Partitioned Camera-OCT based 6 DOF Visual Servoing for Automatic Repetitive Optical Biopsies

Mouloud Ourak, *Student Member, IEEE*, Brahim Tamadazte, *Member, IEEE*
and Nicolas Andreff, *Member, IEEE*

***Abstract*— This paper addresses the design of a partitioned vision-guided scheme for repetitive optical biopsies. More precisely, our approach uses two image modalities to perform 6 degrees of freedom (DOF) positioning task. The development aims to partition the control into 3 DOF controlled by the B-scan images acquired with an optical coherence tomography (OCT) system and the remaining 3 DOF controlled by the white light images provided by a CCD camera. Moreover, for the control and instead of conventional visual features (e.g., points, lines, moments, etc.) extracted using algorithms combined with visual tracking approaches, our visual servoing method uses the multiresolution wavelet coefficients. The developed method was experimentally validated using a parallel kinematic structure equipped with a Telesto-II OCT benchtop. The validation task consisted of an automatic spatial repositioning of the robotic structure to precisely retrieve the position of an initial optical biopsy. Several tests are achieved, which clearly demonstrate the reliability of the proposed controller.**

I. INTRODUCTION

A *biopsy* procedure is the sampling of suspicious cells or tissues for *ex-situ* examination under an optical microscope by a pathologist or analyzed chemically. The tissue removal can be achieved with different techniques, e.g., puncturing needle, endoscopic system, or during surgical interventions. However, the removal of tissues is a complicated procedure which can be even dangerous (e.g., risk of malignant degeneration) or medically contraindicated in some cases (e.g., inflammatory tissue disorganization). During the last two decades, alternative non-invasive techniques have been carried out, commonly known as optical biopsies [1]. The aim of those techniques is to reduce the need for surgical tissue removal and replace *ex-situ* histological examination with *in-situ* real-time optical measurements. To achieve that, optical biopsy technologies have been developed thanks to recent advances in fiber-optics, light sources and detectors, such as fluorescence endoscopy, confocal microendoscopy and optical coherence tomography [2].

OCT has the advantages compared to conventional histopathology as well as to the other optical biopsies to acquire an image with better resolution than ultrasound (US) images, and with large depth compared to the confocal microscopy. This makes OCT as the gold standard for the diagnosis of many diseases [3]. Also, OCT images can be used during surgical procedures as a visual sensor in aim to guide the surgical instruments.

OCT is based on the principle of low coherence interferometry providing very good lateral and axial resolutions depending on the design (between 4 to 10 μm and 5 to 15 μm, respectively). It also has the ability to reach rather high penetration depths, between 1 to 5 mm [4]. OCT was firstly applied in ophthalmology due to the transparent nature of eyes, their minimal scattering and high light penetration. Recently, imaging of non-transparent tissues has been achieved using longer wavelengths, near infrared, where optical scattering is reduced. Therefore, OCT is now used in several fields, such as: cardiology, gastroenterology, or urology. Although many OCT endoscopic probes have been reported in the literature [5], [6], none of these systems use OCT images to servo the endoscope system. Indeed, vision-based control (known as *visual servoing*) is a very useful approach for robotic control, especially when flexibility, accuracy and repeatability are required [7], [8].

Visual servoing has been widely investigated for different robotic applications as well as using various image modalities. In the late 2000s, a new type of visual control has emerged known as *visual servoing set free from image processing*. It consists of removing completely the visual information extraction and matching as well as the visual tracking. These new paradigms use pure image signal such as pixel intensity [9], [10], mutual information [11], Fourier transform [12], [13], etc. These methods demonstrate more accuracy and robustness due to the redundancy of the information, comparing to the traditional visual servoing techniques.

In the literature, visual servoing is applied for various imaging modalities in the case of optical biopsy such as: ultrasounds (US) [14], confocal microscopy [15], OCT [16], etc. In opposition to the previous methods where only one modality is used during the process, this paper focuses on the development of a new partitioned visual servoing approach as the vision/force visual servoing scheme [17]. This partitioned approach combines 3 DOF controlled by the OCT B-scan and the remaining 3 DOF controlled with the CCD camera mounted on the same benchtop. This approach is chosen due to the very long time volumetric OCT image acquisition. Furthermore, the final goal is the repetitive optical biopsy repositioning in SE(3) (i.e., 3D) for the diagnostic monitoring.For instance, the physician can retrieve the excat position of a previous biopsy using an OCT-based visual servoing. In this work, the control law

This work is conducted with a financial support from the project NEMRO (ANR-14-CE17-0013-001) funded by the Agence Nationale de la Recherche, France. It is also performed in the framework of the Labex ACTION (ANR-11-LABX-0001-01)

The authors are with FEMTO-ST institute, AS2M department, Univ. Bourgogne Franche-Comté/CNRS/ENSMM, 25000 Besanon, France. brahim.tamadazte@femto-st.fr

follows the method proposed in [18] for wavelet-based visual servoing, but unlike the paper [16], where we control only the positioning task of the OCT B-scan, the paper extends the control to the volumetric positioning thanks to the partitioned (OCT/CCD camera) approach. The proposed controller is experimentally validated on a set-up composed by a benchtop (i.e., OCT probe + CCD camera) placed over a robotic platform (sample holder). The validation tests consist of an *in-vitro* SE(3) positioning task representing an automatic retrieval of the 3D pose of an initial optical biopsy.

The paper is organized as follows: Section II defines the positioning task studied in this paper as well as the introduction of the proposed visual servoing controller. Section III describes both experimental materials used to validate the control law and the obtained results, while Section IV emphasizes the key points of the paper with a conclusion and perspectives.

II. METHODOLOGY

A. Positioning Task Problem

The typical problem consists of an SE(3) positioning task as can be seen in Fig. 1 where the B-scan OCT image $\boldsymbol{I}_{oct}$ is perpendicular to the white light image $\boldsymbol{I}_{cam}$. To perform an optical biopsy, the operator defines, on the white light image, a line to be scanned by the OCT laser (Fig. 1). To monitor the evolution (texture, shape, etc.) of a pathological tissue, the clinician/biologist needs to perform new optical biopsy at the same initial position. Due to the optical biopsy micrometric resolution, vision-based control is more than necessary in order to be able to retrieve the initial position. Indeed, for this work, we propose partitioned (i.e., OCT and CCD camera) wavelet-based visual servoing to perform the position task on biological sample (e.g., a fly).

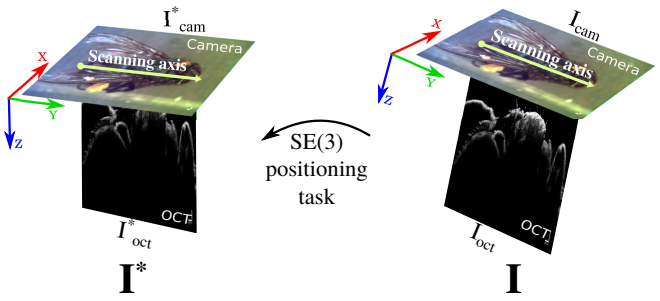

Fig. 1. Representation of SE(3) positioning task used to validate experimentally our controller.

B. Wavelet-based Visual Servoing

Let us consider an image $\boldsymbol{I}(x,y)$ represented in time domain which can be also represented in the frequency domain using Fourier transform. Otherwise, using wavelet transform allows a time-frequency representation of the signal, better suited for studying subtle details, and providing higher precision when modeling a signal. In other words, Fourier transform provides a frequency analysis, but does not allow the temporal location of sharp changes of the signal. A wavelet transform can be defined as an inner product between a 2D signal $\mathcal{F}(x,y)$ and a wavelet function $\mathcal{G}(x,y)$.

$$\left\langle \mathcal{F}(x,y), \mathcal{G}(x,y) \right\rangle = \int \int_{-\infty}^{+\infty} \mathcal{F}(x,y)\ \mathcal{G}(x,y)\ dxdy \quad (1)$$

Different wavelet transforms were reported in the literature as continuous and discrete [19], multiresolution [20], etc. However, this paper is only based on the wavelet multiresolution analysis. Indeed, wavelet coefficients represent very interesting visual features for carrying out an accurate and robust visual servoing controller. This is true because using these coefficients allow to select redundant and pertinent (automatic filtering of high frequencies) visual features from the image. More precisely, our approach uses the approximative coefficients of the wavelet transform to build the down sample image $\boldsymbol{I}_{2^j}$ and the detail coefficients to compute the interaction matrix $\mathbf{L}_{\mathbf{w}(2^j)}$ that links the coefficients time-variation of resolution j to the camera/robot instantaneous velocities $\mathbf{v}_c$.

Based on the optical flow constraint equation (OFCE) applied to wavelet coefficients time-variation [21], and our preliminary work [18], it is possible to write

$$\dot{\mathcal{I}}_{2^j} = \mathbf{L}_{\mathbf{w}(2^j)}\mathbf{v} \quad (2)$$

where

$$\mathcal{I}_{2^j} = vec(\boldsymbol{I}_{2^j}) = \left(\boldsymbol{I}_{(2^j)(1,1)}, \boldsymbol{I}_{(2^j)(1,2)}, \cdots, \boldsymbol{I}_{(2^j)(M,N)}\right)^\top \quad (3)$$

is the visual signal, and

$$\mathbf{L}_{\mathbf{w}(2^j)} = \left[\mathbf{L}_{\mathbf{w}(2^j)(1,1)}, \mathbf{L}_{\mathbf{w}(2^j)(1,2)}, \cdots, \mathbf{L}_{\mathbf{w}(2^j)(M,N)}\right]^\top \quad (4)$$

is the compound multiresolution interaction matrix that links the variation of $\mathcal{I}_{2^j}$ to the camera velocity $\mathbf{v}$. It is built by stacking, each multiresolution interaction matrix $\mathbf{L}_{\mathbf{w}(2^j)(x,y)}$ for each pixel $\boldsymbol{p} = (x,y)^\top$

$$\mathbf{L}_{\mathbf{w}(2^j)(x,y)} = -\begin{bmatrix} \mathbf{g}^H_{(2^j)(x,y)} & \mathbf{g}^V_{(2^j)(x,y)} \end{bmatrix} \mathbf{L}_{2D(x,y)} \quad (5)$$

where

$$\mathbf{L}_{2D(x,y)} = \begin{pmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \\ 0 & -1/Z & y/Z & -(1+y^2) & -xy & -x \end{pmatrix} \quad (6)$$

is the 2D point interaction matrix reported in [7], where Z is the depth, and

$$\mathbf{g}^H_{(2^j)} \triangleq \left\langle \boldsymbol{I}_{(2^{j+1})}, \frac{\partial \Gamma^H}{\partial x} \right\rangle \quad (7)$$

$$\mathbf{g}^V_{(2^j)} \triangleq \left\langle \boldsymbol{I}_{(2^{j+1})}, \frac{\partial \Gamma^V}{\partial y} \right\rangle \quad (8)$$

$$\boldsymbol{I}_{(2^j)} = \left\langle \boldsymbol{I}_{(2^{j+1})}, \Gamma^0 \right\rangle \quad (9)$$

where $\mathbf{g}^H_{(2^j)}$ and $\mathbf{g}^V_{(2^j)}$ are the horizontal and vertical wavelet coefficients, respectively. The image $\boldsymbol{I}_{(2^j)}$ is the new image with resolution j.
Γ^0, $\frac{\partial \Gamma^V}{\partial y}$ and $\frac{\partial \Gamma^H}{\partial x}$ are the functions used to convert image

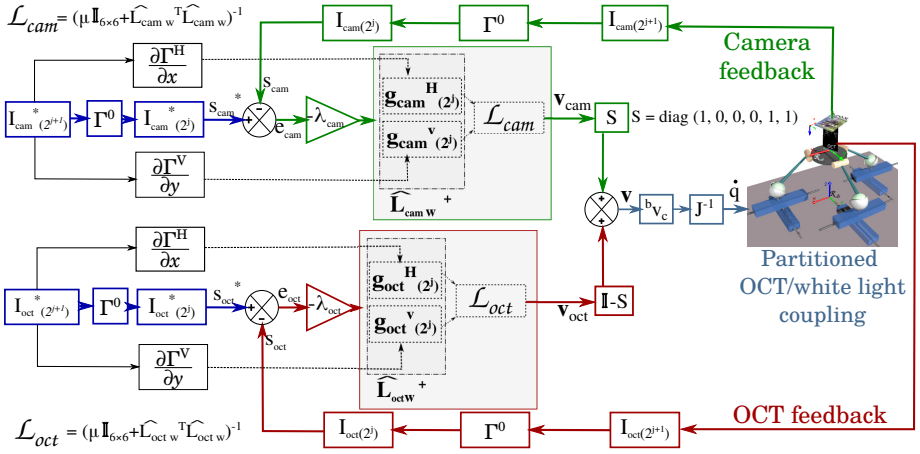

Fig. 2. Global view of the proposed partitioned visual servoing block diagram.

$\boldsymbol{I}_{(2^{j+1})}$ to image $\boldsymbol{I}_{(2^j)}$ and generate detail coefficients (for more information, please refer to [18]).

Finally, in order to ensure a stable and smooth control, we use the *Levenberg-Marquardt* method [22]. Otherwise, the robot spatial velocity is expressed in the camera frame $\mathcal{R}_c$ as

$$\mathbf{v} = -\lambda \left(\mu\, \mathbb{I}_{6\times 6} + \widehat{\mathbf{L}_{\mathbf{w}(2^j)}}^{\top} \widehat{\mathbf{L}_{\mathbf{w}(2^j)}}\right)^{-1} \widehat{\mathbf{L}_{\mathbf{w}(2^j)}}^{+} \left(\mathcal{I}_{2^j} - \mathcal{I}^*_{2^j}\right) \tag{10}$$

where λ and μ are the controller gains, and $\mathbb{I}_{6\times 6}$ is an identity matrix.

C. Control Law

The partition of our controller is inspired by the well-known vision/force coupling approaches reported in the literature [17]. In our work, instead of using force and vision sensors, we use OCT and CCD camera as a multimodal imaging system. Thereby, the controller is divided two separate paths, as shown in Fig. 2. Indeed, white light images were used to control 3 DOF (i.e., x, R_y, and R_z) when OCT images were used for the other 3 DOF (i.e., y, z, and R_x) (where R_x, R_y and R_z are the rotations carried by x, y and z respectively). However, the orthogonality of the two controllers is required to avoid the conflict during the positioning task. Furthermore, a selection matrix $\mathbf{S}$ and its complementary $\mathbb{I} - \mathbf{S}$ are introduced in the control loop (Fig. 2) to avoid any conflict problem. Thus, both image modalities are combined as follows:

$$\mathbf{v} = \mathbf{S}\mathbf{v}_{cam} + \big(\mathbb{I} - \mathbf{S}\big)\mathbf{v}_{oct} \tag{11}$$

where $\mathbf{v}_{cam} = (v_x, \omega_y, \omega_z)^{\top}$ are the output camera velocities, and $\mathbf{v}_{oct} = (v_y, v_z, \omega_x)^{\top}$ are those of the OCT system, $\mathbf{v}$ is the combined velocities send to the robot motors, and the selection matrix is set to $\mathbf{S} = diag\big(1, 0, 0, 0, 1, 1\big)$.

In the experimental scenario, both OCT system and CCD camera are mounted in an *eye-to-hand* configuration system. The CCD images are traditionally used to define the OCT working distance (i.e., sharp image means the perfect OCT working distance) as well as the bounding box to be scanned by the OCT laser. This combination provides an original configuration to design a hybrid vision-based control law.

III. EXPERIMENTAL VALIDATION

A. Experimental Set-up

To evaluate the performances of the proposed controller, an experimental set-up was built. It includes two main parts:

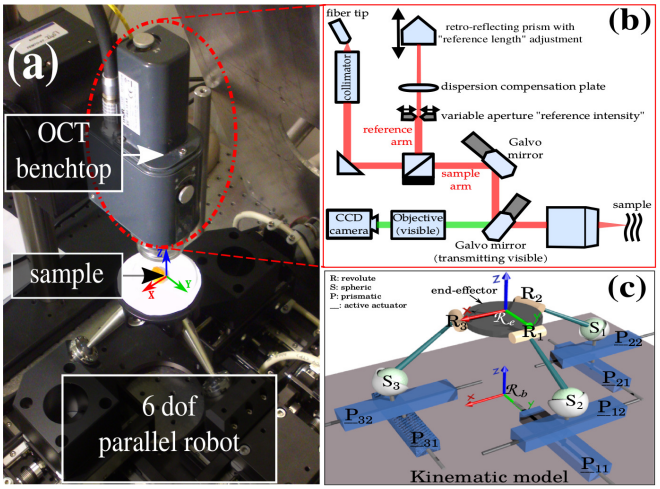

Fig. 3. (a) photography of the experimental set-up, (b) inner architecture of the multimodal imaging system, and (c) the kinematic structure of the parallel robotic system.

- an OCT imaging system (a Telesto-II 1325 nm spectral domain) from ThorLabs[1] (Fig. 3(a)) providing 1D depth (A-scan), 2D cross-sectional (B-scan) or 3D volumetric (3D scan) images with micrometric resolution (5.5 μm and 7 μm for axial and lateral resolutions, respectively) and millimeters depth (3.5 mm of penetration). The Telesto-II allows a maximum field-of-view of $10 \times 10 \times 3.5$ mm^3 with a maximum A-Scan line rate of 76 kHz. The studied sample is also viewed through a

[1] www.thorlabs.de

CCD camera (640 × 480 pixels of resolution) placed in the same axis as the OCT one. Both are positioned, in an *eye-to-hand* configuration, on the top of the robotic-based sample holder (Fig. 3(a)).

- a 6 DOF parallel robotic structure (Fig. 3(c)) which consists of **3PPSR** robot SpaceFAB SF-3000 BS from Micos[2]. This robot is characterized by the following features: translation range $(\Delta T_x, \Delta T_y, \Delta T_z)^\top_{max}$ = (50, 100, 12.7)$^\top$ (mm) and rotation range $(\Delta R_x, \Delta R_y, \Delta R_z)^\top_{max}$ = (10, 10, 10)$^\top$ (deg), a linear resolution of 0.2 μm (repeatability of $\pm$0.5 μm) and an angular resolution of 0.0005 deg (repeatability of $\pm$0.0011 deg).

In addition, two computers equipped the testbench: a 3.2 GHz i5 core Intel CPU with a Linux OS dedicated to the computer vision and control algorithms and a 2.33 GHz Xeon Intel CPU with Windows 7 OS is used for the robot inner control laws (inner PID loop, static and differential kinematic models). The computers communicate between them with an asynchronous TCP/IP protocol.

B. Validation Scenarios

In this section, we address the different scenarios used to demonstrate the functioning of the proposed vision-based controller.

1) The first scenario consists of 3 DOF positioning task using only the translation stages i.e., T_x, T_y and T_z of the robotic platform in order to: firstly evaluate the potential for convergence, secondly the precision rate of our controller without rotation.
2) The second one is the extension of the first scenario to an SE(3) positioning task including the rotations R_x, R_y, and R_z.
3) The third one consists of several SE(3) positioning tests under different conditions of use in aim to judge the repeatability of the controller.

The SE(3) positioning task is performed as follows: the operator defines, using the OCT image, an optical biopsy at a desired pose $\mathbf{r}^*$. Afterwards, from any initial position $\mathbf{r}$, the robot must move back towards the desired position. During the task, the error $\mathbf{e}_i$ (i represents the different robot DOF), the norm of the error $\|\mathbf{e}\|$ as well as the robot joint velocities $\dot{\mathbf{q}}_i$ are recorded and plotted.

C. Experimental Results

1) Scenario 1: Fig. 4 shows some images captured during the 3 DOF positioning task achievement. First, the left column represents the CCD camera image sequence i.e., Fig. 4(1-a) the desired image, Fig. 4(1-b) the initial one, Fig. 4(1-c) the initial difference between the initial and desired images, and Fig. 4(1-d) the final error. Second, the right column depicts the B-scan images grabbed during the positioning task. Likewise, Fig. 4(2-a) is the desired OCT image, Fig. 4(2-b) the initial one, Fig. 4(2-c) is the initial error and Fig. 4(2-d) is the final error. As can be seen on

[2]www.pimicos.com

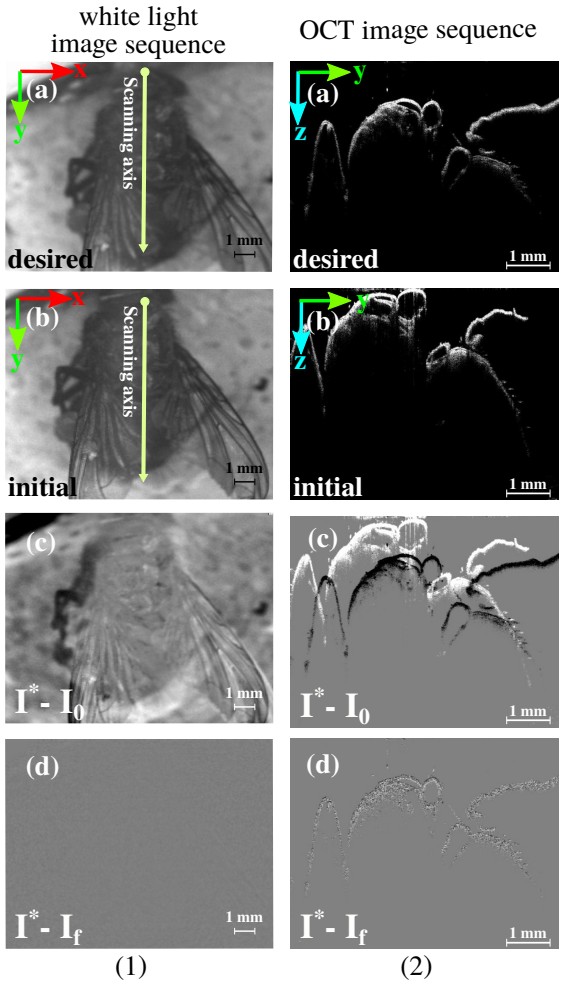

Fig. 4. Image sequence captured during the translation positioning task (i.e., 3 DOF): (a) desired image, (b) initial image, (c) initial image difference, and (d) final difference when the controller reaches the target.

the final image error (totally gray), the controller reaches accurately $\mathbf{r}^*$.

Fig. 5(a) depicts the Cartesian error $\mathbf{e}_i$ decay in each translation stage as well as the decay of the norm of the error (Fig. 5(b)). Thereby, it can be underlined that the different components converge, simultaneously, to their respective desired values. The presence of some oscillations on the actuators controlled from the OCT B-scan images is due to the unfavorable signal to noise ratio.

TABLE I

NUMERICAL VALUES OF TRANSLATION POSITIONING TASK ((T_i, e_i) IN mm) AND (R_i, e_{R_i}) IN deg).

stages	T_x	T_y	T_z	R_x	R_y	R_z
$\mathbf{r}$	0.80	0.80	75.00	x	x	x
$\mathbf{r}^*$	0.00	0.00	74.40	x	x	x
$\mathbf{r}_f$	-0.012	0.014	74.381	x	x	x
	$\mathbf{e}_x$	$\mathbf{e}_y$	$\mathbf{e}_z$	$\mathbf{e}_{R_x}$	$\mathbf{e}_{R_y}$	$\mathbf{e}_{R_z}$
$\mathbf{e}_0$	**0.80**	**0.80**	**0.60**	**x**	**x**	**x**
$\mathbf{e}_f$	**0.010**	**0.011**	**0.019**	**x**	**x**	**x**

TABLE I summarizes the numerical values of the initial,

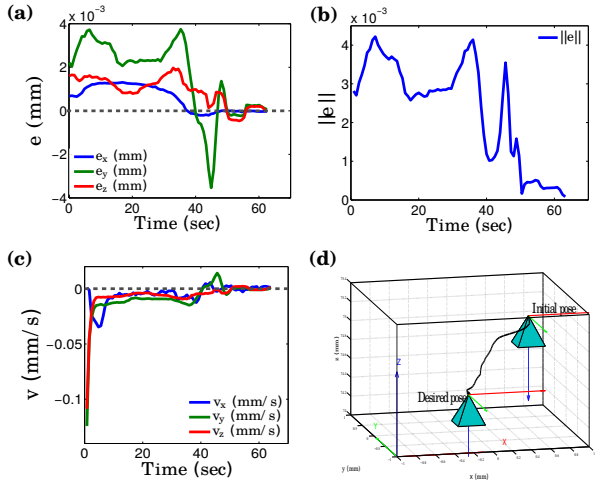

Fig. 5. Curves of the 3 DOF positioning task: (a) pose error *vs.* time, (b) norm of the error $\|\mathbf{e}\|$ *vs.* time, (c) translational velocities $(v_x, v_y, v_z)^\top$ (in mm/s), and (d) 3D robot trajectory.

desired, final positions, as well as the final error measured when the controller reaches the desired position using the high resolution robot encoders supplied by the robot software. For instance, the initial error is estimated to $\overline{\mathbf{e}}_{0_{trans}} = \mathbf{0.733}$ mm when the final error is $\overline{\mathbf{e}}_{f_{trans}} = \mathbf{0.0133}$ mm which represents only $1,81\%$ of the initial error.

2) Scenario 2: In the second experiment, we extended the previous test to the whole robotic stages, i.e., 6 DOF. Fig. 6 presents an image sequence (left column depicts the white light images and right column those corresponding to OCT B-scan images) grabbed during the SE(3) positioning task. Fig. 6 (1-a) to (1-d) show the initial $\mathbf{r}$, desired $\mathbf{r}^*$, initial difference image and final error, respectively. As can see, the controller reaches accurately the desired position. Similar to the first validation test, we recorded the error $\mathbf{e}_i$ decay in each DOF, the norm of the error $\|\mathbf{e}\|$ as well as the 6 joint robot velocities $\dot{\mathbf{q}}_i$. Therefore, Fig. 7(a)-(b) illustrates both the Cartesian error decay in each robotic DOF and norm of the error. It can be underlined that despite the presence of some small oscillations, the robot reaches accurately the desired position. Otherwise, Fig. 7(c)-(d) depict the evolution of the joint velocities sent to the robot during the positioning task performing. The camera 3D trajectory recorded during the positioning task is plotted in Fig. 7(e).

TABLE II

NUMERICAL VALUES PERFORMED FOR AN $SE(3)$ POSITIONING TASKS $((T_i, e_i)$ IN mm AND (R_i, e_{R_i}) IN $deg)$.

stages	T_x	T_y	T_z	R_x	R_y	R_z
$\mathbf{r}$	0.953	0.804	75.1	2.497	2.001	3.002
$\mathbf{r}^*$	0.00	0.00	74.4	0.00	0.00	0.000
$\mathbf{r}_f$	-0.006	0.023	74.371	0.485	0.176	-0.032
	$\mathbf{e}_x$	$\mathbf{e}_y$	$\mathbf{e}_z$	$\mathbf{e}_{R_x}$	$\mathbf{e}_{R_y}$	$\mathbf{e}_{R_z}$
$\mathbf{e}_0$	**0.953**	**0.804**	**0.70**	**2.497**	**2.001**	**3.002**
$\mathbf{e}_f$	**0.006**	**0.023**	**0.029**	**0.437**	**0.181**	**0.120**

TABLE II summarizes numerical values of the initial $\mathbf{r}_0$, desired $\mathbf{r}_d$ and final $\mathbf{r}_f$ in the Cartesian space. The mean

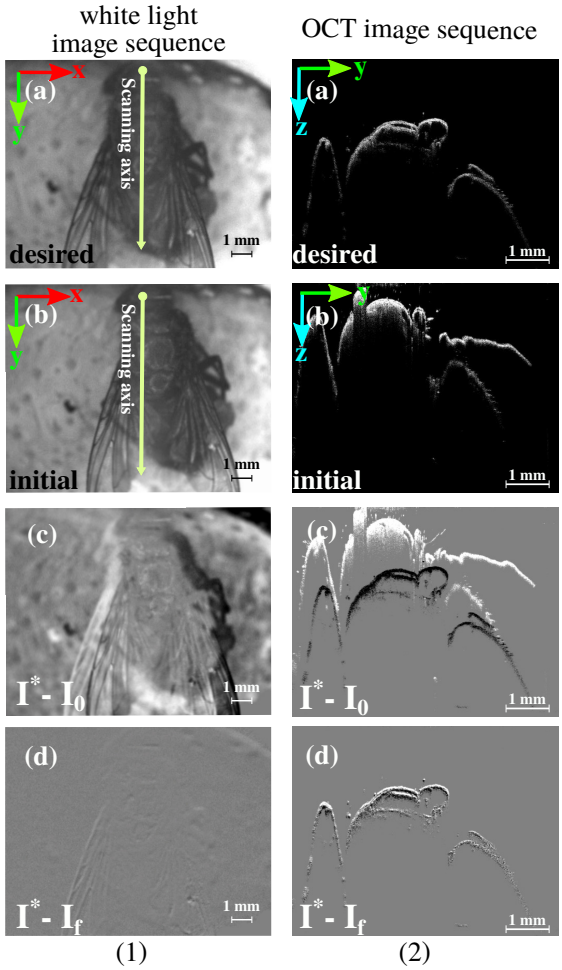

Fig. 6. Image sequence captured during $SE(3)$ (i.e., 6 DOF) positioning task: (a) desired image, (b) initial image, (c) initial image difference, and (d) final difference when the controller reaches the target.

translation (x, y and z axes) error is $\overline{\mathbf{e}}_{f_{trans}} = \mathbf{0.019}$ mm, where the mean rotation (R_x, R_y and R_z axes) error is $\overline{\mathbf{e}}_{f_{rot}} = \mathbf{0.246}$ deg.

3) Scenario 3: The SE(3) positioning task was repeated several times using different initial positions in aim to avoid biased results. It was experimentally demonstrated that in each test, the controller reaches successfully the desired position. TABLE III gives some samples of positioning tasks performed by the proposed controller. The mean Cartesian pose error in translation is estimated to $\overline{\mathbf{e}}_{f_{trans}} = \mathbf{0.048}$ mm (with a standard deviation (STD) of **0.042** mm) when the rotation error is $\overline{\mathbf{e}}_{f_{rot}} = \mathbf{0.417}$ deg (with a STD of **0.393** deg).

IV. CONCLUSION AND FUTURE WORK

In this paper the efficiency of the developed partitioned (OCT/CCD camera) wavelet-based 6 DOF visual servoing controller was demonstrated. The latter uses two image modalities i.e., CCD camera and OCT imaging system. The designed interaction matrix links the time-variation of the wavelet coefficients to the robot spatial velocity. The partitioned controller was experimentally validated using a 6 DOF

parallel robotic kinematic structure placed under the imaging systems in an *eye-to-hand* configuration. The ground-truth validations were consisted of several SE(3) positioning tasks which were successfully performed. The obtained results have demonstrated the efficiency of the controller in terms of accuracy (some tens of micrometers and few hundred of millidegree in the translation and rotation Cartesian space, respectively), convergence, and repeatability.

In future work, we will aim to perform more realistic tests using biological tissues include time variations of the sample (biological degradation, deformation, inflammation, etc.) as well as improving the controller frame rate especially for the wavelet computation. Also, it will concern the use of only OCT volumetric images as signal inputs instead B-scan images on the vision-based controller.

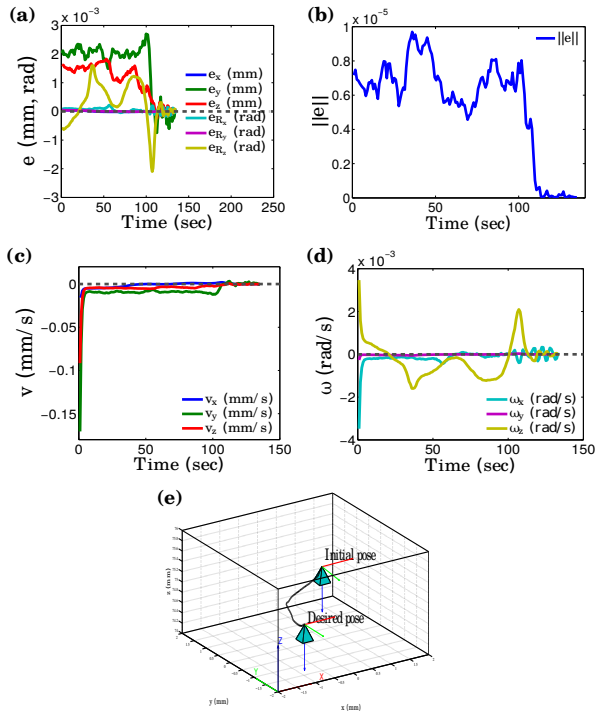

Fig. 7. Curves of the 6 DOF positioning task: (a) pose error *vs.* time, (b) norm of the error $\|\mathbf{e}\|$ *vs.* time, (c) translational velocities $(v_x, v_y, v_z)^\top$ (in mm/s), and (d) rotational velocities $(\omega_x, \omega_y, \omega_z)^\top$ (rad/s) (e) 3D robot trajectory.

TABLE III

NUMERICAL VALUES OF REPETITIVE $SE(3)$ POSITIONING TASKS $((T_i, \mathbf{e}_i)$ IN mm AND $(R_i, \mathbf{e}_{R_i})$ IN $deg)$.

stages	T_x	T_y	T_z	R_x	R_y	R_z
$\mathbf{r}$	0.80	0.80	75.00	2.00	2.00	3.00
$\mathbf{r}^*$	0.00	0.00	74.40	0.00	0.00	0.00
$\mathbf{r}_f(1)$	-0.011	0.045	74.33	1.016	0.233	-0.025
$\mathbf{r}_f(2)$	-0.035	-0.076	74.306	0.197	0.581	-0.078
$\mathbf{r}_f(3)$	-0.185	-0.038	74.382	0.793	-0.100	-1.672
$\mathbf{r}_f(4)$	-0.021	-0.006	74.371	0.459	0.278	-0.053
$\mathbf{r}_f(5)$	-0.038	0.004	74.375	-0.183	0.471	-0.087
$\mathbf{r}_f(6)$	-0.039	0.071	74.324	0.680	0.527	-0.077
	$\mathbf{e}_x$	$\mathbf{e}_y$	$\mathbf{e}_z$	$\mathbf{e}_{R_x}$	$\mathbf{e}_{R_y}$	$\mathbf{e}_{R_z}$
$\mathbf{e}_0$	0.80	0.80	0.60	2.00	2.00	3.00
$\mathbf{e}_f(1)$	0.011	0.045	0.070	1.016	0.233	0.025
$\mathbf{e}_f(2)$	0.035	0.076	0.094	0.197	0.581	0.078
$\mathbf{e}_f(3)$	0.185	0.038	0.018	0.793	0.1	1.672
$\mathbf{e}_f(4)$	0.021	0.006	0.029	0.459	0.278	0.053
$\mathbf{e}_f(5)$	0.038	0.004	0.025	0.183	0.471	0.087
$\mathbf{e}_f(6)$	0.039	0.071	0.076	0.68	0.527	0.077
$\overline{\mathbf{e}_f}$	**0.055**	**0.040**	**0.052**	**0.555**	**0.365**	**0.332**
STD $(\mathbf{e}_f)$	**0.065**	**0.031**	**0.032**	**0.335**	**0.189**	**0.657**

REFERENCES

[1] I. J. Bigio and J. R. Mourant, "Optical biopsy," *Encyclopedia of Optical Engineering*, vol. 1577, p. 1593, 2003.

[2] T. D. Wang and J. Van Dam, "Optical biopsy: a new frontier in endoscopic detection and diagnosis," *Clinical Gastroenterology and Hepatology*, vol. 2, no. 9, pp. 744–753, 2004.

[3] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser, "Optical coherence tomography-principles and applications," *Reports on progress in physics*, vol. 66, no. 2, p. 239, 2003.

[4] D. Huang, E. A. Swanson, C. P. Lin *et al.*, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.

[5] G. J. Tearney, M. E. Brezinski, B. E. Bouma, S. A. Boppart *et al.*, "In vivo endoscopic optical biopsy with optical coherence tomography," *Science*, vol. 276, no. 5321, pp. 2037–2039, 1997.

[6] M. J. Gora, J. S. Sauk, R. W. Carruth *et al.*, "Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure," *Nature medicine*, vol. 19, no. 2, pp. 238–240, 2013.

[7] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Rob. and Auto. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.

[8] S. Hutchinson, G. Hager *et al.*, "A tutorial on visual servo control," *IEEE Trans. on Rob. and Auto.*, vol. 12, no. 5, pp. 651–670, 1996.

[9] C. Collewet and E. Marchand, "Photometric visual servoing," *IEEE Trans. on Rob.*, vol. 27, no. 4, pp. 828–834, 2011.

[10] B. Tamadazte, G. Duceux, N.-F. Piat, and E. Marchand, "Highly precise micropositioning task using a direct visual servoing scheme," in *IEEE Int. Conf. on Rob. and Auto.*, pp. 5689–5694, 2011.

[11] A. Dame and E. Marchand, "Mutual information-based visual servoing," *IEEE Trans. on Rob.*, vol. 27, no. 5, pp. 958–969, 2011.

[12] N. Marturi, B. Tamadazte, S. Dembélé, and N. Piat, "Visual servoing schemes for automatic nanopositioning under scanning electron microscope," in *IEEE Int. Conf. on Rob. and Auto.*, pp. 981–986, 2014.

[13] N. Marturi, B. Tamadazte, S. Dembélé, and N. Piat, "Image-guided nanopositioning scheme for SEM," in *IEEE Trans. on Automation Science and Engineering*, DOI: 10.1109/TASE.2016.2580660, 2016.

[14] 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.

[15] B. Rosa, M. S. Erden, T. Vercauteren, B. Herman *et al.*, "Building large mosaics of confocal edomicroscopic images using visual servoing," *IEEE Trans. on Biomedical Eng.*, vol. 60, no. 4, pp. 1041–1049, 2013.

[16] M. Ourak, A. De Simone, B. Tamadazte, G. J. Laurent *et al.*, "Automated in-plane oct-system positioning towards repetitive optical biopsies," *IEEE Int. Conf. on Rob. and Auto.*, p. 6, 2016. (in press)

[17] Y. Mezouar, M. Prats, and P. Martinet, "External hybrid vision/force control," in *Int. Conf. on Advanced Rob.*, 2007.

[18] M. Ourak, B. Tamadazte, O. Lehmann, and N. Andreff, "Wavelets-based 6 dof visual servoing," *IEEE Int. Conf. on Rob. and Auto.*, p. 6, 2016. (in press)

[19] Y. Meyer, "Wavelets-algorithms and applications," *Society for Industrial and Applied Mathematics*, vol. 1, p. 142, 1993.

[20] S. Mallat, "A theory for multiresolution signal decomposition : The wavelet representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[21] C. Bernard, "Wavelets and ill posed problems: optic flow and scattered data interpolation," Ph.D. dissertation, Centre de mathématiques appliquées, Ecole Polytechnique, France, 1999.

[22] M. Lourakis, "A brief description of the levenberg-marquardt algorithm implemented by levmar," *Foundation of Research and Technology*, no. 4, pp. 1–6, 2005.