USA, May 2008.

[11] G. Kreisselmeier. Stabilized least-squares type adaptive identifiers.*In IEEE Trans. on Automatic Control,* vol. 35, n. 3, pp. 306-309, March 1990.

[12] R. Johansson, "System modeling and identification", *Prentice-hall*, Englewood Cliffs, 1993.

[13] R. Mebarki, A. Krupa, C. Collewet. Automatic guidance of an ultrasound probe by visual servoing based on B-mode image moments. *In 11th Int. Conf. on Medical Image Computing and Computer Assisted Intervention, MICCAI'08,* New York, USA, September, 2008.

[14] 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, *Medical Image Analysis (MedIA)*, vol. 3, no. 3, pp. 285–300, 1999.