

# Wavelet-based visual servoing using OCT images

L.-A. Duflot, B. Tamadazte, N. Andreff and A. Krupa

**Abstract**—This paper deals with the development of an Optical Coherence Tomography (OCT) based visual servoing. The proposed control law uses the wavelet coefficients of the OCT images as the signal control inputs instead of the conventional geometric visual features (points, lines, moments, etc.). An important contribution is the determination of the interaction matrix that links the variation of the wavelet coefficients to the OCT probe (respectively to the robotic platform) spatial velocity. This interaction matrix, required in the visual control law, is obtained from time-derivation of the wavelet coefficients. This work is carried out in a medical context which consists of automatically moving a biological sample in such a way to go back to the position of a sample region that corresponds to a previous optical biopsy (OCT image). For instance, the proposed methodology makes it possible to follow accurately the progress of a pathological tissue between an optical biopsy and a former one. The developed method was experimentally validated using an OCT imaging system placed in an eye-to-hand configuration viewing the robotic platform sample holder. The obtained results demonstrated the feasibility of this type of visual servoing approach and promising performances in terms of convergence and accuracy.

## I. INTRODUCTION

### A. Motivation

The wavelet transform is a mathematical tool introduced in 1910 by *Alfred Haar* but has grown in popularity only in the late 1980s with the arrival of the early work of *Yves Meyer* [1]. The rise of digital image processing and the need of compression in order to process, encode, transmit, and save images with increasing resolutions have led to the development of advanced and efficient compression methods. The wavelet framework was particularly used for elaborating the JPEG2000 compression [2]. In fact, wavelet multiscale decomposition of an image is obtained by repeatedly smoothing and subsampling an image signal while also storing the detail information lost at each stage of this process. Such decomposition is not only useful for defining increasingly coarse approximations of images, but also to obtain sparse representations which means that most coefficients describing the detail information lost in each transformation step are typically close to zero. Therefore, the coarse approximation coefficients of an image can be considered as signal control inputs in the formulation of a robotic visual

control law. A 6-degrees-of-freedom (DOF) visual servoing approach based on the low-pass approximation of an image obtained from a discrete wavelet transform was proposed in [3] using a conventional projective camera as a sensor. In the same manner, shearlet transform (considered as the extension of the wavelets one) was applied in the context of medical image-guided tasks, where a 6 DOF ultrasound (US) based visual servoing method was proposed for automatically positioning an ultrasound probe actuated by a robotic arm [4]. These methods have demonstrated very interesting behavior during the different experimental validations of the respective control laws. These preliminary investigations already demonstrated the feasibility and potential benefits of considering wavelet-based multiscale representations in a visual servoing scheme, in particular with respect to unfavorable conditions (noise, partial image occlusion, illumination changes). In fact, using multiscale wavelet decomposition could serve as an interesting compromise between purely geometric feature-based approaches [5] and image-intensity-based (photometric) ones [6], [7]. This is especially valuable in case of less textured images such as medical images: US, OCT, in which the detection, extraction and tracking of visual features over time is very challenging.

Our work takes place in an ambitious project dealing with the development of a microrobotic-based nasal endoscopy system embedding a fiber-based OCT system. The medical aim is to perform optical biopsy of the olfactory mucosa in order to prevent an eventual smell deficiency which can be directly related to the appearance of neurodegenerative diseases. In this paper we present one important aspect of this project that concerns the development of a visual control law in which the robotic sensor is an OCT imaging system. The objective is to automatically align with a robotic platform a tissue sample observed by the OCT system in such a way to reach a desired OCT image previously captured at an early stage of the pathology for examination purpose. OCT is an emerging technological imaging device used to obtain high resolution cross-sectional tomographic images. First developed in 1991 [8], this device, which allows reaching a micrometric resolution and millimetric penetration depth, was tested on retina and on the coronary artery. Actually, OCT is more and more often used for medical mini- and non-invasive applications, in particular for ophthalmology and dermatology. This device uses low-coherence interferometry with near-infrared light in order to produce cross-section images of a scattering medium. Nevertheless, the acquisition of 3D volumes remains slow (nearly one minute for a centimetric size area) due to mechanical scanning as well as the size of the acquired data.

L.-A. Duflot and A. Krupa are with Univ Rennes, Inria, CNRS, IRISA, Campus universitaire de Beaulieu, Rennes 35042, France [lesley-ann.duflot@inria.fr](mailto:lesley-ann.duflot@inria.fr) - [alexandre.krupa@inria.fr](mailto:alexandre.krupa@inria.fr)

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](mailto:brahim.tamadazte@femto-st.fr) - [nicolas.andreff@femto-st.fr](mailto:nicolas.andreff@femto-st.fr)

## B. Contributions

A CCD camera is rigidly attached to our OCT probe. That is why, one previous work was already using a wavelet decomposition of the OCT scan to control the in-plane movements [9]. The image captured by the CCD camera was used to control the other DOF. Our goal now is to control all the DOF of the robot based on OCT information. Given that volume acquisition remains too slow to apply a visual servoing scheme, this paper will propose a direct visual servoing using only OCT cross-sectional image as visual information. All the DOF can theoretically be controlled with the work presented here. Nevertheless, since rotations are difficult to manage with our robot with such resolutions, only 3 DOF will be controlled experimentally. They correspond to the two in-plane ( $xz$ ) and the out-of-plane  $y$  translations of the OCT image plane (see Fig. 3 for the definition of the  $x$ ,  $y$  and  $z$  axes). The wavelet decomposition is not the same as the one used in [9] and use more coefficient as signal input in order to upgrade the controller. The work presented in [10] is also making a 6 DOF visual servoing using OCT probe with a geometric feature-based approach.

## C. Plan

In the sequel, Section II first recalls the general framework of the wavelet transform and then presents our contribution on the design of a visual servoing approach that considers the wavelet coefficients as visual features (signal control inputs). Section III shows the experimental validation of the proposed wavelet-based visual servoing approach using a Fourier-Domain OCT (FD-OCT) imaging system as a robotic sensor. This section also provides a discussion on the performances of the developed method.

# II. FRAMEWORK

## A. Wavelet Signal Decomposition

Real-world data, like signal or images, are composed by smooth (homogeneous) regions and non-smooth ones (e.g., edges). These edges often represent the most interesting information for signal processing. In contrast to Fourier transforms, which only yield information about which frequencies are present in a signal but not their locations, wavelet transforms allow a simultaneous and efficient analysis of the behavior of a signal in both the time (in 1D case) or the space (in 2D case) and the frequency domains.

In fact, a wavelet is a rapidly decaying wave-like oscillation, centered in 0 and existing for a given location, contrary to sine functions in the Fourier transform. A wavelet system is constructed by shifting and scaling a wavelet function  $\psi$  in the Hilbert space  $L^2(\mathbb{R})$  of square integrable functions. Then the generating wavelet  $\psi$  shifted by  $k$  and scaled by  $j$  is described by

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k) \quad (1)$$

In the late 1980s, pioneers like *Ingrid Daubechies* [11] and *Stéphane Mallat* [12] introduced discrete wavelet transforms more adapted for implementation. In this paper, we opt

for the *Daubechies* wavelets included within the MATLAB WAVELET TOOLBOX.

Note that the basic formula of a continuous wavelet transform of a function  $f \in L^2(\mathbb{R})$  is defined by

$$(\mathcal{W}_\psi f)(j, k) = \langle f, \psi_{j,k} \rangle = \int_{\mathbb{R}} \psi\left(\frac{x - k2^{-j}}{2^{-j}}\right) f(x) dx \quad (2)$$

Thus, the wavelet transform allows extracting different levels of details coefficients  $B_1$ ,  $B_2$  and  $B_3$  and smooth approximation coefficients  $A_3$  of a given signal  $S$ .

In practice, a discrete wavelet transform is performed by applying successively high-pass and low-pass filters to a signal  $S$  as represented in Fig. 1.

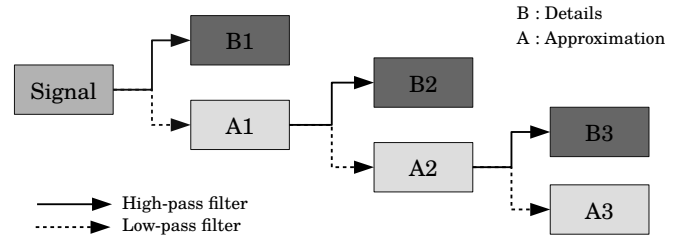


Fig. 1: Diagram of a discrete wavelet decomposition of a signal  $S$ .

Therefore, in the case of 2D signal (i.e., image), we consider low-pass and high-pass filtering along each direction leading to 3 types of detail coefficients: horizontal, vertical and diagonal ones and approximation coefficients. In other words, the details coefficient previously introduced by  $B$  are divided in three types: horizontal  $h$ , vertical  $v$  and diagonal  $d$  coefficients (see Fig. 2).

## B. Control Law

The objective of a visual servoing scheme is to control the motion of a robotic structure in order to allow that a set of visual features  $s$  ( $s \in \mathbb{R}^k$ ), defining a robot pose  $r(t) \in SE(3)$  (i.e.,  $s = s(r(t))$ ), reaches a set of desired features  $s^*$  ( $s^* \in \mathbb{R}^k$ ) expressing the desired robot pose  $r^* \in SE(3)$ . In other words, the aim is to minimize the visual error given by

$$e = s(r(t)) - s^* \quad (3)$$

For this purpose, it is needed to establish the link between the velocity twist vector, which in our case contains only the translation components  $v = (v_x \ v_y \ v_z)^T$  of the OCT system frame, and the variations of  $s$  by  $\dot{s} = L_s v$  where  $L_s \in \mathbb{R}^{k \times 3}$  is called the *interaction matrix*.

In order to obtain a decrease of the error  $e$  such as  $\dot{e} = -\lambda e$ , with  $\lambda$  being a positive gain, the classical controller usually considered in the literature is given by [5]

$$v = -\lambda \widehat{L}_s^+(s(t) - s^*) \quad (4)$$

where  $\widehat{L}_s^+$  is the Moore-Penrose pseudo-inverse of an estimation of the interaction matrix. We propose therefore to

use this control law to control the 3 DOF robotic platform in the OCT system frame.

In this work, the visual information  $s$  used in the control loop consists of the wavelet coefficients computed by the discrete *Daubechies* wavelet transform  $W_\psi(\cdot)$  on a square image  $\mathbf{I}$  of size  $(N \times N)$  as

$$\begin{pmatrix} \mathbf{A}_3 & \mathbf{h}_3 & \mathbf{h}_2 & \mathbf{h}_1 \\ \mathbf{v}_3 & \mathbf{d}_3 & & \\ & \mathbf{v}_2 & \mathbf{d}_2 & \\ & & \mathbf{v}_1 & \mathbf{d}_1 \end{pmatrix} = \mathcal{W}_\psi(\mathbf{I}) \quad (5)$$

More precisely, we take into account only one quarter of the coefficients which are  $\mathbf{B}_2$ ,  $\mathbf{B}_3$  and  $\mathbf{A}_3$  such as

$$\mathbf{B}_m = [\mathbf{d}_m; \mathbf{h}_m; \mathbf{v}_m] \quad \text{with index } m = \{2, 3\} \quad (6)$$

The resulting visual feature vector is therefore built as

$$\begin{aligned} s = & [\mathbf{B}_2(1, 1), \dots, \mathbf{B}_2(i, j), \dots, \mathbf{B}_2(3N/4, N/4), \\ & \mathbf{B}_3(1, 1), \dots, \mathbf{B}_3(i, j), \dots, \mathbf{B}_3(3N/8, N/8), \\ & \mathbf{A}_3(1, 1), \dots, \mathbf{A}_3(i, j), \dots, \mathbf{A}_3(N/8, N/8)] \end{aligned} \quad (7)$$

As example, Fig. 2 shows the *Daubechies* wavelet decomposition of an OCT image (a  $xz$  image slice) of a LEGO toy element. The first level of details  $\mathbf{B}_1$  corresponds to the higher frequency assimilated to noise and is therefore not kept to compute the visual features. As  $N = 512$  in our test case, then the size of  $s$  is equal to  $256 \times 256 = 65536$  coefficients.

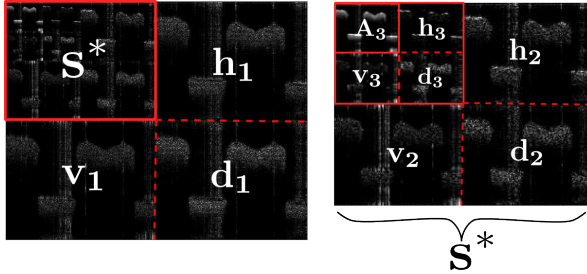


Fig. 2: Wavelet transform of an OCT image of a LEGO toy element

The estimation of the interaction matrix  $\widehat{\mathbf{L}}_s$  that links the variations of  $s$  to the velocity vector  $\mathbf{v}$  is computed once at the desired position by calculating the wavelet transform of the image gradient along each direction

$$\widehat{\mathbf{L}}_s = [L_{\mathbf{B}_2}(1, 1), \dots, L_{\mathbf{B}_2}(i, j), \dots, L_{\mathbf{B}_2}(3N/4, N/4), L_{\mathbf{B}_3}(1, 1), \dots, L_{\mathbf{B}_3}(i, j), \dots, L_{\mathbf{B}_3}(3N/8, N/8), L_{\mathbf{A}_3}(1, 1), \dots, L_{\mathbf{A}_3}(i, j), \dots, L_{\mathbf{A}_3}(N/8, N/8)]^T \quad (8)$$

with

$$\begin{aligned} L_{\mathbf{B}_2}(i, j) &= [\nabla_x(\mathbf{B}_2)_{(i,j)}, \nabla_y(\mathbf{B}_2)_{(i,j)}, \nabla_z(\mathbf{B}_2)_{(i,j)}] \\ L_{\mathbf{B}_3}(i, j) &= [\nabla_x(\mathbf{B}_3)_{(i,j)}, \nabla_y(\mathbf{B}_3)_{(i,j)}, \nabla_z(\mathbf{B}_3)_{(i,j)}] \\ L_{\mathbf{A}_3}(i, j) &= [\nabla_x(\mathbf{A}_3)_{(i,j)}, \nabla_y(\mathbf{A}_3)_{(i,j)}, \nabla_z(\mathbf{A}_3)_{(i,j)}] \end{aligned} \quad (9)$$

The functions  $\nabla_x(\cdot)$ ,  $\nabla_y(\cdot)$  and  $\nabla_z(\cdot)$  give the components along  $x$ ,  $y$  and  $z$  of the wavelet transform of the 3D gradient of the considered image.

They are calculated from

$$\begin{aligned} \nabla_x(\cdot) &= \mathcal{W}_\psi(\nabla \mathbf{I}_x) \\ \nabla_y(\cdot) &= \mathcal{W}_\psi(\nabla \mathbf{I}_y) \\ \nabla_z(\cdot) &= \mathcal{W}_\psi(\nabla \mathbf{I}_z) \end{aligned} \quad (10)$$

where  $\nabla \mathbf{I}_x = \partial \mathbf{I} / \partial x$ ,  $\nabla \mathbf{I}_y = \partial \mathbf{I} / \partial y$ ,  $\nabla \mathbf{I}_z = \partial \mathbf{I} / \partial z$  are the value of the image gradient for the point of coordinates  $(x, y, z)$  with a derivative filter as the Sobel one and  $\mathcal{W}_\psi(\cdot)$  is given by (5). Whereas the computation of  $\nabla \mathbf{I}_x$  and  $\nabla \mathbf{I}_z$  uses the OCT slice at the desired position, the computation of  $\nabla \mathbf{I}_y$  needs two parallel slices on each side of the desired slice to be calculated (see Fig. 3).

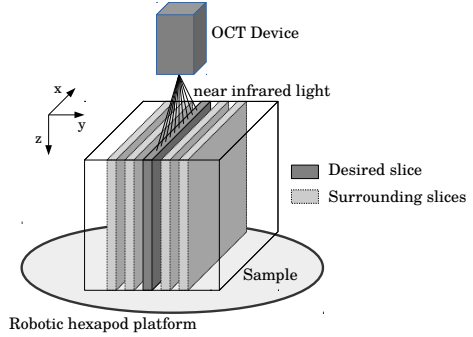


Fig. 3: Diagram showing the operation of a typical OCT imaging system

### III. EXPERIMENTAL VALIDATION

#### A. OCT Device

As mentioned in the introduction, OCT is an emerging technology for performing high-resolution, cross-sectional, tomographic imaging in real time. OCT system operation is based on a low-coherence interferometric technique, usually using near infrared light [8]. In fact, OCT images are obtained by the measure of the reverberate time delay and the intensity of the back-scattered light from the viewed object that gives data about the different scrambling layers of the tissue.

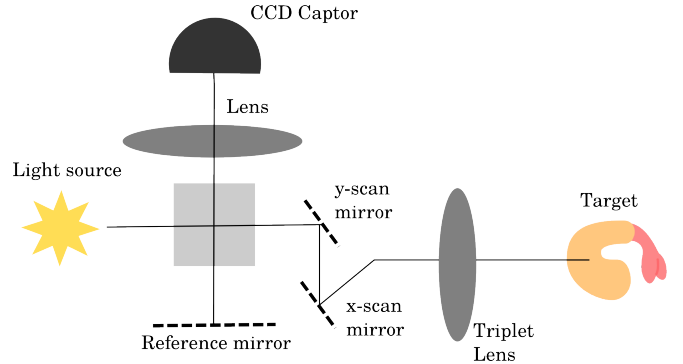


Fig. 4: Functioning of a Fourier-Domain OCT system

The OCT images (depending on the used lighting source wave-length) have micrometer resolution and millimeter penetration depth while US images offer a millimeter resolution

with several centimeters of depth. Due to the numerous advantages of the OCT, it is considered "*optical biopsy*", in various medical fields, which can be a substitute of the conventional "*physical biopsy*". The OCT imaging system used is shown in Fig. 5. It consists of a Fourier-Domain Telesto-II OCT from THORLABS. It is characterized by  $5.5\mu\text{m}$  and  $7\mu\text{m}$  axial and lateral resolutions, respectively, and provides up to 3.5mm of depth. The Telesto-II allows a maximum field-of-view of  $10 \times 1 \times 13.5\text{mm}^3$  with a maximum A-Scan (1D optical core) acquisition rate of 76 kHz.

### B. Robotic Setup

The proposed methods and materials were experimentally tested using a robot placed as a sample holder i.e., under the OCT imaging system, in a eye-to-hand configuration (Fig. 5). More precisely, the robotic platform is a 6 DOF 3PPSR parallel robot (a Space-FAB SF-3000 BS from Micros) characterized by a translation range of  $(t_x, t_y, t_z)_{\text{max}} = (50, 100, 12.7)$  (mm) and a rotation range of  $(r_x, r_y, r_z)_{\text{max}} = (10, 10, 10)$ ( $^\circ$ ), a linear resolution of  $0.2\mu\text{m}$  and an angular resolution of  $0.0005^\circ$ ; and two computers communicating with an asynchronous TCP/IP protocol, a 3.2 GHz i5 core Intel CPU (with a Linux OS) for computer vision and control law implementation and a 2.33 GHz Xeon Intel CPU (with Windows OS) for robot inner control law implementation. Note that, with a non-optimized implementation of the method (essentially scripted using MATLAB), the control loop takes approximately 100 milliseconds.

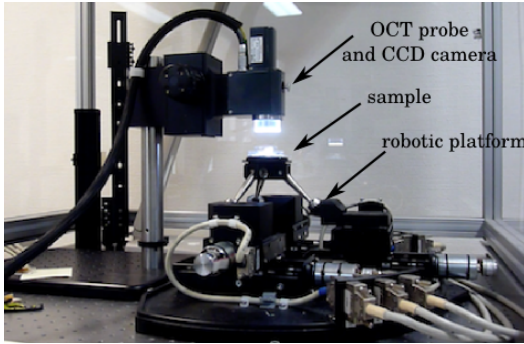


Fig. 5: Photography of the experimental setup including the 6 DOF robotic platform as well as the OCT imaging system.

During the visual servoing process, we grab also the white light image of the CCD camera rigidly attached in a co-axial configuration with the OCT light path. Therefore, the CCD images are used to essentially verify the good functioning of the developed control law. In order to validate the visual servoing controller in which the signal control inputs are the wavelet details coefficients, we choose two scenarios: using an artificial sample (a LEGO piece) and a biological sample (part of a shrimp) (Fig. 6). Two tests are presented in detail in the following; "LEGO piece test 1" corresponds to subsection III-C and "shrimp sample test 1" corresponds to the subsection III-D.

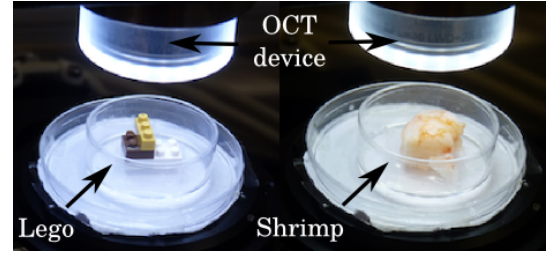


Fig. 6: Sample materials used in the experimental validation: LEGO and shrimp

### C. LEGO Experiment

The first experiment was performed using a LEGO of size  $1.6 \times 0.8 \times 1.6 \text{ cm}^3$ . First, the sample is placed at a desired position  $\mathbf{r}^*$  and moved to an arbitrary (initial) position  $\mathbf{r}$ . So, the initial error is estimated (using the robot's high resolution encoders) to be  $\Delta T_{\text{initial}} = (-0.5, -0.5, 0.3)$  (mm). The image difference between the initial OCT image  $\mathbf{I}_{\text{oct}}$  and desired one  $\mathbf{I}_{\text{oct}}^*$  is depicted in Fig. 7.(b). Note that, the displacement along the  $x$  and  $z$  axes are clearly visible on the OCT slice, contrary to the one along  $y$  axis, this is due to the LEGO structure (Fig. 7.(b)). However, this difference along  $y$  axis can be shown in Fig. 8.(b).

The first experiment achievement is shown in Fig. 7, where Fig. 7.(a) is the initial image (position), Fig. 7.(b) the initial difference image, and Fig. 7.(c) the final difference image at convergence. As can be noticed, this final difference image contains mostly noise which proves that final and desired images are completely superimposed demonstrating the good accuracy of the control law. This is also demonstrated by the CCD images sequence illustrated (Fig. 8). The final error is estimated (again using the robot's encoders) to  $\Delta T_{\text{final}} = (-0.0118, -0.0149, 0.0048)$  (mm).

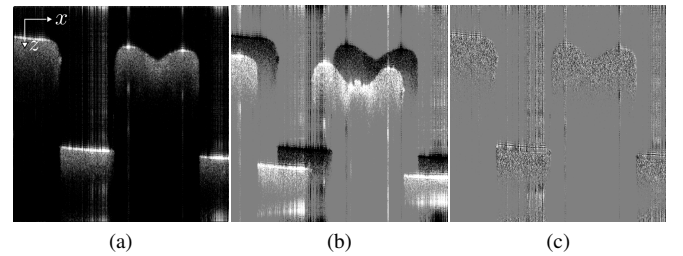


Fig. 7: Test 1 (OCT): (a) desired OCT image  $\mathbf{I}_{\text{oct}}^*$ , (b) initial difference image  $\mathbf{I}_{\text{oct}}^* - \mathbf{I}_{\text{oct}}$ , (c) the final difference image  $\mathbf{I}_{\text{oct}}^* - \mathbf{I}_{\text{oct}}$ .

The Cartesian error decay in each DOF is recorded and plotted in Fig. 9 as well as the norm of the global error  $\|\mathbf{e}\| = \|\mathbf{s} - \mathbf{s}^*\|$  which is shown in Fig. 10. One can underline that all the errors converge towards zero, in approximately 80 iterations. However, it is important to remember that the control is performed on the OCT image (with an unfavorable signal-to-noise ratio); this does not generate exponential convergence of the Cartesian errors, unlike in case of a pose-based visual servoing scheme.

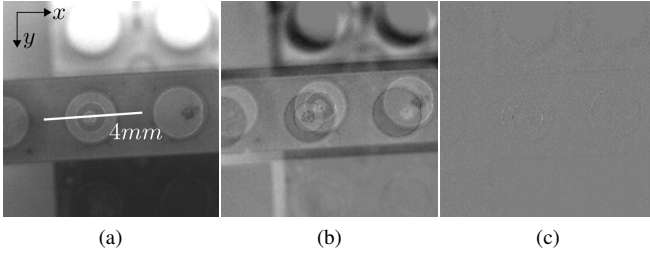


Fig. 8: Test 1 (CCD camera): (a) desired white light image  $\mathbf{I}_{cam}^*$ , (b) the initial difference image  $\mathbf{I}_{cam}^* - \mathbf{I}_{cam}$ , and (c)  $\mathbf{I}_{cam}^* - \mathbf{I}_{cam}$  acquired at convergence.

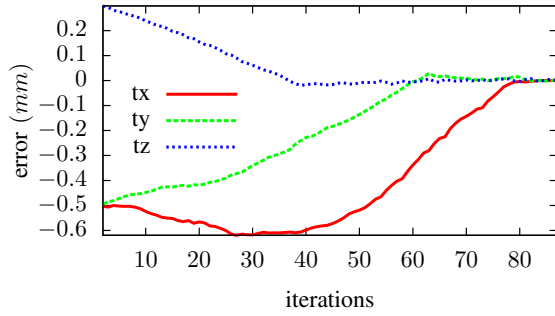


Fig. 9: LEGO: Positioning error

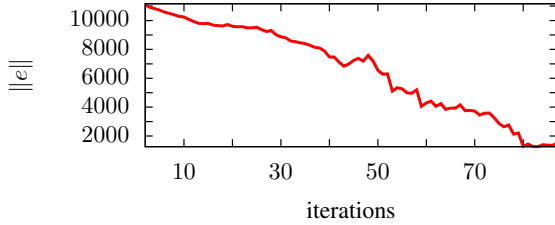


Fig. 10: LEGO: Wavelet domain norm error

#### D. Shrimp Experiment

Concerning the second experiment, the LEGO sample was replaced by a biological one, i.e., a part of a shrimp of size  $2 \times 1 \times 0.5 \text{ cm}^3$  approximately. The scenario remains the same, the operator records the desired position and then moves the sample (using the robotic platform) to an arbitrary initial position. The visual servoing task should therefore automatically return the robot to the desired position. For this experiment, the initial error is  $\Delta T_{initial} = (-0.5, -0.5, 0.3)$  (mm). The image difference between the initial OCT image and the desired one is shown in Fig. 11.(b). Contrary to the first task, the displacement along the  $y$  axis is visible (Fig. 11.(b)). As experimentally demonstrated in Fig. 11.(c), which is completely gray, the visual servoing control law converges accurately towards the desired position. The final error (i.e., at convergence) is  $\Delta T_{final} = (-0.0074, 0.0035, -0.0091)$  (mm).

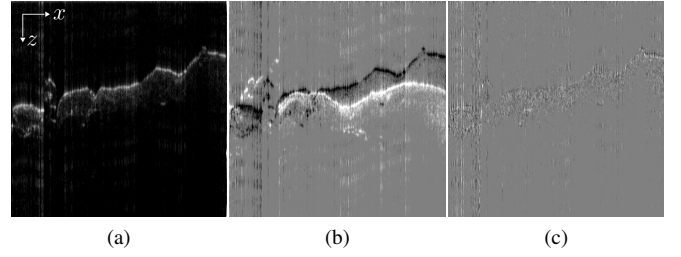


Fig. 11: Task 2 (OCT): (a) desired OCT image  $\mathbf{I}_{oct}^*$ , (b) initial difference image  $\mathbf{I}_{oct}^* - \mathbf{I}_{oct}$  and (c) the final difference image  $\mathbf{I}_{oct}^* - \mathbf{I}_{oct}$  at convergence.

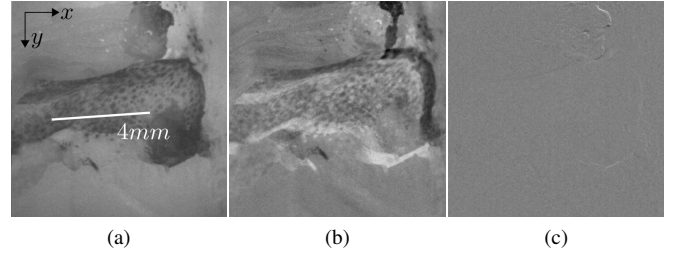


Fig. 12: Task 2 (CCD camera): (a) the white light image at the desired image  $\mathbf{I}_{cam}^*$ , (b) the initial difference image  $\mathbf{I}_{cam}^* - \mathbf{I}_{cam}$ , and (c) the final difference  $\mathbf{I}_{cam}^* - \mathbf{I}_{cam}$  acquired at convergence.

In a similar way to the first experiment, the CCD camera images were recorded and shown in Fig. 12 which confirm the accurate convergence of the OCT-based control law.

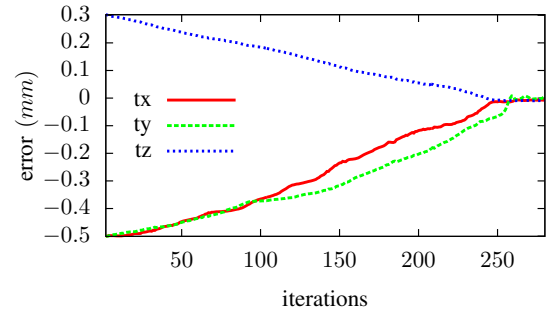


Fig. 13: Shrimp: Positioning Error

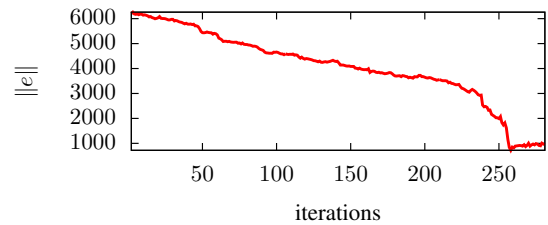


Fig. 14: Shrimp: Wavelet domain norm error



Again, the error in each DOF for the second test decreases to zero as shown in Fig. 13. As can be highlighted, all DOF converge to the desired position similarly to the first test. Also, the norm of the error  $\|e\| = \|s - s^*\|$  was recorded and plotted in Fig. 14 and its evolution towards zero confirms the good convergence behavior of the OCT-based control law.

#### E. Others Experiments

Table I presents the initial positioning errors and the final positioning ones (measured at convergence) of all the experiments done to validate our method. Two tests were performed using the LEGO piece and three using the shrimp sample. The mean positioning error considering all axes is **0.01mm** and the standard deviation (STD) considering all axes is **0.02mm**. This accuracy shows the precision of the presented framework. We can observe that the  $y$  axis (i.e. the out-of-plane axis) is slightly more than half as accurate as the others despite the fact that the interaction matrix was computed using only four slices adjacent to the slice considered as the desired image.

TABLE I: Summary of the different performed tests

Case	error (mm)	$\Delta T_x$	$\Delta T_y$	$\Delta T_z$
LEGO piece test 1	$e_{initial}$	<b>-0.5</b>	<b>-0.5</b>	<b>0.3</b>
	$e_{final}$	-0.0118	-0.0149	0.0048
LEGO piece test 2	$e_{initial}$	<b>-0.3</b>	<b>-0.6</b>	<b>0.5</b>
	$e_{final}$	-0.0002	-0.0020	-0.0111
shrimp sample test 1	$e_{initial}$	<b>-0.5</b>	<b>-0.5</b>	<b>0.3</b>
	$e_{final}$	-0.0074	0.0035	-0.0091
shrimp sample test 2	$e_{initial}$	<b>-0.5</b>	<b>-0.2</b>	<b>0.3</b>
	$e_{final}$	-0.0209	0.0677	-0.0087
shrimp sample test 3	$e_{initial}$	<b>0.2</b>	<b>-0.2</b>	<b>0.2</b>
	$e_{final}$	-0.0058	0.0054	-0.0070
	mean( $ e_{final} $ )	0.0092	0.0187	0.0081
	STD( $ e_{final} $ )	0.0077	0.0278	0.0024

#### IV. CONCLUSION AND DISCUSSION

The main objective of this paper was the study of an OCT-based direct visual servoing scheme. In fact, instead of considering geometric visual features in the interaction matrix, we took into account the wavelet coefficients. Furthermore, the OCT imaging system that is usually used only for diagnosis purposes (*in-situ optical biopsies acquisition*) was employed here also as a robotic sensor. The proposed controller was experimentally tested using a highly accurate parallel robotic structure placed under the OCT system in the role of sample holder platform. Despite of the nature of the OCT images (unfavorable signal-to-noise ratio), the proposed approach demonstrated interesting behaviors, especially in term of accuracy. Because of this unfavorable signal-to-noise ratio, taking directly the pixels of the image in the visual servoing scheme fails to provide an effective visual servoing scheme. The work presented in [9] was already using wavelet coefficients to overcome the limitations of a direct visual servoing based on pixel intensities. In the framework presented here, a larger amount of coefficients is used to compute the visual features of the control law and the interaction matrix is computed differently. With

our experimental setup, we obtained an average final error positioning of **0.01mm** which represents only 3% of the average initial displacement.

Further work will concern the extension of this investigation to a 6 DOF control law by taking into account the rotations. This haven't be done yet because of technical issues but the framework presented here contains all the theoretical notions to provide a 6 DOF direct visual servoing. More than five tests should also be proposed as well as a comparison with a pixel intensities based visual servoing. Another way of improvement will be the implementation of the whole proposed materials in a more efficient software framework, e.g., a C++ one which can improve consequently the dynamic of the controller. Finally, OCT allows to generate volumetric data. A visual servoing based on a C-scan could be a groundbreaking work for optical biopsies. Nevertheless, this process is time costly (near 1 minute for a  $1cm^3$  sample). One idea should be to combine some compressive sensing acquisition scheme (which uses often wavelet transforms) with this direct visual servoing framework.

#### ACKNOWLEDGMENT

This work is partially supported by DESWEEP project funded by the *Région de Bretagne* 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] Y. Meyer, "Ondelettes, fonctions splines et analyses graduées," *Lectures given at the University of Torino, Italy*, 1986.
- [2] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The jpeg2000 still image coding system: an overview," *IEEE Transactions on Consumer Electronics*, vol. 46, pp. 1103–1127, Nov 2000.
- [3] M. Ourak, B. Tamadazte, O. Lehmann, and N. Andreff, "Wavelets-based 6 dof visual servoing," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, pp. 3414–3419, IEEE, 2016.
- [4] L.-A. Dufloy, A. Krupa, B. Tamadazte, and N. Andreff, "Shearlet-based vs. photometric-based visual servoing for robot-assisted medical applications," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'16*, 2016.
- [5] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.
- [6] C. Collewet and E. Marchand, "Photometric visual servoing," *IEEE Transactions on Robotics*, vol. 27, no. 4, pp. 828–834, 2011.
- [7] B. Tamadazte, N. L. F. Piat, and E. Marchand, "A direct visual servoing scheme for automatic nanopositioning," *IEEE/ASME Transactions on Mechatronics*, vol. 17, pp. 728–736, Aug 2012.
- [8] D. Xu, Yong, and J. U. Kang, "Real-time compressive sensing spectral domain optical coherence tomography," *Opt. Lett.*, vol. 39, pp. 76–79, Jan 2014.
- [9] M. Ourak, B. Tamadazte, and N. Andreff, "Partitioned camera-oct based 6 dof visual servoing for automatic repetitive optical biopsies," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2337–2342, Oct 2016.
- [10] H. Yu, J. H. Shen, K. M. Joos, and N. Simaan, "Calibration and integration of b-mode optical coherence tomography for assistive control in robotic microsurgery," *IEEE/ASME Transactions on Mechatronics*, vol. 21, pp. 2613–2623, Dec 2016.
- [11] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on pure and applied mathematics*, vol. 41, no. 7, pp. 909–996, 1988.
- [12] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.