Automatic calibration of a robotized 3D ultrasound imaging system by visual servoing

A. Krupa

HAL Id: inria-00350295

https://inria.hal.science/inria-00350295

Submitted on 6 Jan 2009

Automatic Calibration of a Robotized 3D Ultrasound Imaging System by Visual Servoing

Alexandre Krupa

IRISA - INRIA Rennes

Campus de Beaulieu, 35042 Rennes Cedex, France

Email: Alexandre.Krupa@irisa.fr

Abstract— **Three-dimensional free-hand ultrasound imaging consists of capturing a set of ultrasound images with a 2D ultrasound system and their respective locations in order to position them in a 3D reference frame. Usually the clinician performs the acquisition manually through the use of an optical or magnetic localization system attached to the ultrasound probe. To assist the clinician, we propose to use a robotic system to automatically move the ultrasound probe and measure its position. As for manual 3D ultrasound imaging, it is crucial to know precisely the spatial calibration parameters of the ultrasound system in order to perform accurate 3D imaging. Therefore, we propose to automate the spatial calibration procedure. A robotic task is developed to automatically position the ultrasound image on the intersection point of a cross-wire phantom used for spatial calibration. To perform this task, a new visual servoing technique based on 2D ultrasound images is used to control automatically the motion of the ultrasound probe held by a medical robot.**

Index Terms— **Visual servoing, 3D ultrasound imaging, medical robotics.**

I. Introduction

Three-dimensional ultrasound imaging is used in numerous medical applications such as vascular imaging, cardiology, obstetrics, and neurosurgery. This imaging modality is, for example, used to detect or survey the evolution of atherom plates attached to the walls of arteries that can be deadly for the patient. Two different techniques are used to perform 3D ultrasound (US) imaging. The first one uses a 3D ultrasound sensor, however it currently provides low voxel resolution and a small field of view. The second technique, referred to as "3D free-hand ultrasound imaging" consists of measuring the relative displacement between each image captured by a two-dimensional ultrasound system in order to position them in a 3D frame [1]. For this second technique a localization system, which can be magnetic, optic, acoustic, or mechanical is fixed to the 2D US probe. In the context of 3D free-hand ultrasound imaging, we propose to use a medical robotic manipulator as the localization system and also to automatically move the 2D US probe. This will allow repeat examinations of a patient on different dates in order to observe quantitatively the pathology evolution under the same conditions. Towards that end, a robotic system [2] has already been developed to perform 3D US acquisition of cardiovascular pathologies by automatically moving a 2D ultrasound probe along a given trajectory. Nevertheless, to ensure the accuracy of the 3D reconstruction, it is crucial that the two scaling factors of the ultrasound image and the relative position between the ultrasound image reference frame and the localization system (which is in this case the robot end-effector) are known precisely. Usually, a spatial calibration procedure is manually performed before the 3D acquisition. The approach of the calibration procedure is to capture a set of 2D US images of a known object for different measured probe positions and then to off-line estimate the spatial parameters by coupling visual features extracted from each US image to the geometrical properties of the object. For example, in [3], a method is presented whereby the intersection point of a fixed, cross-wire phantom, immersed in water, has to be centered in the US image for different positions and orientations of the US probe. In [4] another method is presented using a plane phantom.

The calibration procedure is laborious for the clinician and its quality influences the accuracy of the 3D reconstruction. Thus, in this paper, we propose to automate the spatial calibration procedure by developing a robotic task using a new ultrasound image-based visual servoing. Up until now, only a few studies have been made on visual servoing using information from 2D ultrasound images. In [5], visual servoing is used to center within the 2D US image a point corresponding to the section center of an artery during the probe displacement along a 1D trajectory. Of the 6 degrees of freedom (DOF) available to the robot holding the US probe, only 3 DOF in the US observation plane are controlled by visual servoing, while the 3 other DOF are teleoperated by the user. In another work [6], the authors present a robotic system for needle insertion with the US probe rigidly fixed to the base of a small 2 DOF robot held by a 5 passive DOF mechanical architecture. The probe is positioned in such a way that the needle is always visible in the US image as a straight line. Once again, only 2 DOF (1 translation and 1 rotation) in the US observation plane are controlled by visual servoing. More recently, a study [7] has been presented where 4 DOF, which are not necessary in the observation plane of the probe, are controlled by visual servoing in order to automatically move a robotized laparoscopic instrument.

The paper is organized as follows: the next Section briefly introduces the principle of the Detmer spatial calibration method [3] that we propose to automate. In the Section III, the robotic task is formulated, the visual features are defined, and the visual servoing control law is developed. The image processing used to extract the visual features from the US image is described in Section IV. Then Section V presents

experimental results of the robotic task performed with a 6 DOF medical robot and the validation of the calibration.

II. SPATIAL CALIBRATION

Let $\mathcal{F}_s$ and $\mathcal{F}_n$ be respectively the cartesian frames attached to the probe sensor and the end-effector of the robot as illustrated in Fig. 1.a. $\mathcal{F}_0$ is the base frame of the robot and $\mathcal{F}_c$ is the frame chosen for the 3D reconstruction. The 3D coordinates expressed in $\mathcal{F}_c$ of a point P belonging to the ultrasound plane is given by the following relation:

$$^c\mathbf{P} = {}^c\mathbf{T}_0\ {}^0\mathbf{T}_n\ {}^n\mathbf{T}_s\ {}^s\mathbf{P} \tag{1}$$

where $^c\mathbf{T}_0$, $^0\mathbf{T}_n$ and $^n\mathbf{T}_s$ are the different homogeneous transformations between each frame and $^s\mathbf{P} = (x, y, 0, 1)$ denotes the homogeneous coordinates of P expressed in the probe frame (the left subscript ^{c} denotes that the components are expressed in $\mathcal{F}_c$). Since the image coordinates are measured in pixel, the following variable transformations have to be made: $x = (u - u_0)S_x$ and $y = (v - v_0)S_y$, where (u, v) and (u_0, v_0) are respectively the pixel coordinates of P and the image center, S_x and S_y are the height and width of a pixel (scaling factors).

In practice, in order to obtain a good accuracy of the coordinates $^c\mathbf{P}$ expressed in $\mathcal{F}_c$ (with a given $^c\mathbf{T}_0$), it is crucial that the robot provides a very low position error on $^0\mathbf{T}_n$ and that the spatial calibration parameters of the system which include the rigid transformation $^n\mathbf{T}_s$ and the two scaling factor (S_x, S_y) of the image are known precisely. Several methods were proposed in the literature to estimate these parameters. We apply the one introduced by Detmer [3] which consists of capturing a set of ultrasound images of a fixed point for different positions and orientations of the end-effector. To do this in practice, the center of a cross-wire phantom immersed in water is used as the point. Then if we consider that this point corresponds to the origin of $\mathcal{F}_c$ we obtain the following relation for each image:

$$^c\mathbf{P} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^c\mathbf{T}_0\ {}^0\mathbf{T}_n\ {}^n\mathbf{T}_s \begin{pmatrix} (u - u_0)S_x \\ (v - v_0)S_y \\ 0 \\ 1 \end{pmatrix} \tag{2}$$

Several measurements allow then to write a system of non-linear equation which can be soled by an iterative method in order to obtain the rigid homogeneous transformation $^n\mathbf{T}_s$ and the image scale factors S_x and S_y (see [3], [8] for more details). Nevertheless in practice it is really difficult to position the US plane on the intersection of the cross-wire phantom. So we propose to automatically perform it by visual servoing.

III. VISUAL SERVOING

First, let us formulate the robotic task to achieve. The goal is to position the intersection point P^* (see Fig. 1.b) between the two converging straight lines $\mathcal{D}_1$ and $\mathcal{D}_2$ forming the cross-wire phantom in the ultrasound image with a set of different positions and orientations of the probe. So the visual task consists of centering the points P_1 and P_2 on a desired
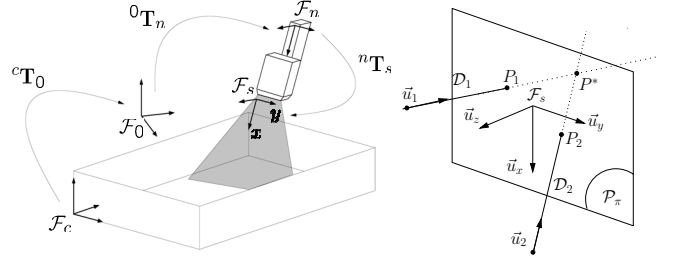

Fig. 1. (a) 3D ultrasound imaging with a 2D probe - (b) Coupling between the ultrasound observation plane and the two converging straight lines forming the cross-wire object

target in the image plane $\mathcal{P}_\pi$. We choose as visual features $\mathbf{s} = (x_1, y_1, x_2, y_2)$ the 2D coordinates $\mathbf{p}_1 = (x_1, y_1)$ and $\mathbf{p}_2 = (x_2, y_2)$ of points P_1 and P_2 expressed in the ultrasound probe frame $\mathcal{F}_s$.

The variation of the visual information $\mathbf{s}$ to the kinematic screw $\mathbf{v} = (\boldsymbol{v}, \boldsymbol{\omega})$ of the US probe is given by :

$$\dot{\mathbf{s}} = \mathbf{L}_\mathbf{s}\mathbf{v} = \begin{bmatrix} -1 & 0 & \frac{u_{1_x}}{u_{1_z}} & \frac{u_{1_x}}{u_{1_z}}y_1 & -\frac{u_{1_x}}{u_{1_z}}x_1 & y_1 \\ 0 & -1 & \frac{u_{1_y}}{u_{1_z}} & \frac{u_{1_y}}{u_{1_z}}y_1 & -\frac{u_{y_1}}{u_{1_z}}x_1 & -x_1 \\ -1 & 0 & \frac{u_{2_x}}{u_{2_z}} & \frac{u_{2_x}}{u_{2_z}}y_2 & -\frac{u_{2_x}}{u_{2_z}}x_2 & y_2 \\ 0 & -1 & \frac{u_{2_y}}{u_{2_z}} & \frac{u_{2_y}}{u_{2_z}}y_2 & -\frac{u_{y_2}}{u_{2_z}}x_2 & -x_2 \end{bmatrix}\mathbf{v} \tag{3}$$

Where $\boldsymbol{v} = (v_x, v_y, v_z)$ is the translational velocity vector and $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ is the angular velocity vector of the probe frame with respect to the robot base frame. In classical visual servoing, $\mathbf{L}_\mathbf{s}$ is called the interaction matrix (see [9]) and is determined from the geometrical model of the considered system. For our system, it is analytically obtained from the modelling of the coupling between the ultrasound observation plane and the two converging straight lines (modelling details are given in [10]). As we can note $\mathbf{L}_\mathbf{s}$ depends only on the components of the unitary vectors $^s\mathbf{u}_1 = (u_{1_x}, u_{1_y}, u_{1_z})$ and $^s\mathbf{u}_2 = (u_{2_x}, u_{2_y}, u_{2_z})$ of the straight lines $\mathcal{D}_1$ and $\mathcal{D}_2$ and the 2D coordinates $\mathbf{p}_1$ and $\mathbf{p}_2$, all expressed in the probe frame. The condition to compute $\mathbf{L}_\mathbf{s}$ is that u_{1_z} or $u_{2_z} \neq 0$. This is verified when the straight lines are not collinear to the observation plane.

The visual servoing task can be expressed as a regulation to zero of the visual error [9]:

$$\mathbf{e}_1(\mathbf{r}(t)) = \mathbf{s}(\mathbf{r}(t)) - \mathbf{s}^* \tag{4}$$

where $\mathbf{s}^*$ is the reference value of the visual features to be reached and $\mathbf{s}$ is the value of the visual features currently observed by the US probe. The features depend on the relative position $\mathbf{r}$ between the probe and the scene. Note that the rank of $\mathbf{L}_\mathbf{s}$ is 4 except when the two points are joined together. In this last case, the rank is reduced to 3. This means that when the points are centered on a target in the image, the dimension of the visual task switches from 4 to 3 and then only 3 DOF are constrained. We will see in the Section V how to cope with the rank change. Consequently, as the robot holding the ultrasound probe has $n = 6$ DOF, it is possible to

use the remaining 3 DOF to perform a secondary task such as changing the orientation of the probe.

A. Redundancy formalism

In our approach we use the redundancy formalism [11] to combine the visual task and the task of changing the orientation of the probe. This formalism has first been used for visual servoing in [9], and in numerous applications since (e.g. avoiding visual features occlusion [12], or human-machine cooperation using vision control [13]). The idea is to use the DOF left by a main task $\mathbf{e}_1$ of dimension $m < n$, to realize a secondary task $\mathbf{g}^\top = \frac{\partial h}{\partial \mathbf{r}}$ at best without disturbing the first one. Generally, the realization of a secondary goal is expressed as a minimization of a cost function h under the constraint that the main task is achieved, i.e $\mathbf{e}_1 = 0$. The determination of DOF which are left by the main task requires the computation of the null space of the interaction matrix $\mathbf{L}_{\mathbf{e}_1}$ of the task $\mathbf{e}_1$. We have of course $\mathbf{L}_{\mathbf{e}_1} = \mathbf{L}_{\mathbf{s}}$. Performing the two tasks consists then in minimizing a global task function given by [11]:

$$\mathbf{e} = \hat{\mathbf{L}}_{\mathbf{s}}^+ \mathbf{e}_1 + (\mathbf{I}_n - \hat{\mathbf{L}}_{\mathbf{s}}^+ \hat{\mathbf{L}}_{\mathbf{s}}) \mathbf{g}^\top \tag{5}$$

where $\hat{\mathbf{L}}_{\mathbf{s}}^+$ is the pseudo-inverse of an estimation $\hat{\mathbf{L}}_{\mathbf{s}}$ of the interaction matrix and $(\mathbf{I}_n - \hat{\mathbf{L}}_{\mathbf{s}}^+ \hat{\mathbf{L}}_{\mathbf{s}})$ an orthogonal projector operator which projects $\mathbf{g}^\top$ in the null space of $\hat{\mathbf{L}}_{\mathbf{s}}$ in order that the second task does not disturb the first one.
In practice the secondary task we propose consists of minimizing the $\theta\mathbf{u}$ angle vector of the relative rotation between the current orientation of the probe and a desired orientation. To do this we define the following quadratic cost function:

$$h = \frac{1}{2}\theta\mathbf{u}^\top \theta\mathbf{u} \tag{6}$$

and by computing the gradient of h and its partial time derivative, we get:

$$\mathbf{g} = \begin{bmatrix} \mathbf{0}_{[1\times3]} & \theta\mathbf{u}^\top \end{bmatrix} \quad \text{and} \quad \frac{\partial\mathbf{g}}{\partial t} = 0 \tag{7}$$

B. Control law

Usually, the control law is obtained by trying to make the global task $\mathbf{e}$ exponentially decrease in order to behave like a first order decoupled system. If the observed object is static (which is our case because the cross-wire phantom is motionless), this is achieved by applying the following control screw velocity to the probe [9]:

$$\mathbf{v} = -\lambda\mathbf{e} - (\mathbf{I}_n - \hat{\mathbf{L}}_{\mathbf{s}}^+ \hat{\mathbf{L}}_{\mathbf{s}})\frac{\partial\mathbf{g}^\top}{\partial t} \tag{8}$$

where:
- λ is the proportional coefficient involved in the exponential convergence of $\mathbf{e}$.
- $\hat{\mathbf{L}}_{\mathbf{s}}$ is an approximation of the interaction matrix. An on-line estimation of $\mathbf{L}_{\mathbf{s}}$ is presented in paragraph III-C.

In practice, we consider the input of the robot controller as the kinematic screw $\mathbf{v}_n$ of the end-effector. It is linked to the kinematic screw $\mathbf{v}$ of the US probe by:

$$\mathbf{v}_n = \begin{pmatrix} {}^n\mathbf{R}_s & [{}^n\mathbf{t}_s]_\times {}^n\mathbf{R}_s \\ \mathbf{0}_3 & {}^n\mathbf{R}_s \end{pmatrix} \mathbf{v} \tag{9}$$

where ${}^n\mathbf{t}_s$ and ${}^n\mathbf{R}_s$ are the translation vector and the rotation matrix from the end-effector frame to the probe frame. These two parameters with the image scaling factors S_x and S_y correspond to the spatial parameters of the ultrasound system. Since these parameters are not perfectly known before using the off-line Detmer calibration method, we set them to rough values. We will see in Section V that this will not affect the task performance due to the well-known robustness property of image-based visual servoing.

C. On-line estimation of the interaction matrix

The interaction matrix $\mathbf{L}_{\mathbf{s}}$ depends on the unitary vectors ${}^s\mathbf{u}_1$ and ${}^s\mathbf{u}_2$ of straight lines $\mathcal{D}_1$ and $\mathcal{D}_2$ and the 2D coordinates $\mathbf{p}_1$, $\mathbf{p}_2$. In practice $\mathbf{p}_1$, $\mathbf{p}_2$ are measured from their pixel coordinates but ${}^s\mathbf{u}_1$ and ${}^s\mathbf{u}_2$ are not known and we have to on-line estimate them. To do this, we use a recursive least-squares algorithm delivered below. For the sake of simplicity, in the sequel we give the method for one straight line $\mathcal{D}$. As the straight lines are motionless it is better to estimate them in a fixed frame such as the robot base frame wherein projection ${}^0\mathbf{u}$ of ${}^s\mathbf{u}$ is constant and the following minimal representations of $\mathcal{D}$ can be used:

$$x = az + c \quad \text{and} \quad y = bz + d \tag{10}$$

Here x, y, z are the coordinates expressed in the robot base frame of any point belonging to the straight line $\mathcal{D}$ and a, b, c, d are constant parameters. This minimal representation is always valid if $\mathcal{D}$ is not collinear to the plane described by X and Y axis of the robot base frame $\mathcal{F}_0$. By rewriting (10) in a vectorial system form we obtain:

$$\mathbf{Y} = (x, y) = \mathbf{\Phi}^\top \theta \quad \text{with} \quad \mathbf{\Phi}^\top = \begin{pmatrix} z & 0 & 1 & 0 \\ 0 & z & 0 & 1 \end{pmatrix} \tag{11}$$

where $\theta = (a, b, c, d)$ is the parameters vector to estimate. This system can be solved if we have at least the coordinates measurement of 2 points. In our approach we use a least-squares algorithm in order to take into account the coordinates of the point P measured at each iteration k. The method consists of computing the estimation value $\hat{\theta}_{[k]}$ that minimizes the following quadratic sum of the modeling error [14]:

$$J(\hat{\theta}_{[k]}) = \sum_{i=0}^{k} (\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\theta}_{[k]})^\top (\mathbf{Y}_{[i]} - \mathbf{\Phi}_{[i]}^\top \hat{\theta}_{[k]}) \tag{12}$$

Therefore $\hat{\theta}_{[k]}$ is obtained by nullifying the gradient of $J(\hat{\theta}_{[k]})$ and is given by the following recursive estimation:

$$\hat{\theta}_{[k]} = \hat{\theta}_{[k-1]} + \mathbf{F}_{[k]}\mathbf{\Phi}_{[k]}(\mathbf{Y}_{[k]} - \mathbf{\Phi}_{[k]}^\top \hat{\theta}_{[k-1]}) \tag{13}$$

where $\mathbf{F}_{[k]}$ is a covariance matrix such that $\mathbf{F}_{[k]} = \mathbf{F}_{[k]}^\top > 0$ and whose recursive expression is:

$$\mathbf{F}_{[k]}^{-1} = \mathbf{F}_{[k-1]}^{-1} + \mathbf{\Phi}_{[k]}\mathbf{\Phi}_{[k]}^\top \tag{14}$$

In practice we set initial values $\mathbf{F}_{[0]} = f_0\mathbf{I}_4$ with $f_0 > 0$ and $\hat{\theta}_{[0]} = \theta_0$. Once $\hat{\theta} = (\hat{a}, \hat{b}, \hat{c}, \hat{d})$ is computed, the estimated

unitary vector of $\mathcal{D}$ expressed in the $\mathcal{F}_0$ and in $\mathcal{F}_s$ is finally given by:

$$^0\hat{\mathbf{u}} = \frac{(\hat{a}, \hat{b}, 1)}{\|(\hat{a}, \hat{b}, 1)\|} \quad \text{and} \quad ^s\hat{\mathbf{u}} = {}^s\mathbf{R}_0 {}^0\hat{\mathbf{u}} \qquad (15)$$

where $^s\mathbf{R}_0$ is the rotation matrix between the probe frame and the robot base frame. Note that $^s\mathbf{R}_0 = {}^s\mathbf{R}_n {}^n\mathbf{R}_0$ is computed from the inverse of the rotation matrix $^n\mathbf{R}_s$ included in the spatial parameters. Therefore the estimation of the straight lines is sensitive to the accuracy of the spatial parameters. We finally obtain an estimation $\hat{\mathbf{L}}_\mathbf{s}$ of the interaction matrix by substituting in (3) the estimated unitary vectors $^s\hat{\mathbf{u}}_1$, $^s\hat{\mathbf{u}}_2$ and the current coordinates $\mathbf{p}_1$ and $\mathbf{p}_2$ measured in the US image. An adaptive visual servoing is then performed by updating $\hat{\mathbf{L}}_\mathbf{s}$ at each iteration of the control law (8).

IV. Image processing

This Section concerns the image processing used to extract image coordinates of point P_1 and P_2. The cross-wire phantom is realized with two nylon yarns (0.5 mm diameter) which are fixed in a box full of water such that they make two converging straight lines (see Fig. 3.a). From the captured image seen in Fig. 2.a we can see that Sections of nylon yarns are visible with high intensity level pixels. To extract the pixels of the two points, we apply a binary thresholding on the whole US image following by a binary closing. Then, two lists of pixels belonging to point P_1 and P_2 are obtained by detecting the pixels of connected components. The computation of the pixels center of mass for each list gives finally the image coordinates of P_1 and P_2. This coordinates extraction works very well as long as the two points are not joined together. If the points begin to join together, the detection of the connected components give only one list of pixels and the center of mass computed from this list is shifted from the real centers of points P_1 and P_2. To cope with this problem which may introduce discontinuity in the control law, we use a k-mean algorithm to separate the two image points. To do this, first we affect randomly each pixel to a region k with $k = 1$ or 2. Then the k-mean algorithm consists to reallocate each pixel to the region k such that the distance from this pixel to the center of mass of the region k is minimum. Once all the pixels are affected, the center of mass of each region is updated and the pixels affectation is repeated until the centers of mass remain stable. At the end we obtain two pixels regions like these shown in blue and white in Fig. 2.b. We assume hereafter that the nylon yarns sections are seen as small regular disks in the US image with identical radius. In this case, the k-mean algorithm gives the same area $A_1 = A_2 = A$ for the two truncated disks shown in Fig. 2.b as well as their center of mass coordinates M_1 and M_2. However M_1 and M_2 can not be used as the image coordinates of P_1 and P_2 because they are shifted from the real disks centers due to the truncation. Nevertheless, if the distance d_c (see Fig. 2.b) is known, the image coordinates of P_1 and P_2 can then be easily located on the straight line defined by M_1 and M_2. So we propose to compute the distance d_c from the area A and
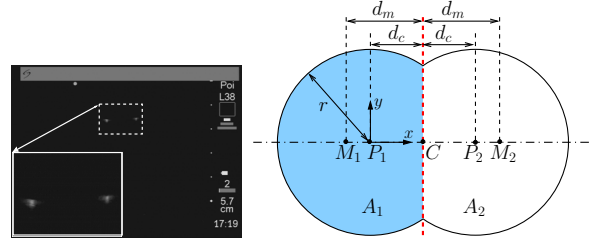

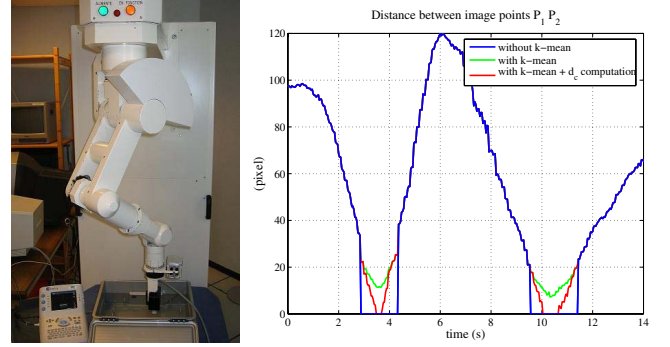Fig. 2. (a) Ultrasound image with image points zoomed - (b) Points fusion

Fig. 3. (a) Medical robot holding the ultrasound probe and phantom used for spatial calibration - (b) Measured image distance between points P_1 and P_2 when the probe is manually moved in order to overlay the two points

the distance d_m both provided by the k-mean algorithm. In our approach, by determining the analytical expression of the truncated disk area A and the expression of the x coordinate of M_1 which corresponds to $d_c - d_m$, we obtained the following two nonlinear equations system to solve r and d_c:

$$\begin{cases} f(r, d_c) &= d_c\sqrt{r^2 - d_c^2} + r^2 \arctan(d_c/\sqrt{r^2 - d_c^2}) \\ &\quad + \frac{r^2\pi}{2} - A = 0 \\ g(r, d_c) &= \frac{2(r^2 - d_c^2)^{3/2}}{3A_1} + d_c - d_m = 0 \end{cases}$$

Analytically it is very hard to solve it, so we use the popular numerical Newton-Raphson method to estimate r and d_c. Fig. 3.b displays the evolution of the distance between points P_1 and P_2 measured in the US image, when we manually move the probe in order to overlay the two points in the image. To compare the performance of the image coordinates extraction, we show the distance computed, firstly without considering fusion of the points, then by using the k-mean algorithm to separate the points when they join together (distance $= 2d_m$), and finally with the k-mean algorithm followed by the d_c computation (distance $= 2d_c$). We can easily see that the better method is the last one because continuity is improved when the points join together and offset is null once the two points are entirely layered.

V. Experimental results

A. Experimental setup

The ultrasound system is composed of a SonoSite 180PLUS (2D imaging system) connected to a broadband 10-5 MHz linear transducer (probe). We use a PC running Linux for real-time image processing and for controlling, via a LAN network,

a 6 DOF medical manipulator holding the US probe (see Fig. 3.a). This robot was specially designed for 3D ultrasound imaging and is similar to the Hippocrate robot [2]. The 2D imaging system is linked to a PCI capture board which grabs gray level 576×768 images at 25 fps.

B. Test of the robotic task for automatic spatial calibration

As we have mentioned in Section III, the rank of the interaction matrix switches from 4 to 3 when points P_1 and P_2 join together. Of course joining the two points is the goal of the visual task. So in order to avoid numerical instabilities, we propose to force the rank of the interaction matrix used in the control law to 3 when the two points join together (distance ≤ 5 pixels).

In a first experiment, we set rough values for the spatial calibration parameters used by the control law. The image scale factors are set to $S_x=S_y=$2e-4 (meter/pixel), the translation from the end-effector frame to the probe frame is $^n\mathbf{t}_s=(0.17, 0, 0)$ (meter) and the X-Y-Z Euler angles describing the rotation $^n\mathbf{R}_s$ are $\alpha=\beta=\gamma=0$ (deg). We set the reference values of the visual features in order to position the images of points P_1 and P_2 on a target point located in the image at $(144, 384)$ pixels. For the secondary task, the relative rotation between the current and desired orientation of the probe is set to the $\theta\mathbf{u}$ angle vector $(0, 5, 15)$ (deg). The gain of the control law is fixed to $\lambda = 0.2$ and the initial estimation of the straight lines parameters are set to wrong values $\hat{\theta}_{[0]}=(0.2, 0.2, 0, 0)$ for straight line $\mathcal{D}_1$, $\hat{\theta}_{[0]}=(-0.2, -0.2, 0, 0)$ for straight line $\mathcal{D}_2$. Fig. 4 displays the evolution of the visual error of P_1 and P_2 image coordinates, the $\theta\mathbf{u}$ angle error of the secondary task and the rank of the interaction matrix. We can see that the visual error overall decreases towards zero but that the visual task is lightly coupled with the secondary one (see at time $t=22$ s when the error briefly increases). This coupling is due to an estimation error on $\hat{\mathbf{L}}_\mathbf{s}$ used in the control law (8) and the inaccurate spatial parameters needed in equation (9). Indeed the components of the unitary vectors of the straight lines used to compute $\hat{\mathbf{L}}_\mathbf{s}$ are not well estimated in the robot base frame as it is shown in Fig. 5, where dashed lines correspond to ideal values measured on the phantom. This is also due to the inaccurate spatial parameters which are needed in the estimation algorithm (eq. (10)) to express in the robot base frame the 3D coordinates of a point measured in the probe frame. Note that when the rank of the interaction matrix $\hat{\mathbf{L}}_\mathbf{s}$ decreases to 3, the angle error decreases exponentially since the secondary task can use the 3 DOF necessary to orientate the probe whereas if the rank is 4 the secondary task can not be achieved. As the visual task is lightly coupled with the secondary one, we can also see that the rank of $\hat{\mathbf{L}}_\mathbf{s}$ switch from 3 to 4 when the points disjoint (distance > 5 pixels) and that the secondary task stops to converge due to the priority of the visual task. The image trajectories of the points are drawn in Fig. 6. We can note that at the start the points do not move towards the correct direction due to the wrong initial estimated values of $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, but estimation values are quickly readjusted by the on-line least-square algorithm and
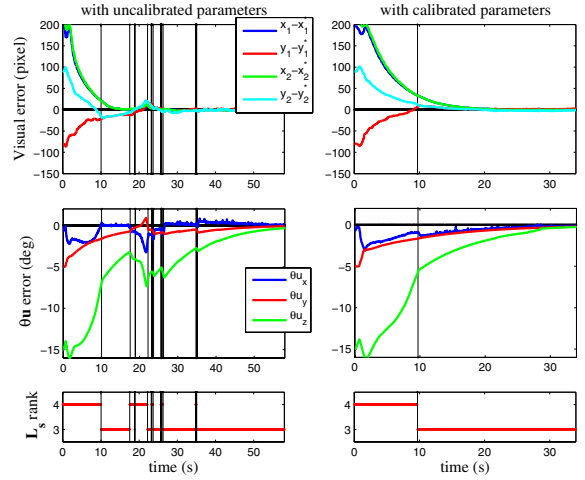

Fig. 4. Evolution of the visual error $\mathbf{s} - \mathbf{s}^*$ (top figures), $\theta\mathbf{u}$ angle error (middle figures) and the $\hat{\mathbf{L}}_\mathbf{s}$ rank (bottom figures) for uncalibrated (left figures) and calibrated (right figures) spatial parameters

the points converge towards the image target point $(144, 384)$ (pixels). Although the spatial parameters of the US system are not accurate, the robotic task is well achieved thanks to the robustness property of the image-based visual servoing.

We repeated this robotic task 20 times for different desired target points in the image and different desired orientations of the US probe. For each experiment, once the task was achieved, the relative position and orientation between the end-effector and the robot base frame as well as the pixel coordinates of the point $P^*=P_1=P_2$ were measured. This data was then given as input to the Detmer algorithm [3] referred in Section II in order to off-line estimate the spatial calibration parameters. This algorithm provides us the following parameters: S_x=1.22e-4, S_y=1.27e-4, $^n\mathbf{t}_s$=(0.206, -0.007, 0.001) (meter) and X-Y-Z Euler angles α=1.7, β=-2.5, γ=-3.7 (deg).

To verify the accuracy of these parameters, we perform the robotic task in the same conditions than the first experiment but using in the control law the estimated spatial parameters. We present in figures 4-6 the same measurements than for the first experiment. However, we can now see that the two points converge quasi exponentially towards the image target and that the angle error also decreases exponentially towards zero once the rank of $\hat{\mathbf{L}}_\mathbf{s}$ switches from 4 to 3. Now the secondary task is achieved without disturbing the visual task thanks to the accurate estimation of the spatial parameters. The estimated components of the unitary vectors converge also better towards the good values and the points trajectories become nearly straight lines in the image once the estimation of unitary vectors are readjusted. This well behaviour of the robotic task means that the spatial parameters are now more accurate than these used in the first experiment.

C. Validation for 3D ultrasound imaging

To validate the spatial calibration we perform the 3D ultrasound imaging of a straight nylon string by moving the probe with the robot. The 3D trajectory applied to the end effector
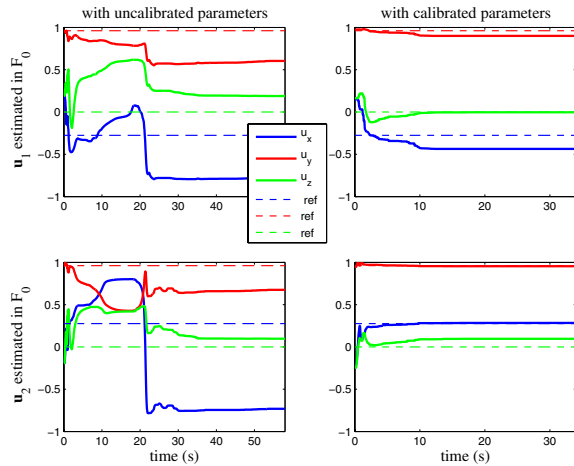
Fig. 5. Evolution of the estimated unitary vectors $^0\hat{\mathbf{u}}_1$ and $^0\hat{\mathbf{u}}_2$ expressed in the robot base frame for uncalibrated (left figures) and calibrated (right figures) spatial parameters. The dashed lines correspond to ideal values measured on the cross-wire phantom
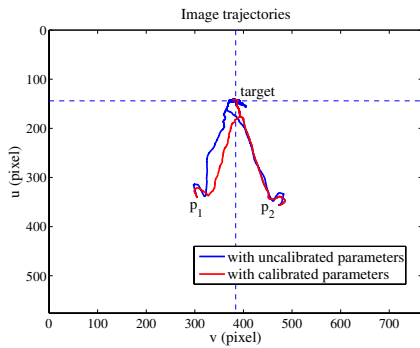

Fig. 6. Image trajectories of points P_1 and P_2 during the robotic task for uncalibrated (blue color) and calibrated (red color) spatial parameters
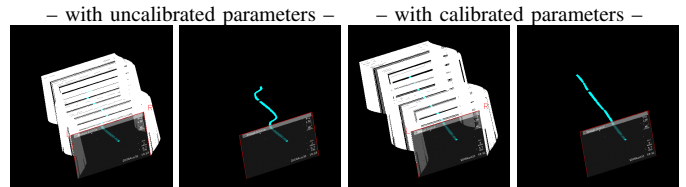

Fig. 7. 3D reconstruction of a straight nylon string (visualisation with Stradx software)

intersection point of a cross-wire phantom. A robust image processing method has also been developed to extract, from the ultrasound images, the coordinates of two points joining together. Experimental results show that the visual servoing is robust to large model errors and that the spatial calibration procedure is well performed.

of the robot is set in order to have a constant translational motion along the Z axis of the probe frame $\mathcal{F}_s$, sinusoidal translation motions on the 2 lateral axis of $\mathcal{F}_s$ and sinusoidal rotation motions around the 3 axis. A set of ultrasound images with their respective probe locations is captured when the robot tracks the trajectory and is used as input to the 3D visualisation Stradx software [15]. Fig.7 displays the 2D ultrasound images positioned in the robot base frame and the 3D shape of the string obtained when either inaccurate spatial parameters or calibrated parameters are used. Here again, we can see that our automatic calibration is well performed because the 3D shape resulted for the calibrated parameters is straight contrary to the shape obtained for uncalibrated parameters. In the future, we plan to compare accuracy of the obtained spatial parameters with other calibration methods like this described in [4].

VI. CONCLUSION

In order to automatically perform the spatial calibration procedure of a robotized 3D ultrasound imaging system, a new visual servoing technique based on 2D ultrasound images has been presented to position the ultrasound image on the

REFERENCES

[1] T. R. Nelson and D. H. Pretorius, "Three-dimensional ultrasound imaging," *Ultrasound in Medicine and Biology*, vol. 24, no. 9, pp. 1243–1270, 1998.

[2] 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.

[3] P. R. Detmer, G. Basheim, T. Hodges, K. W. Beach, E. P. Filer, D. H. Burns, and D. E. Strandness, "3d ultrasonic image feature localization based on magnetic scanhead tracking : in vitro calibration and validation," *Ultrasound in Medicine and Biology*, vol. 20, no. 9, pp. 923–936, 1994.

[4] F. Rousseau, P. Hellier, and C. Barillot, "Robust and automatic calibration method for 3d freehand ultrasound," in *Int. Conf. Medical Image Computing and Computer Assisted Intervention*, Montreal, Canada, November 2003, pp. 440–448.

[5] P. Abolmaesumi, S. E. Salcudean, W. H. Zhu, M. Sirouspour, and S. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Trans. Robotics and Automation*, vol. 18, pp. 11–23, February 2002.

[6] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata, "An ultrasound-driven needle insertion robot for percutaneous cholecystostomy," *Physics in Medicine and Biology*, vol. 49, no. 3, pp. 441–455, February 2004.

[7] M. A. Vitrani, G. Morel, and T. Ortmaier, "Automatic guidance of a surgical instrument with ultrasound based visual servoing," in *IEEE Int. Conf. on Robotics and Automation*, Barcelona, Spain, April 2005, pp. 510–515.

[8] R. Prager, R. Rohling, A. Gee, and L. Berman, "Automatic calibration for 3-d free-hand ultrasound," Cambridge University Engineering Department, Tech. Rep. CUED/F-INFRNG/TR 303, 1997.

[9] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robotics and Automation*, vol. 8, no. 3, pp. 313–326, June 1992.

[10] A. Krupa and F. Chaumette, "Control of an ultrasound probe by adaptive visual servoing," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Edmonton, Canada, August 2005, pp. 2007–2012.

[11] C. Samson, M. L. Borgne, and B. Espiau, "Robot control : The task function approach," *Clarendon Press, Oxford, United Kingdom*, 1991.

[12] E. Marchand and G.-D. Hager, "Dynamic sensor planning in visual servoing," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Lueven, Belgium, May 1998, pp. 1988–1993.

[13] G. D. Hager, "Human-machine cooperative manipulation with vision-based motion constraints," in *Workshop on visual servoing, IROS'02*, October 2002, pp. 1988–1993.

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

[15] R. Prager, A. Gee, and L. Berman, "Stradx: real-time acquisition and visualisation of freehand three-dimensional ultrasound," *Medical Image Analysis*, vol. 3, no. 2, pp. 129–140, 1999.