Towards Ultrasound-based Visual Servoing using Shearlet Coefficients.

Lesley Duflot, Alexandre Krupa, Brahim Tamadazte, Nicolas Andreff

Towards Ultrasound-based Visual Servoing using Shearlet Coefficients

Lesley-Ann Duflot, Alexandre Krupa, Brahim Tamadazte, and Nicolas Andreff

***Abstract*— This paper deals with the development of a new 6 degrees-of-freedom (DOF) vision-based controller for robot-assisted medical applications. The main objective is to use visual information extracted from Ultrasounds (US) images to control an US probe held by a robot. Instead of the conventional use of visual features e.g., geometric features (points, lines, moments, etc.) in the visual control law design, the described method uses the shearlet coefficients. More precisely, the time-variation of the coarsest level of shearlet decomposition coefficients are linked to the US probe (respectively to the robot) spatial velocity and then related to the task-function control law. The proposed control law was experimentally tested and validated using a realistic abdominal phantom. The obtained results demonstrated promising performances in terms of accuracy, repeatability, robustness and convergence behavior.**

I. INTRODUCTION

2D and 3D US imaging devices, also known as echography systems, are widely used for non-invasive observations of inner organs i.e., diagnosis. Also, real-time US images can be used as visual feedback for guiding non/mini-invasive surgical procedures. Unfortunately, echography examination implies high demands on the surgeon since hand-eye-coordination becomes very mastery. Therefore, several robotic systems were developed to hold the US probe allowing thus remote examination of a patient. These systems are mainly based on master-slave robotic architectures allowing the clinician to tele-operate the probe motions [1]. Since few years ago, different solutions based on the *visual servoing* framework [2] were proposed to control the robot probe holder directly from information extracted from the US images. For example, in [3] several autonomous control modes based on US image-based visual servoing were developed to provide different kinds of assistance during robotic US examination. The objectives of visual servoing during an echography examination can be summarized by: 1) maintaining visibility of an organ [4], 2) automatic positioning of the probe in such a way to retrieve a pre-operative image [5] and 3) automatic compensation of physiologic patient motion in such a way to actively stabilize the US image on a moving organ [9].

The efficiency of a visual servoing approach is strongly depended on the choice of the appropriate visual features. These features can be points [4], lines [6], moments [5], etc. The success of these approaches depends highly on the ability to detect, match and track, in the US images, the visual features over time. In the literature, these techniques are referred to *visual tracking*. To overcome these limitations, original visual servoing approaches have been introduced which show that the design of an image-based visual controller can totally remove the visual tracking process [7], [8]. Recently, in [9], the authors have proposed an alternative 6 DOF visual servoing scheme by taking as visual inputs directly the intensity of the US image pixels. This method has proven its reliability in terms of convergence behavior robustness and accuracy thanks to the redundancy of the information used as inputs to the controller.

This paper addresses another kind of featureless visual servoing undressed so far: the use of *shearlet coefficients* for control. Shearlet is a recent mathematics tool for signal (respectively image) representation introduced in 2005 [10] which can be considered as a natural extension of wavelet [11]. It was proven that shearlet are more appropriate to represent anisotropic image features such as edges comparing to the wavelet. From this, several applications using shearlet representation have emerged: image inpainting [12], denoising and speckle filtering on US images [13].

The main scientific contribution of our work is the design of a visual servoing approach that uses shearlet coefficients instead of pure photometric information to control the 6 DOF of a robotized US probe. One of the advantages using shearlet coefficients in the design of the control law is to obtain a set of *noiseless* and redundant visual signal for a more robust and accurate visual servoing. The latter was experimentally tested and validated using a 6 DOF US robotic system holder and a realist abdominal phantom. The validation tests were conducted under different conditions i.e., nominal and unfavorable. The obtained results were more than satisfactory especially in terms of convergence (smooth domain), robustness (robust to the unfavorable conditions), repeatability and accuracy (average error of **0.19mm** and **0.23**$^\circ$ in translation and rotation axes, respectively).

In the remainder of this paper, Section II reviews the general basics of the shearlet representation of an image. The derived visual servoing controller is detailed in Section III while its experimental validation using a 6 DOF US robotic system holder and an abdominal phantom is discussed in Section IV.

II. SHEARLET TRANSFORM

The shearlet image representation used to design the proposed US-based visual controller constitutes an important

[1]L. A. Duflot and A. Krupa are with the Inria Rennes-Bretagne Atlantique, Campus universitaire de Beaulieu, Rennes 35042, France lesley-ann.duflot@inria.fr - alexandre.krupa@inria.fr

[2]L. A. Duflot, B. Tamadazte and N. Andreff are with FEMTO-ST, AS2M, Univ. Bourgogne Franche-Comté, Univ. de Franche-Comté/CNRS/ENSMM, 25000 Besançon, France brahim.tamadazte@femto-st.fr - nicolas.andreff@femto-st.fr

mathematical tool that may allow development of promising new visual servoing approaches. Thereby, for a better understanding of the controller design methodology, it is obvious to introduce some basics of the shearlet transform, more precisely the *cone-based shearlet system*. As mentioned in [14], the shearlet theory provides an efficient mathematical tool for image sparse representation including geometry and multiscale analysis. It can be considered as an extension of the wavelet transform, by increasing their directional sensitivity, to be more adapted for anisotropic image objects (e.g., edges, key-points, etc.).

Indeed, shearlet coefficients construction can be summarized by the reposition of three operators: *dilatation* $\mathtt{D}_{\mathbf{A}_\alpha}$, *shearing* $\mathtt{D}_{\mathbf{S}_\beta}$ and *translation* $\mathtt{T}_t$, applied on a generating waveform function ϕ (also called mother shearlet function). We invite the reader to refer to [14] for more details.

Firstly, an operator $\mathtt{D}_{\mathbf{M}}$ is needed to compute the *dilatation* and the *shearing* function.

$$\mathtt{D}_{\mathbf{M}}\phi(x) = |det(\mathbf{M})|^{-\frac{1}{2}}\phi(\mathbf{M}^{-1}x) \qquad \mathbf{M} \in GL_d(\mathbb{R}) \tag{1}$$

with $x \in \mathbb{R}^2$, $GL_d(\mathbb{R})$ the group of d-dimensional invertible matrices, *det*(**M**) the determinant of the matrix **M**.

To obtain the *dilatation function* $\mathtt{D}_{\mathbf{A}_\alpha}$ we use this operator $\mathtt{D}_{\mathbf{M}}$ with the parameter **M** as a parabolic scaling matrix given by

$$\mathbf{A}_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & \sqrt{\alpha} \end{pmatrix} \qquad \alpha \in \mathbb{R}^{+*} \tag{2}$$

Also, to obtain the *shearing function* $\mathtt{D}_{\mathbf{S}_\beta}$, we use the operator $\mathtt{D}_{\mathbf{M}}$ with the parameter **M** as

$$\mathbf{S}_\beta = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \qquad \beta \in \mathbb{R} \tag{3}$$

where β is the slope that parametrizes the orientation of the *generating function*.

Finally, the *translation function* $\mathtt{T}_t \in L^2(\mathbb{R}^2)$ is applied using

$$\mathtt{T}_t\phi(x) = \phi(x-t) \qquad (x,t) \in (\mathbb{R}^2)^2 \tag{4}$$

From this, the shearlet system $SH(\phi)$ associated to $\phi \in L^2(\mathbb{R}^2)$ is given by

$$SH(\phi) = \left\{ \phi_{(\alpha,\beta,t)} = \mathtt{T}_t\mathtt{D}_{\mathbf{A}_\alpha}\mathtt{D}_{\mathbf{S}_\beta}\phi, \quad \alpha \in \mathbb{R}, \quad \beta \in \mathbb{R}, \quad t \in \mathbb{R}^2 \right\} \tag{5}$$

By replacing (2), (3) and (4) in (5), we get

$$\begin{aligned} \phi_{(\alpha,\beta,t)}(x) &= \alpha^{-\frac{3}{4}}\phi\left(\mathbf{A}_\alpha^{-1}\mathbf{S}_\beta^{-1}(x-t)\right) \\ &= \alpha^{\frac{-3}{4}}\phi \begin{pmatrix} \frac{1}{\alpha} & \frac{-\beta}{\alpha} \\ 0 & \frac{1}{\sqrt{\alpha}} \end{pmatrix} \left(x-t\right) \end{aligned} \tag{6}$$

Moreover, according to [14] shearlet decomposition is possible if and only if ϕ satisfies the following *admissible condition*

$$\int_{\mathbb{R}^2} \frac{|\widehat{\phi}(\xi_1,\xi_2)|^2}{\xi_1^2} d\xi_1 d\xi_2 < \infty \qquad (\xi_1,\xi_2) \in \mathbb{R}^2 \tag{7}$$

Thereby, the mother shearlet function $\phi \in L^2(\mathbb{R}^2)$ is defined such that

$$\begin{aligned} \widehat{\phi}(\xi) &= \widehat{\phi}(\xi_1,\xi_2) \\ &= \widehat{\phi}_1(\xi_1)\widehat{\phi}_2\left(\frac{\xi_2}{\xi_1}\right) \end{aligned} \tag{8}$$

where, $\widehat{\phi}$ is the Fourier transformation of ϕ, ϕ_1 is a wavelet function and ϕ_2 is a bump-like function. Then, the shearlet transform of the image $I \in L^2(\mathbb{R}^2)$ is

$$I \rightarrow SH_\phi I(\alpha,\beta,t) = < I, \phi_{\alpha,\beta,t} > \tag{9}$$

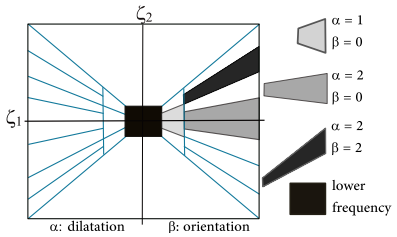

Fig. 1. Representation of a frequency plane tiled by the shearlet system. The horizontal cone is represented in blue.

As shown in Fig.1, the direction ξ_1 is *naturally* favored to the direction ξ_2. To tackle this problem, the frequency plane is divided into two cones: the vertical one corresponds to $\|\frac{\xi_2}{\xi_1}\| < 1$ and $\|\xi_1\| > 1$, and the horizontal one corresponds to $\|\frac{\xi_1}{\xi_2}\| < 1$ and $\|\xi_2\| > 1$. Concerning the low-frequency, it is filtered by a square centered in the origin ($\|\xi_1\| < 1$ and $\|\xi_2\| > 1$).

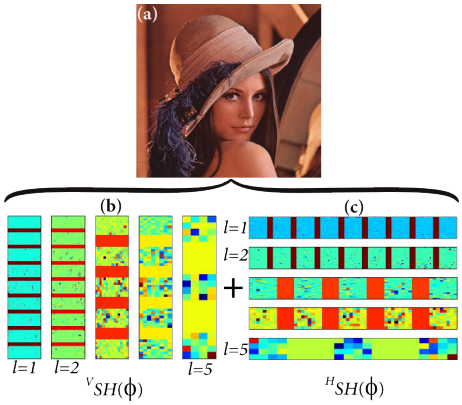

Fig. 2. Shearlet coefficients (level $l = 1...5$) computed from the *Lena* photography (a). (b) the horizontal coefficients ${}^{H}SH(\phi)$ and (c) the vertical ones ${}^{V}SH(\phi)$.

Figure 2 shows the representation of the shearlet coefficients of a 128 × 128 pixels image (e.g., *Lena* photography). These coefficients correspond to the different levels $l = 1,\ldots 5$ ($l = 1$ corresponds to the highest resolution coefficients and $l = 5$ the lowest ones). Each level also contains a certain number of subimages, depending on the number of the chosen shearing directions.

III. SHEARLET-BASED VISUAL SERVOING

A. Basics of Visual Servoing

According to [2], the aim of a visual servoing is to control the motion of a robot in order to allow a set of visual signal $\mathbf{s}$ ($\mathbf{s} \in \mathbb{R}^k$) defining a robot pose $\mathbf{r}(t) \in SE(3)$ $\Big($i.e., $\mathbf{s} = \mathbf{s}\big(\mathbf{r}(t)\big)\Big)$ to reach a set of desired signal $\mathbf{s}^*$ ($\mathbf{s}^* \in \mathbb{R}^k$) by minimizing a visual error given by

$$\mathbf{e} = \mathbf{s}\big(\mathbf{r}(t)\big) - \mathbf{s}^* \tag{10}$$

The variation of $\mathbf{s}$ is linked to the velocity screw vector $\mathbf{v} = (v_x \quad v_y \quad v_z \quad \omega_x \quad \omega_y \quad \omega_z)^T$ of the visual sensor frame by $\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v}$ where $\mathbf{L_s} \in \mathbb{R}^{k\times 6}$ is the *interaction matrix*.

By taking into account (10), we obtain the variation of the visual error $\mathbf{e}$ due to the visual sensor velocity such as $\dot{\mathbf{e}} = \mathbf{L_s}\mathbf{v} - \dot{\mathbf{s}}^*$. In order to obtain an exponential decrease of this visual error, the following classical control law is usually employed [2]

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_s}}^{+}\big(\mathbf{s}(t) - \mathbf{s}^*\big) \tag{11}$$

where λ is a positive gain and $\widehat{\mathbf{L_s}}^{+}$ is the Moore-Penrose pseudo-inverse of an estimation of the interaction matrix. In this study we choose to fix $\widehat{\mathbf{L_s}} = \mathbf{L_s}^*$ to be the interaction matrix $\mathbf{L_s}^*$ obtained at the desired position of the visual sensor as it is often done.

B. Control Law using Shearlet Coefficients

In this work, we propose to use the k shearlet coefficients of the coarsest level $l = 5$ (in the right of (Fig. 2 (b)) and in the bottom of (Fig. 2 (c)) as the set of visual features $\mathbf{s}$ for the image I, such as

$$\mathbf{s} = < I, \phi_{\alpha,\beta,t} >_{(l=5)} \qquad \alpha = 1, \quad \beta = [-1, 0, 1] \tag{12}$$

Since an analytical formulation of the interaction matrix related to the shearlet coefficients is difficult to derive, at this stage of development, we propose to numerically compute the constant interaction matrix $\mathbf{L_s}^*$ used in the control law (11) thanks to an off-line initial robotic procedure that we performed at the desired location of the probe. The principle consists in applying successively and independently small relative displacements corresponding to 3 translations Δx, Δy, Δz and 3 rotations $\Delta\theta_x$, $\Delta\theta_y$, $\Delta\theta_z$ along and around each axis of the Cartesian frame attached to the robot end-effector (as depicted in Fig. 7). The measured visual feature differences (scalar values) $\Delta s^j_{\Delta x}$, $\Delta s^j_{\Delta y}$, $\Delta s^j_{\Delta z}$, $\Delta s^j_{\Delta\theta_x}$, $\Delta^j s_{\Delta\theta_y}$, $\Delta s^j_{\Delta\theta_z}$ induced by these motions for each visual feature component j allow then to computed the j^{th} row of the interaction matrix $\mathbf{L_s}^*$ as follow

$$\mathbf{L}^*_j = \begin{pmatrix} \frac{\Delta s^j_{\Delta x}}{\Delta x} & \frac{\Delta s^j_{\Delta y}}{\Delta y} & \frac{\Delta s^j_{\Delta z}}{\Delta z} & \frac{\Delta s^j_{\Delta\theta_x}}{\Delta\theta_x} & \frac{\Delta s^j_{\Delta\theta_y}}{\Delta\theta_y} & \frac{\Delta s^j_{\Delta\theta_z}}{\Delta\theta_z} \end{pmatrix} \tag{13}$$

Finally, the whole interaction matrix is obtained by stacking all the rows $\mathbf{L}^*_j$ related to each visual feature component $\mathbf{L_s}^* = \big[\mathbf{L}^*_1, \ldots, \mathbf{L}^*_k\big]$ and the control law (11) is applied with this numerically estimated constant interaction matrix $\mathbf{L_s}^*$.

C. Convergence domain

In order to judge the effectiveness of the proposed controller in term of convergence domain, we have computed the cost-function shape when varying the different robot DOF using the following relationship

$$C(s) = \Big(((\mathbf{s} - \mathbf{s}^*)^\top (\mathbf{s} - \mathbf{s}^*))/N_{pix}\Big)^{\frac{1}{2}} \tag{14}$$

where, N_{pix} is the number of pixel in the initial image. As it can be seen in Fig. 3, the cost-function C has an almost perfect convex shape where the global minimum is clearly identified.

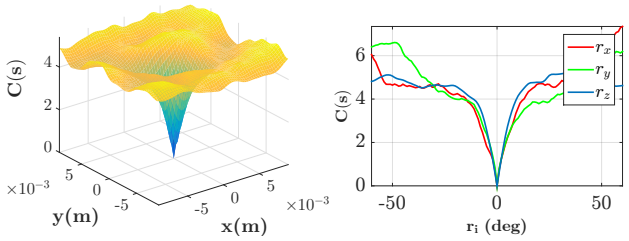

Fig. 3. (a) the shearlet-based visual servoing cost-function for x and y translation and (b) for the rotations around x, y and z axis.

IV. GROUND-TRUTH VALIDATION

To validate the proposed methods, we develop an automatic 6 DOF positioning task by controlling the 2D US robot holder motion. This scenario was performed in both simulation and experimental modes. The obtained results are presented and discussed below.

A. Simulation Results

Before the proposed approach can be validated on the experimental work-flow, it was tested using a developed C++ simulator which was very appropriate to validate the theoretical developments in different conditions of use without any risk of damaging the US probe. The simulator, miming a 2D US probe and an abdomen phantom, was implemented using the: Open-source Visual Servoing Platform (ViSP)[1], Visualization ToolKit library (VTK)[2] and Shearlab library[3] (available in MatLab version and implemented by ourselves in C++). An example of a US volume as well as a 2D US slice are shown in Fig. 4. This US volume is composed by 335 parallel 2D images of size 250 × 250 with a pixel size of 0.6mm × 0.6mm and interval of 0.3mm between each image (captured with a real 3D US Probe).

The test consists of performing a 6 DOF automatic positioning task using the shearlet-based visual controller. To do that, the positioning task consists of reaching a desired position $\mathbf{r}^* \in SE(3)$ (i.e., desired image) from an arbitrary initial position $\mathbf{r} \in SE(3)$ (i.e., an initial image). At each US image acquisition, the shearlet coefficients are computed, more precisely the first 96 coarsest coefficients. The spatial

[1] www.irisa.fr/lagadic/visp
[2] www.vtk.org
[3] www.shearlab.org

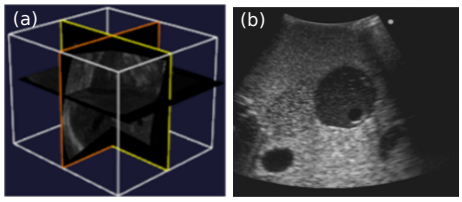

Fig. 4. (a) 3D US volume provided by the developed simulator and (b) an extracted 2D US slice.

velocities computed by the controller (11) are applied to the virtual probe with a gain $\lambda = 0.5$ in order to reach the desired position. Figure 5 depicts an image sequence acquired during the positioning task: Fig. 5(a)-(b) show the initial and desired images, respectively, and Fig. 5(c) represents the initial image difference that is observed for an initial error of $\Delta\mathbf{e}_{init}$ (mm, deg) = (9, $\times$, 5, 2, 2, 5) while Fig. 5(d) shows the image difference at the last recorded iteration, which shows the convergence of the control law.

$$\mathbf{I}_{diff} = \frac{(\mathbf{I} - \mathbf{I}^*) + 255}{2} \tag{15}$$

The pose error obtained at the convergence using high resolution robot encoders supplied by the simulation software, was measured to be $\Delta\mathbf{e}_{final}$ **(mm, deg) =** $\big($**0.1,** $\times$ **, 0.2, 0.02, 0.18, 0.08**$\big)$. Figure 6 (left) shows the pose error decay for each DOF when Figure 6 (right) illustrates the velocities sent to the robot DOFs. After about thirty seconds, we can see that the system has converged smoothly.

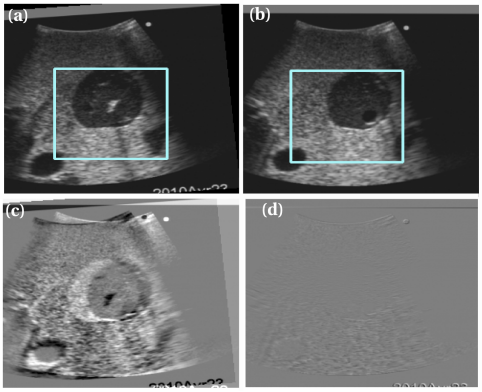

Fig. 5. Image sequence captured during the positioning task: (a) the initial image $\mathbf{I}(\mathbf{r_0})$, (b) the desired image $\mathbf{I}(\mathbf{r}^*)$, (c) the difference $\mathbf{I}_{diff}$ between the initial and the desired image and (d) the difference between the final and the desired images.

B. Experimental Results [nominal conditions]

The proposed controller was tested and validated in experimental conditions using a realist testbench (Fig. 7) equipped with

- a 6 DOF anthropomorphic robotic arm of type of *Adept Viper* s850;
- a 2-5MHz 2D-US transducer (C60, Sonosite 180+) with a depth of 12cm providing 480 $\times$ 640 pixels US images with a pixel size of 0.35 $\times$ 0.36mm^2;

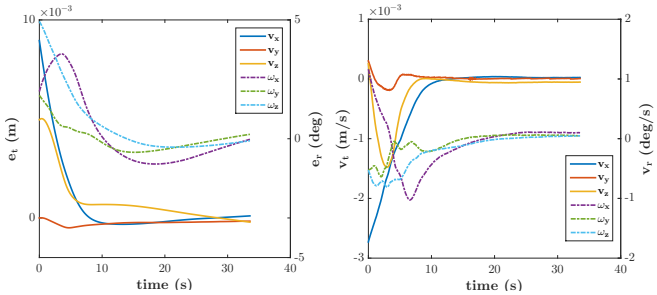

Fig. 6. Positioning task *[simulation case]*: (left) pose error decay in each DOF (e_t for translations and e_r for rotations) and (right) velocity evolution in each DOF (v_t and v_r respectively for the translational and rotational components of the robot control velocity screw)

- a 6 axes force sensor ATI Gamma SI-65-5;
- a 3-GHz PC running under a Linux distribution in which were implemented the developed approaches in C++ and the control communication with the robot;
- an abdominal phantom of type of AB-41900-030 Kyoto Kagaku-ABDFAN.

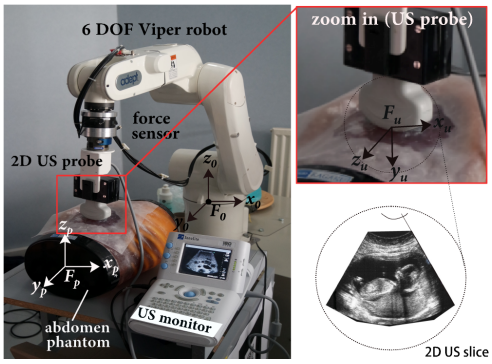

Fig. 7. Photography of the experimental set-up.

The objective of this experiment is to reproduce the simulation tests on a realist testbench closer to a clinical use of the US imaging system. For validation purpose, the US probe attached to the 6 DOF *Viper* robot is positioned to a reference location (preoperative image) where the desired shearlet components were measured and the related interaction matrix was numerically estimated, then without moving the phantom, we position randomly the probe at a new location which is considered as the initial position (intraoperative image). The preoperative image can be considered as a desired image acquired few days earlier and the intraoperative one captured during another examination. Thereby, the objective of our approach is to retrieve automatically and accurately the preoperative image during a second examination.

Furthermore, the y-translation corresponds to the contact direction between the US probe and the phantom. Indeed, it is advisable to be careful (low motion) along the y axis because of the risk of damaging the probe and/or phantom. To tackle this problem, a force control is used to control this y-translation instead of the visual controller (for more details, please refer to [9]) to maintain a desired contact

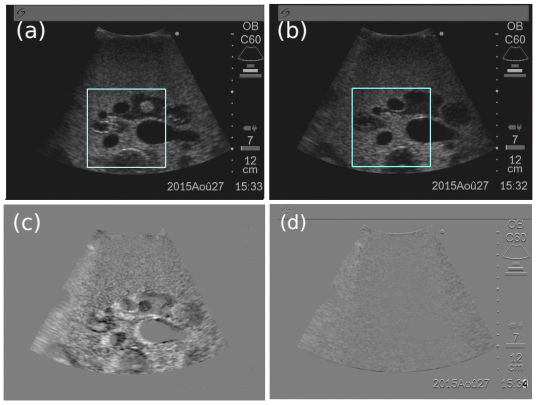

Fig. 8. Image sequence captured during a positioning task: (a) the initial image $\mathbf{I}(\mathbf{r_0})$, (b) the desired image $\mathbf{I}(\mathbf{r}^*)$, (c) the difference between the initial and the desired image and (d) the difference between the final and the desired images $\mathbf{I}_{diff}$.

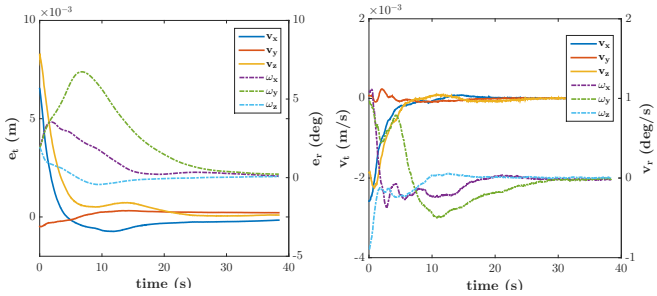

Fig. 9. Positioning task *[experimental case, nominal conditions]*: (left) pose error decay in each DOF and (right) velocity evolution in each DOF.

force of 2N while the others 5 DOF are controlled by our shearlet-based visual servoing approach.

Figure 8 represents the case where the initial pose error between the initial US image $\mathbf{I}(\mathbf{r_0})$ and the desired US image $\mathbf{I}(\mathbf{r}^*)$ was measured to be $\Delta\mathbf{e}_{init}$ (mm, deg) = (7, ×, 8, 2,2, 2). The final positioning error (Fig. 9 (left)), computed using the high resolution robot encoders supplied by the robot software, was measured to be $\Delta\mathbf{e}_{final}$ **(mm, deg) = (0.16, ×, 0.09, 0.09, 0.22, 0.07)** demonstrating a good accuracy and convergence behavior.

The positioning task was repeated several times using different initial positions (in order to avoid biased results). It was experimentally demonstrated that in each test the controller reaches successfully the desired position. TABLE I summarizes some samples of positioning tasks performed by the proposed controller. Note that the y-translation is not taken into account since this axis was controlled by force control.

C. Experimental Results [unfavorable conditions]

In order to judge the robustness of the shearlet-based visual servoing, another experimental test was performed. During this test, the reference US image (Fig. 10(b)) was acquired in unfavorable conditions. Indeed, we learn the shearlet coefficients on the desired image blurred with a Gaussian filter of 31 pixels size. During the task, the current

TABLE I

ROBUSTNESS STUDY OF THE CONTROLLER ($\triangle T_i$ (MM), $\triangle R_i$ (DEG), $\mathbf{e}_0$ AND $\mathbf{e}_f$ REPRESENT THE INITIAL AND FINAL ERROR RESPECTIVELY.)

$\mathbf{e}_f$ and $\mathbf{e}_0$	$\triangle T_x$	$\triangle T_y$	$\triangle T_z$	$\triangle R_x$	$\triangle R_y$	$\triangle R_z$
$\mathbf{e}_0$	1	×	2	5	5	5
$\mathbf{e}_f$	0,41	×	0,09	0,06	0,13	0,14
$\mathbf{e}_0$	10	×	5	5	5	5
$\mathbf{e}_f$	0,3	×	0,26	0,37	0,5	0,21
$\mathbf{e}_0$	7	×	8	2	2	2
$\mathbf{e}_f$	0,16	×	0,09	0,09	0,22	0,07
$\mathbf{e}_0$	8	×	5	2	2	5
$\mathbf{e}_f$	0,41	×	0,11	0,11	0,3	0,85
$\mathbf{e}_0$	4	×	5	5	5	5
$\mathbf{e}_f$	0,3	×	0,06	0,1	0,22	0,08
$\mathbf{e}_0$	2	×	4	5	5	4
$\mathbf{e}_f$	0,09	×	0,02	0,03	0,35	0,31
MEAN(e)	**0,28**	×	**0,10**	**0,13**	**0,29**	**0,28**
STD (e)	**0,13**	×	**0,08**	**0,12**	**0,13**	**0,29**

US images were captured in nominal conditions (Fig. 10(a)). The initial pose error is $\Delta\mathbf{e}_{init}$ (mm, deg) = (9, ×, 3, 3, 1, 4).

Figure 10 depicts some US images captured during the positioning task process. More precisely, Fig. 10(a) and (b) show the initial image and the desired image (in poor quality) when Fig. 10(c) illustrates the difference between the initial and desired image $\mathbf{I}_{diff}$ and Fig. 10(d) the difference between the final and the desired image demonstrating the accurate convergence of the controller towards the desired position.

The positioning error decay in each DOF and the evolution of the spatial velocity over the time are reported in Fig. 11(left) and (right). As can be underlined, the controller converges exponentially and smoothly to the desired position with a good accuracy. The obtained final error was measured to be $\Delta\mathbf{e}_{final}$ **(mm, deg) = (0.229, ×, 0.353, 0.59, 0.56, 0.19)** very close to the ones obtained in the case of nominal conditions demonstrating the robustness of our method *w.r.t.*

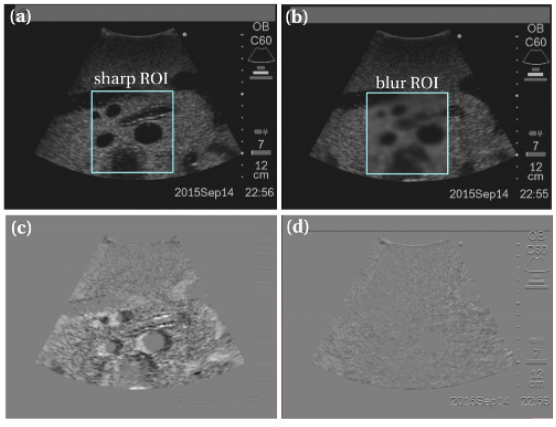

Fig. 10. Some snapshots captured during the positioning task *[unfavorable conditions]*: (a) the initial image $\mathbf{I}(\mathbf{r_0})$, (b) the desired image $\mathbf{I}(\mathbf{r}^*)$, (c) the difference $\mathbf{I}_{diff}$ between the initial image and the desired image and (d) the difference between the final and the desired images.

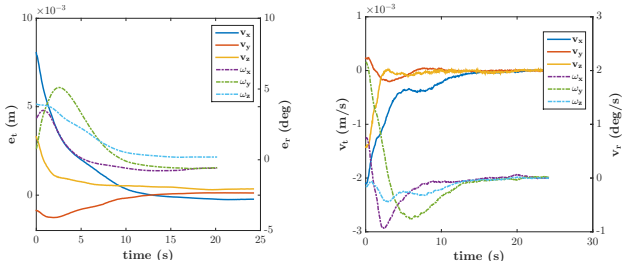

Fig. 11. Positioning task *[experimental case, unfavorable conditions]*: (left) pose error decay in each DOF, and (right) velocity evolution sent to the robot DOFs.

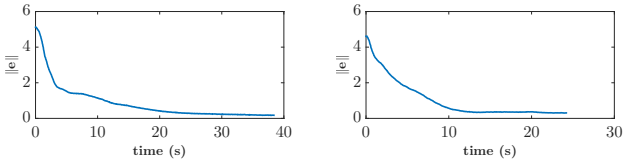

Fig. 12. (left) shows the decrease of the norm of the visual error in the case of nominal conditions, and (right) illustrates the evolution of the norm of the visual error in the case of unfavorable conditions.

blurred images. Furthermore, Fig. 12 (left) and (right) depict the decrease of the norm of the visual error $\|e\|$ given by eq. (14) in nominal and unfavorable conditions, respectively.

D. Towards Tracking using Shearlet-based Visual Servoing

The developed controller was also used to achieve tracking tasks. The latter consists in automatically moving the US probe to follow phantom motions. To do this, we use the same controller by considering the desired image as being the initial one. Therefore the robotic system aims to automatically stabilize the US images through the compensation of the target external motions. The objective of this task is to compensate the physiological motions (e.g., breathing and heartbeat) during an US-based examination. For this experiment, we manually move the phantom along the 6 DOF with an amplitude of ΔA (mm, deg) = (20, 2, 2, 3, 17, 3). As can be shown in Fig. 13 and in the video accompanying this paper, our shearlet-based controller is able to efficiency compensate the applied external movements.

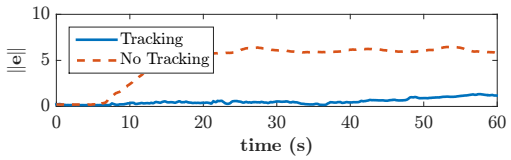

Fig. 13. Evolution of the error in the Shearlet domain with and without tracking.

V. CONCLUSION & FUTURE WORKS

In this paper was presented a new 6 DOF US image-based visual servoing paradigm. Instead of using conventional visual information in the design of the control lax, our method makes use of the shearlet transform coefficients. This implied that the interaction matrix links directly the US images shearlet coarsest coefficients time-variation to the robot motion. The method was validated in both simulations and experiments using a 2D US probe attached to a 6 DOF anthropomorphic robotic arm. The obtained results in the cases of automatic probe positioning and phantom motion compensation demonstrated the efficiency of the proposed controller in terms of accuracy (average final positioning error of **0.19mm** and **0.23**° in the translation and rotation, respectively), repeatability, robustness and convergence behavior.

Further work will be undertaken to improve the developed controller by providing an analytical formulation of the interaction matrix avoiding therefore the need of its off-line numerical estimation. We will also consider the problem of US images acquisition time and quality. This can be tackled by considering the *Compressed Sensing* techniques well-adapted for medical images e.g., US, CT-scan, MRI, etc. The compressed sensing technique can also be taken into account during shearlet-based visual servoing.

VI. ACKNOWLEDGEMENTS

This work is partially supported by DESWEEP project funded by the *Région de Bretagne, France* and from the project NEMRO (ANR-14-CE17-0013-001) funded by the ANR, France. It is also performed in the framework of the Labex ACTION (ANR-11-LABEX-01-01).

REFERENCES

[1] F. Courreges, P. Vieyres, R. Istepanian, P. Arbeille, and C. Bru, "Clinical Trial and evaluation of a mobile, robotic tele-ultrasound system," *J. of Telemedicine and Telecare*, vol. 11, pp. 46–55, 2005.

[2] F. Chaumette and S. Hutchinson, "Visual servo control. part I: basic approaches," *IEEE Rob. & Aut. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.

[3] A. Krupa, D. Folio, C. Novales, P. Vieyres, and T. Li, "Robotized tele-echography: an assisting visibility tool to support expert diagnostic," *IEEE Syst. J.*, pp. 1–10, 2015.

[4] P. Abolmaesumi, S. Salcudean, W.-H. Zhu, M. Sirouspour, and S. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Trans. on Rob. and Aut.*, vol. 18, no. 1, pp. 11–23, 2002.

[5] R. Mebarki, A. Krupa, and F. Chaumette, "2D ultrasound probe complete guidance by visual servoing using image moments," *IEEE Trans. on Rob.*, vol. 26, no. 2, pp. 296–306, 2010.

[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, p. 441, 2004.

[7] D. Koichiro, "A direct interpretation of dynamic images with camera and object motions for vision guided robot control," *Int. J. of Computer Vision*, vol. 37, no. 1, pp. 7–20, 2000.

[8] B. Tamadazte, N. Piat, and E. Marchand, "A direct visual servoing scheme for automatic nanopositioning," *IEEE/ASME Trans. on Mech.*, vol. 17, no. 4, pp. 728–736, 2012.

[9] C. Nadeau and A. Krupa, "Intensity-based ultrasound visual servoing: modeling and validation with 2d and 3d probes," *IEEE. Trans. on Rob.*, vol. 29, no. 4, pp. 1003–1015, 2013.

[10] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 25–46, 2008.

[11] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[12] E. King, G. Kutyniok, and W.-Q. Lim, "Image inpainting: theoretical analysis and comparison of algorithms," in *SPIE Opt. Eng. App.*. Int. Society for Optics and Photonics, pp. 885 802–885 802, 2013.

[13] D. Gupta, R. Anand, and B. Tyagi, "Enhancement of medical ultrasound images using multiscale discrete shearlet transform based thresholding," in *Int. Symp. on Electronic Syst. Des.*, pp. 286–290, 2013.

[14] G. Kutyniok and D. Labate, "Introduction to shearlets," in *Shearlets*. Springer, pp. 1–38, 2012.