Intensity-based direct visual servoing of an ultrasound probe

C. Nadeau, A. Krupa

HAL Id: hal-00639689
https://inria.hal.science/hal-00639689

Submitted on 9 Nov 2011

Intensity-based direct visual servoing of an ultrasound probe

Caroline Nadeau and Alexandre Krupa

Abstract— **This paper presents a new image-based approach to the control of a robotic system equipped with an ultrasound imaging device. For diagnostic applications, the proposed method makes it possible to position an ultrasound probe on a desired organ section and to track it by compensating for rigid motions of the organ. Both in-plane and out-of-plane motions of the probe are controlled by the proposed method. The main contribution of this work is the direct use of the ultrasound image as visual feature which spares any segmentation or image processing time consuming step. Simulation and robotic experiments are performed on a realistic abdominal phantom and validate this ultrasound intensity-based visual servoing approach.**

Index Terms— **Visual servoing, ultrasound guided robotic system, intensity-based control**

I. INTRODUCTION

Among the different medical imaging modalities, ultrasound (US) imaging has many benefits for the patients as well as for the specialists. Indeed, this modality is cheap, real time and contrarily to MRI or CT devices, the US transducer is not cumbersome and can be easily used in an operating room. Moreover US waves are safe for human body and do not interact with ferromagnetic medical instruments. Therefore the use of US imaging during a medical intervention does not bring any additional constraint. Because of such advantages, US is a promising imaging modality to deal with image guided robotized systems. However the information carried by an US slice is far different from the one carried by a camera view which is traditionally used in visual control. In particular, the major remaining challenges in US visual servoing concern the image processing and the control of the out-of-plane motions of the US device.

Previous works dealing with US image-based robotic systems mainly focus on two different system configurations. In the eye-to-hand configuration, the US probe is observing a surgical instrument mounted on the robot end-effector. The robotic manipulation offers a better accuracy than the human one and the proposed applications concern needle insertion procedures [1] or cardiac surgery [2], [3]. In [1], two degrees of freedom (dof) of a needle-insertion robot are controlled by visual servoing to perform a percutaneous cholecystostomy while compensating involuntary patient motions. The target and the needle are automatically segmented in the US images and their respective poses are used to guide the robot. In [2], a robotic system is proposed to track a surgical instrument

C. Nadeau is with Université de Rennes I, IRISA, INRIA Rennes-Bretagne Atlantique, Lagadic research group, 35042 Rennes, France. A. Krupa is with INRIA Rennes-Bretagne Atlantique, IRISA, Lagadic research group, 35042 Rennes, France `Caroline.Nadeau@irisa.fr`, `Alexandre.Krupa@inria.fr`

and move it to a desired target. 3D US images are processed to localize here again the respective positions of the target and the instrument tip, then the position error is used to control the surgical robot. In [3], the four dof of a surgical forceps inserted in a beating heart through a trocar are controlled by US image-based visual servoing. In relation with this work, the authors of [4] developed a predictive control scheme to keep the forceps visible in the US image.

The other configuration, namely eye-in-hand configuration, allows the direct control of the US sensor mounted on the robot end-effector for diagnostic purpose [5], [7] or medical procedure [6]. In [5], the three in-plane dof of a robotic system are controlled to maintain a visual feature centered in the US image during a manual out-of-plane translation of the US probe. In [6], two US probes and a HIFU transducer are mounted on the end effector of a XYZ stage robot to follow a target kidney stone while compensating physiological motions. For a positioning task, [7] proposed a method to automatically reach a desired cross section of an organ of interest by servoing the six dof of a 2D US probe.

In order to control one to six dof of the robotic manipulator, the efficiency of the visual servoing approaches is highly dependent on the choice of appropriate image features. Depending on the configuration, these features can be created by the intersection of the surgical tool with the US plane (eye-to-hand configuration) [1]–[4] or by anatomical landmarks (eye-in-hand configuration) [5]–[7].

In robotic systems where the US probe itself is controlled, which are more particularly within the scope of this work, the image features can only be anatomic ones. In [5], five features extraction methods are compared to track an anatomic point which is the center of an artery in order to servo the in-plane motions of the probe. These methods are based on image similarity measures such as cross correlation and Sequential Similarity Detection (SSD) or on contour segmentation by a Star [8] or Snake algorithm. In another work [6], the translational motions of a robotic effector are controlled using the center position of a segmented renal stone. Recently, a few authors provided solutions to control the six dof of the probe. In [9], an approach based on the speckle correlation observed in successive US images is detailed. However only tracking tasks can be considered with such approach. Then, different approaches have been proposed to perform positioning tasks using six geometric features built from 2D moments extracted from a single US image [7] or three orthogonal images [10]. However the moment computation requires a previous contour segmentation step whose efficiency is dependent on the organ shape and which is not robust to organ topology changes.

In this paper we propose to avoid any image processing step by using directly the US image intensity in our visual servoing approach. In this case, the visual features are the set of image pixel intensities. The contribution of this paper is therefore to provide an efficient model of the interaction between the variation of these features and the velocity of an actuated 2D US probe to control the six dof of this probe.

The structure of our paper is as follows. We initially describe our US intensity-based approach and detail the computation of the interaction matrix to control in-plane but also out-of-plane motions of the probe in section II. In section III, we present results of the proposed approach for positioning and tracking tasks performed in a simulation environment. Finally, in section IV, a robotic tracking experiment involving a hybrid force/vision control demonstrates the validity of the approach in real environment.

II. ULTRASOUND VISUAL SERVOING

Traditional visual servoing methods refer to vision data acquired with a camera mounted on a robotic system. In this case, the vision sensor provides a projection of the 3D world into a 2D image and the coordinates of a set of 2D geometric primitives can be used to control the six dof of the system. However, a 2D US transducer provides complete information in its image plane but not any outside of this plane. Therefore, the interaction matrix relating the variation of the chosen visual features to the probe motion is far different and has to be modeled.

A. The control law

An image-based visual servoing control scheme consists in minimizing the error $\mathbf{e}(t) = \mathbf{s}(t) - \mathbf{s}^*$ between a current set of visual features $\mathbf{s}$ and a desired one $\mathbf{s}^*$. Considering an exponential decrease of this error, the classical control law [12] is given by:

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L_s}}^+ (\mathbf{s}(t) - \mathbf{s}^*), \tag{1}$$

where λ is the proportional gain involved in the exponential decrease of the error ($\dot{\mathbf{e}} = -\lambda \mathbf{e}$). In an eye-in-hand configuration, $\mathbf{v}_c$ is the instantaneous velocity applied to the visual sensor and $\widehat{\mathbf{L_s}}^+$ is the pseudo-inverse of an estimation of the interaction matrix $\mathbf{L_s}$ that relates the variation of the visual features to the velocity $\mathbf{v}_c$.

According to [12], the control scheme (1) is known to be locally asymptotically stable when a correct estimation $\widehat{\mathbf{L_s}}$ of $\mathbf{L_s}$ is used (i.e., as soon as $\mathbf{L_s}\widehat{\mathbf{L_s}}^{-1} > 0$).

B. Selection of intensity features

In this paper, we propose to avoid any US image preprocessing step by choosing as visual feature the image itself. In this case, the visual features considered are the intensity value of each pixel contained in the US image. The size of the image features vector $\mathbf{s}$ is therefore equal to the size of the US image $\mathbf{I}(\mathbf{r})$ acquired at the pose $\mathbf{r} \in SE_3$ of the imaging device.

$$\mathbf{s}(\mathbf{r}) = \{I_r(u,v), \ \forall (u,v) \in [1,M] \times [1,N]\} \tag{2}$$

where M and N are respectively the width and the height of the US image and where $I_r(u,v)$ represents the intensity of the pixel of coordinates (u,v) in the image $\mathbf{I}(\mathbf{r})$.

C. Computation of the interaction matrix

In a B-mode US scan, the intensity of the US signal is represented in term of pixel luminance intensity. The higher this luminance, the higher the US wave reflexion. Then as the US reflexion only depends on the organ structure and interfaces, its value remains roughly constant for a given anatomic micro structure. As a result, we will consider that the luminance intensity in a B-mode US image of a physical 3D point remains also constant during a time interval dt.

As stated in a previous work dealing with photometry based approach for camera images [11], the hypothesis of intensity conservation allows us to link the time variation of the features vector $\mathbf{s}(\mathbf{r})$ to the motion of the imaging device. As the considered visual features belong to the US image plane, their 3D coordinates $\mathbf{x}$ expressed in the probe frame can be computed from the image pixel coordinates:

$$\mathbf{x} = (x,y,z)^\top = (s_x(u - u_0), s_y(v - v_0), 0)^\top,$$

with (s_x, s_y) the image pixel size and (u_0, v_0) the pixel coordinates of the image center. Given such a 3D point expressed in the imaging device frame and $\mathbf{dx} = (dx, dy, dz)^\top$ an elementary motion of this point in the 3D space, the intensity conservation is expressed by the relationship:

$$I_r(x + dx, y + dy, z + dz, t + dt) - I_r(\mathbf{x}, t) = 0. \tag{3}$$

A first order Taylor series expansion of (3) yields to:

$$\frac{\partial I_r}{\partial x}dx + \frac{\partial I_r}{\partial y}dy + \frac{\partial I_r}{\partial z}dz + \frac{\partial I_r}{\partial t}dt = 0. \tag{4}$$

Then, the time variation of each image feature $I_r(u,v)$ can be expressed as a function of the 3D point motion :

$$\dot{I}_r = -(\nabla I_x \ \nabla I_y \ \nabla I_z) \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix}, \tag{5}$$

with $\nabla I = (\nabla I_x \ \nabla I_y \ \nabla I_z)^\top$ the 3D image gradient.

The interaction matrix L_I of size 1×6 associated to each visual feature is then:

$$L_I = -(\nabla I_x \ \nabla I_y \ \nabla I_z) L_{\mathbf{x}}, \tag{6}$$

where $L_{\mathbf{x}}$ relates the velocity of a 3D point $\mathbf{x}$ to the motion of the US probe $\mathbf{v}_c$ according to the kinematics fundamental relationship:

$$\dot{\mathbf{x}} = L_{\mathbf{x}}\mathbf{v}_c, \quad L_{\mathbf{x}} = \begin{bmatrix} -1 & 0 & 0 & 0 & -z & y \\ 0 & -1 & 0 & z & 0 & -x \\ 0 & 0 & -1 & -y & x & 0 \end{bmatrix}. \tag{7}$$

The final interaction matrix $\mathbf{L_s}$ used in the control law is built by stacking all the previously defined matrices L_I associated to each pixel of the considered sub image $\mathbf{I}(\mathbf{r})$.

D. Image gradient computation

To control the six dof of the US probe, the variation of the visual features is related to both in-plane and out-of-plane motions of the probe. In our US intensity-based approach, this variation is given by the 3D image gradient, which is computed from the current probe image and at least two parallel additional images.

With a conventional 2D US probe mounted on a robotic arm, a small back and forth translational motion along the elevation direction can be applied before each iteration of the visual control algorithm in order to acquire these additional images I_a and I_b. We design then a set of three 3D Gaussian derivatives filters of size $3 \times 3 \times 3$ applied in each pixel of the current image I_0 to compute the 3D gradient (see Fig. 1).
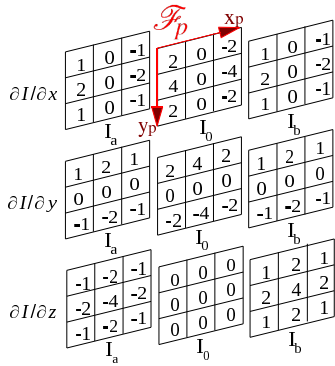

Fig. 1. The three cubic filters are applied in each pixel of the image I_0 to compute the gradient components $(\nabla I_x, \nabla I_y, \nabla I_z)$ using the additional parallel images I_a and I_b acquired on both sides of I_0.

III. SIMULATION VALIDATION

To validate our approach, we use a software simulator that we have developed to reconstruct and display a dense volume from a set of parallel images. In addition to this display functionality, the simulator allows us to move the reconstructed volume with respect to a fixed cartesian frame and to control a 2D virtual probe which generates an US image by cubic interpolation process.

For the simulation experiments, an US complete volume of the right kidney of a realistic abdominal phantom is loaded in the simulator (see Fig. 2). This volume is created from a set of 335 250×250 parallel images, the voxel size is $0.6 \times 0.6 \times 0.3 mm^3$.

A. Positioning task

We first simulate a positioning task, using the simulation environment to obtain a ground truth of the evolution of the pose error of the US probe. We position the probe on the kidney volume and we consider the corresponding US scan as the desired one. In the same time, the corresponding pose of the probe in the simulator frame $\mathbf{r}^*$ is saved. Then the probe is moved away to a new pose where the observed organ section is considered as the initial image.

To avoid continuous out-of-plane motions of the 2D US probe required to compute the 3D image gradient during the
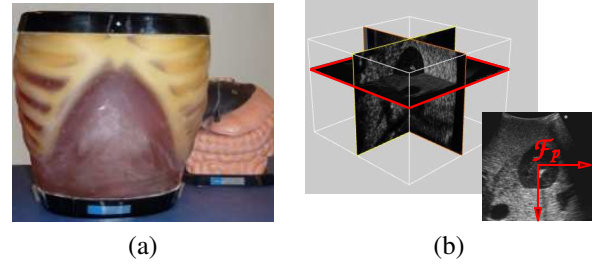

Fig. 2. (a) The US abdominal phantom (Kyoto Kagaku). (b) The volume loaded in the simulator is represented by two orthogonal slices and the virtual probe plane, defined with the frame $\mathscr{F}_p$, is displayed in red.

visual servoing process, we propose to use the interaction matrix estimated at the desired pose $\mathbf{L_s}^*$ in the control law. However, with such an approximation of the interaction matrix, the convergence of the control law (1) is not guaranteed from a far initialization. We propose then to solve the minimization problem with the Levenberg-Marquardt algorithm which is a combination of the Gauss-Newton algorithm with the steepest descent method and ensures a better convergence than the Gauss-Newton one (1) from a far initialization. The implemented control law for the positioning task is then:

$$\mathbf{v}_c = -\lambda \left(\mathbf{H} + \mu \, diag(\mathbf{H}) \right)^{-1} \mathbf{L_s}^{*\top} \left(\mathbf{s}(t) - \mathbf{s}^* \right),$$

with $\mathbf{H} = \mathbf{L_s}^{*\top} \mathbf{L_s}^*$ and $\mu = 0.05$. The 3D gradient components are computed with $5 \times 5 \times 5$ filters by using four additional images. These filters, built on the same Gaussian derivative model than the $3 \times 3 \times 3$ filters described in Fig. 1, increase the robustness of the control scheme.
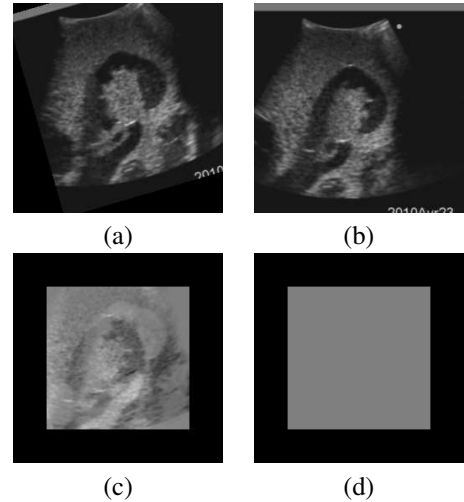

Fig. 3. Automatic probe positioning on a desired cross section of the right kidney. Full 250×250 initial (a) and desired (b) images. (c) and (d) are respectively the initial and final image differences $\mathbf{I}(\mathbf{r}^*) - \mathbf{I}(\mathbf{r})$ where $\mathbf{I}(\mathbf{r})$ is the 170×170 sub image considered as visual feature.

The Fig. 3 shows the visual convergence of the algorithm. The images difference between the current and the desired US scans are displayed for the initial and final probe poses. The uniform gray color of this images difference observed after the convergence of the algorithm proves the success of

the positioning task since the final image perfectly coincides with the desired one.
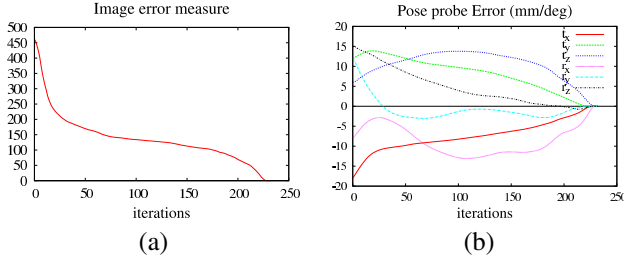

Fig. 4.   Evolution of the visual (a) and pose error (b) in mm and deg, (the $\theta\mathbf{u}$ representation is used to describe the orientation).

The behavior of the automatic positioning is shown through the evolution of the probe pose error (see Fig. 4). We also define an image error measure $\mathscr{C}(\mathbf{r}) = (\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*))^\top (\mathbf{s}(\mathbf{r}) - \mathbf{s}(\mathbf{r}^*))$ to observe the minimization of the visual features. Without any image processing step during the process, the iteration loop is performed in only $10ms$ on a PC equipped with a 3GHz Dual Core Xeon Intel processor. From the initial pose error $\Delta\mathbf{r}_{init} = (-18mm, 12mm, 6mm, -8°, 12°, 15°)$, the desired pose is then reached in $3s$. After the algorithm convergence, the final pose error is on the order of 10^{-6} in meter and degree.

B. Tracking task

For a tracking task, we set the desired image to be the initial observed one. In such a configuration, the interaction matrix can therefore be pre computed at the initial pose without being updated during the servoing task. As previously, the desired interaction matrix $\mathbf{L_s}^*$ is then used in the control law. However, this approximation is well justified in a tracking application and ensures this time a good behavior of the Gauss-Newton control law (1).
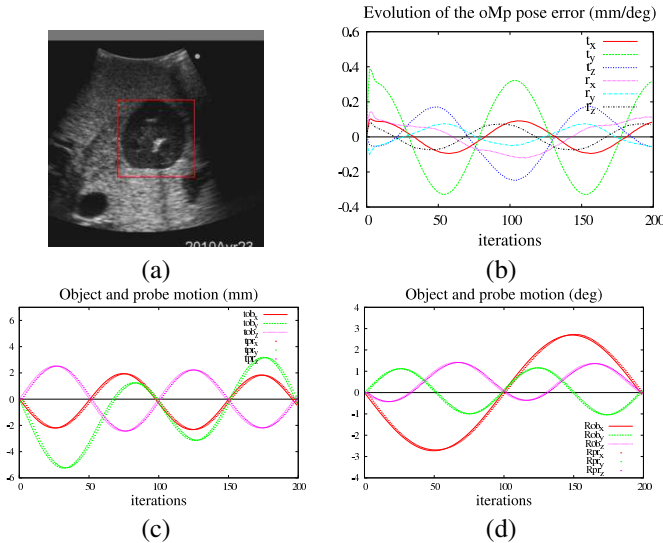

Fig. 5.   Results of the tracking of a kidney cross-section (a) while sinusoidal motions are applied to the US volume. (b) Relative pose of the probe in the object frame. Evolution of the translational (c) and rotational (d) components of the object (line) and probe (dots) poses expressed in the simulator world frame.

The results of the tracking task are shown in Fig. 5. For validation purpose, we define in the simulation environment the object frame as superimposed with the probe frame at the first iteration of the algorithm. The relative pose error oMp between both frames, displayed during the tracking task, remains lower than $0.4mm$ in translation and $0.15°$ in rotation and shows the efficiency of the visual control.

IV. ROBOTIC EXPERIMENT

In a medical robotic application, safety issues imply to combine the visual control with a force control insofar as the US probe relies on the patient skin. We propose then to add to the visual task a force constraint which guarantees a constant force applied on the patient.

A. Force and image control combination

Two sensors are now used to control the system. We implement a hybrid vision/force control based on an external control loop approach. The force control is used to servo the translational motion along the y-axis of the probe frame while the five remaining dof are controlled by the visual servoing control scheme.
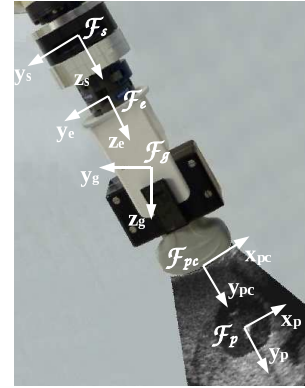

Fig. 6.   The robot end effector (frame $\mathscr{F}_e$) is equipped with a force sensor (frame $\mathscr{F}_s$) and an US probe (frame $\mathscr{F}_p$).

1) Force control: We implement a torque/force control law to guarantee a constant resulting force of $1N$ applied on the contact point pc of the probe with the object surface along the y-axis of the probe frame. $^{pc}H_{pc}$ corresponds to the contact force tensor expressed in the frame $\mathscr{F}_{pc}$, which is centered on the contact point and aligned with the probe frame $\mathscr{F}_p$ (see Fig. 6). It is given by the following relationship:

$$^{pc}H_{pc} = {}^{pc}\mathbf{F}_s \left({}^sH_s - {}^s\mathbf{F}_g \, {}^gH_g \right), \qquad (8)$$

$^a\mathbf{F}_b$ is the transformation matrix used to express in the frame $\mathscr{F}_a$ a force tensor known in the frame $\mathscr{F}_b$:

$$^a\mathbf{F}_b = \begin{pmatrix} {}^a\mathbf{R}_b & \mathbf{0}_{3\times 3} \\ [^a\mathbf{t}_b]_\times \, {}^a\mathbf{R}_b & {}^a\mathbf{R}_b \end{pmatrix}, \qquad (9)$$

where $^a\mathbf{t}_b$ and $^a\mathbf{R}_b$ are the translation vector and the rotation matrix of the frame $\mathscr{F}_b$ with respect to the frame $\mathscr{F}_a$ and $[^a\mathbf{t}_b]_\times$ is the skew symetric matrix related to $^a\mathbf{t}_b$.

^sH_s is the total force tensor measured by the force sensor and $^s\mathbf{F}_g\,^gH_g$ is the gravity force applied on the force sensor due to the mass m_p of the US probe, both are expressed in the force sensor frame. The ^gH_g tensor is defined as $^gH_g = [0\ 0\ 9.81m_p\ 0\ 0\ 0]^\top$ in the frame $\mathscr{F}_g$ centered on the mass center of the probe as indicated in Fig. 6.

We use a 6×6 selection matrix $Ms = diag(0, 1, 0, 0, 0, 0)$ to apply the force control only along the y-axis of the probe. We express then the resulting force tensor in the force sensor frame $\mathscr{F}_s$ and we compute the instantaneous velocity of the sensor $\mathbf{v}_s$ from the following proportional force control law:

$$\mathbf{v}_s = -\frac{K\,^s\mathbf{F}_{pc}\,(Ms\,^{pc}H_{pc} - {}^{pc}H_{pc}{}^*)}{k}, \tag{10}$$

where $^{pc}H_{pc}{}^* = [0\ 1\ 0\ 0\ 0\ 0]^\top$ is the desired contact force, k is the contact stiffness and K is the control gain.

2) Vision control: As we chose for safety reasons to give priority to the force control over the vision control, the latter can fail to converge to the desired image since the y-translational velocity component due to the image control is not applied to the probe. To deal with this issue, we propose to separate the dof controlled by the force control from the others. The five velocity components corresponding to the translations along the x and z axis of the probe and to the three rotations are applied physically to the US device while the last component corresponding to the y translation is virtually applied to a window containing the region of interest (ROI) and included in the US image (see Fig. 7).
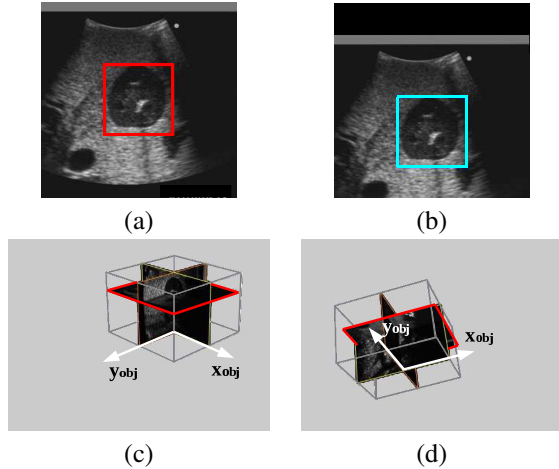

Fig. 7. Tracking application with five dof physically actuated. (a) The US slice to track with in red the desired ROI. (b) The final view of the US probe including the current ROI in cyan. (c) and (d) show the displacement between the initial and final poses of both object and US plane.

The velocity applied to the probe due to the image control is then such as $\mathbf{v}_p = \mathbf{v}_c$ given by (1), except for the component corresponding to the translation along the y-axis of the probe which is set to zero.

3) Sensor fusion: To combine the force and the vision control, we send the following angular velocity $\dot{q}$ to the end effector of the robotic arm:

$$\dot{q} = {}_e\mathbf{J}_e^+\mathbf{v}_e = {}_e\mathbf{J}_e^+({}^e\mathbf{V}_s\,\mathbf{v}_s + {}^e\mathbf{V}_p\,\mathbf{v}_p), \tag{11}$$

where $_e\mathbf{J}_e^+$ is the pseudo inverse of the robot Jacobian. Both image and force sensors being rigidly attached to the robot end effector, the control velocity of the effector $\mathbf{v}_e$ can be expressed in function of the control velocity of each device $\mathbf{v}_s$ and $\mathbf{v}_p$ through the transformation matrices $^e\mathbf{V}_s$ and $^e\mathbf{V}_p$, which are similarly defined in the following way:

$$^e\mathbf{V}_p = \begin{pmatrix} ^e\mathbf{R}_p & [^e\mathbf{t}_p]_\times\,^e\mathbf{R}_p \\ \mathbf{0}_{3\times3} & ^e\mathbf{R}_p \end{pmatrix}, \tag{12}$$

where $^e\mathbf{t}_p$ and $^e\mathbf{R}_p$ are the translation vector and the rotation matrix of the probe frame $\mathscr{F}_p$ expressed in the coordinates system of the end effector $\mathscr{F}_e$.

In addition, the v_{c_y} component computed with the vision control is applied to the considered window of interest to readapt its position in the US image.

B. Results

Experiments are performed with an anthropomorphic robotic arm equipped with a convex 2-5 MHz US transducer and a force sensor (see Fig. 8(a)). The same realistic US abdominal phantom already used for the simulation validation represents the patient body.
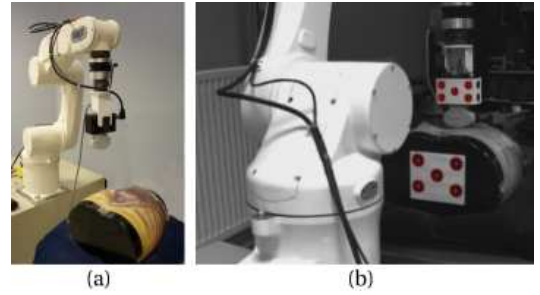

Fig. 8. (a) The ADEPT Viper robotic system. (b) View of the external camera used to compute the probe and object respective poses.

The considered application is a tracking task which allows us to compute the interaction matrix only at the initial pose of the probe. The image 3D gradient is computed with the $5 \times 5 \times 5$ filters and a Kalman filter is implemented to predict the phantom motion and increase the reactivity of the control. The Kalman filter is based on a constant velocity model and takes as input the measures of the image features variation and the probe instantaneous velocity. The estimated object velocity $\hat{\mathbf{v}}_\mathbf{o}$ is finally reinjected into the control law (1) as a prediction term:

$$\mathbf{v}_c = -\lambda\,\mathbf{L_s}^{*\top}\,(\mathbf{s}(t) - \mathbf{s}^*) + \hat{\mathbf{v}}_\mathbf{o}.$$

To validate the efficiency of the task, the relative pose of the probe with respect to the object is used as ground truth data. Both poses of the object and the probe are estimated by virtual visual servoing [13] thanks to a well calibrated camera observing the experiment scene and visual markers fixed on the probe and the phantom (see Fig. 8(b)). The 3D poses of the probe $^{cam}M_{probe}$ and the US phantom $^{cam}M_{ph}$ expressed in the camera frame are then used to compute their relative pose:

$$^{ph}M_{probe} = {}^{cam}M_{ph}^{-1}\,{}^{cam}M_{probe}$$

We position the 2D US probe on the abdominal phantom and we use a ROI of the observed US B-scan as the desired image to consider only relevant anatomic data in the visual control (see Fig. 9). The force/vision servo process is launched after the small automatic back and forth out-of-plane translation used to estimate the 3D image gradient. Then we manually apply various translational and rotational motions to the phantom.
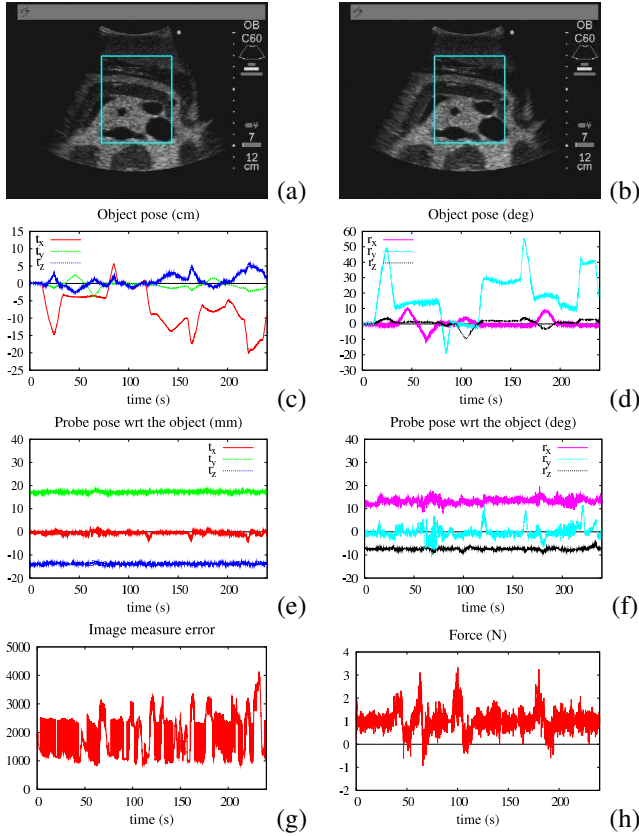

Fig. 9. Tracking application with the robotic arm. The initial (a) and final (b) images are displayed with in cyan the ROI. Although large translational (c) and rotational motions (d) are applied to the phantom, the respective pose of the probe in the object frame (e), (f) and the image measure error (g) remain roughly constant. The force applied to the probe (h) is maintained around 1N during the tracking process as expected from the force control.

The Fig. 9 shows the results of one tracking experiment. When important changes are applied on the object motion, the error of the probe pose with respect to the phantom and the image measure error both increase because of the tracking delay. However the image based algorithm is robust enough and the probe converges to the desired pose. Then, at the end of the tracking task, initial and final respective poses of the probe with respect to the phantom are compared and their difference is computed: $\Delta^{ph}\mathbf{r}_{probe} = (-0.28mm, 0.03mm, 0.24mm, -0.13°, 1.74°, -0.17°)$. Compared to the maximum motion amplitude applied to the phantom along each direction in translation and rotation, this error is on the order of 0.1% in translation and 1% in rotation, which demonstrates the success of the task. More tracking experiments results with visual and pose validations are presented in the attached video.

V. CONCLUSION

This paper presents a new approach of US image based robotic control. In order to avoid any segmentation process and to be robust to changes in the organ topology, the proposed control directly uses the B-scan image provided by the US probe as visual feature. The interaction matrix enabling the control of the six dof of the system from the image intensity is computed from the 3D image gradient of the US scan. As the estimation of this 3D gradient requires additional parallel images, we focus on tracking and local positioning tasks where the interaction matrix can be pre computed at the desired pose and used in the algorithm without being updated. However, in further works, positioning tasks involving the current interaction matrix can be considered using a 3D US probe in order to take into account higher initial pose errors. The challenge remains also in considering physiological non rigid motions.

REFERENCES

[1] J. Hong, T. Dohi, M. Hashizume, K. Konishi, N. Hata, A motion adaptable needle placement instrument based on tumor specific ultrasonic image segmentation. *In 5th Int. Conf. on Medical Image Computing and Computer Assisted Intervention, MICCAI'02*, pp122-129, Tokyo, Japan, September 2002.

[2] 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 IEEE Int. Conf. on Robotics and Automation, ICRA'07*, pp2655-2660, Roma, Italy, April 2007.

[3] 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, ICRA'07*, pp3021-3027, Roma, Italy, April 2007.

[4] M. Sauvee, P. Poignet, E. Dombre, US image based visual servoing of a surgical instrument through non-linear model predictive control, *Int. Journal of Robotics Research*, vol. 27, no. 1, January 2008.

[5] P. Abolmaesumi, S. Salcudean, W. Zhu, M. Sirouspour, and S. DiMaio, Image-guided control of a robot for medical ultrasound. *In IEEE Trans. on Robotics*, vol. 18, no. 1, February 2002.

[6] D. Lee, N. Koizumi, K. Ota, S. Yoshizawa, A. Ito, Y. Kaneko, Y. Matsumoto, and M. Mitsuishi, Ultrasound-based visual servoing system for lithotripsy. *In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'07*, pp.877 -882, 2007.

[7] R.Mebarki, A. Krupa and F. Chaumette, 2D ultrasound probe complete guidance by visual servoing using image moments. *In IEEE Trans. on Robotics*, vol. 26, nr. 2, pp 296-306; 2010.

[8] N. Friedland and D. Adam, Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing. *In IEEE Trans. Med. Imag.*, vol. 8, no. 4, pp. 344 - 353, 1989.

[9] A. Krupa, G. Fichtinger, G. Hager, Real time motion stabilization with B-mode ultrasound using image speckle information and visual servoing.*The International Journal of Robotics Research, IJRR*, 2009.

[10] C. Nadeau, A. Krupa, A multi-plane approach for ultrasound visual servoing: application to a registration task.*In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'10*, Taipei, Taiwan, 2010.

[11] C. Collewet, E. Marchand, F. Chaumette. Visual servoing set free from image processing.*In IEEE Int. Conf. on Robotics and Automation, ICRA'08*, pp. 81-86, Pasadena, CA, May 2008.

[12] B. Espiau, F. Chaumette and P. Rives, A new approach to visual servoing in robotics. *In IEEE Trans. on Robotics*, vol 8(3), pp.313-326, 1992.

[13] E. Marchand, F. Chaumette, Virtual visual servoing: A framework for real-time augmented reality. *In EUROGRAPHICS 2002 Conference Proceeding*, vol. 21(3), pp. 289-298, Sarrebruck, Germany, 2002.