

Hybrid Automatic Visual Servoing Scheme using Defocus Information for 6-DoF Micropositioning

Le Cui, Eric Marchand, Sinan Haliyo, Stéphane Régnier

► **To cite this version:**

Le Cui, Eric Marchand, Sinan Haliyo, Stéphane Régnier. Hybrid Automatic Visual Servoing Scheme using Defocus Information for 6-DoF Micropositioning. IEEE Int. Conf. on Robotics and Automation, ICRA'15, May 2015, Seattle, United States. 2015. <hal-01120734>

HAL Id: hal-01120734

<https://hal.inria.fr/hal-01120734>

Submitted on 26 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hybrid Automatic Visual Servoing Scheme using Defocus Information for 6-DoF Micropositioning

Le Cui¹, Eric Marchand¹, Sinan Haliyo² and Stéphane Régnier²

Abstract—Direct photometric visual servoing uses only the pure image information as a visual feature, instead of using classic geometric features such as points or lines. It was demonstrated efficiently in 6 degrees of freedom (DoF) positioning. However, in micro-scale, using only image intensity as a visual feature performs unsatisfactorily in cases where the photometric variation is low, such as motions along vision sensor’s focal axis under a high magnification. In order to improve the performance and accuracy in those cases, an approach using hybrid visual features is proposed in this paper. Image gradient is employed as a visual feature on z axis while image intensity is used on the other 5 DoFs to control the motion. A 6-DoF micro-positioning task is accomplished by this hybrid visual servoing scheme. The experimental results obtained on a parallel positioning micro-stage under a digital microscope show the robustness and efficiency of the proposed method.

I. INTRODUCTION

Automatic and reliable positioning, handling and assembly of the micro-structures have shown their potential in fabrication of micro-structures and microelectromechanical systems (MEMS) over the past few decades. In the micro-scale, the prevailing spatial sensing technique is vision, through optical or electron microscopy. Visual servoing is hence an unavoidable tool to automate manipulation tasks [1], [2]. The accuracy in position and orientation of a sample in the camera reference frame is one important issue [3], [4]. Many approaches in this domain rely the observation of object features [5], [6]. The pose (i.e., position and orientation) of samples is computed by estimating the position and orientation of these features. Furthermore, available imaging techniques at the micro-scale by an optical or electronic microscope exhibit several limitations such as high signal-to-noise ratio, low refresh rate, contrast, etc., leading to the bottleneck of pose detection.

Recently, a novel visual servoing technique has been proposed to overcome these different issues. This approach uses the photometric information [7], [8] as a visual feature in the image based control law. In this case, only the image intensity is needed while visual tracking, including feature extraction and motion prediction is no longer necessary. This approach is demonstrated in micro-positioning on 3-DoF (translation along x and y axes, and rotation around z

axis) [9], and on 6-DoF with some limitations on translation along z axis [10].

Particularly in micro-nano scale, the motion along z axis is difficult to observe due to microscope projection model [11]. As a microscope provides a long focal length, a motion along z axis causes only a tiny variation on the observed image. Using only image intensity information as a visual feature, the accuracy of the translation on z axis is inferior to other DoFs. Accurate depth information is an important issue for correct positioning. Many approaches in depth detection have been proposed in computer vision. One idea is to employ stereo vision and reconstruct the three-dimensional (3D) image from several two-dimensional (2D) images [12], in which the reliable extraction of features is necessary. Both extraction and feature matching could be computationally expensive. Considering focus information, depth from focus (DFF) [13] has been proposed. The basic of this method is to obtain different focus levels by adjusting the camera parameters (i.e., the distance between the lens and image plane, the focal length and the aperture radius). It involves obtaining many observations for the various camera parameters and estimating the focus using a criterion function. Since many camera parameters and observations should be considered, it is not practical in positioning. Alternatively, the depth from defocus (DFD) based approaches have been also widely discussed [14], [15]. The main idea of these methods is objects at a particular distance from the lens will be focused in an optical system, whereas objects at other distances will be blurred. By measuring the amount of defocus of the object on the observed image, the depth of the object with respect to the lens can be recovered from the geometric optics. In these methods, the defocus parameters can be estimated from the image by an inverse filter in the Fourier domain [14], by modeling the depth and the image as separate Markov random fields [16], or by some spatial domain-based techniques [17], etc. Inspired from the DFD, considering that the image sharpness and focus vary significantly when using an optical path which provides a small depth of field, the image gradient information can be hence used for controlling z motion in a visual servoing scheme.

This study addresses a closed-loop control scheme for visual servoing using image intensity and image gradient for a micro-positioning task in 6-DoF. The proposed method is demonstrated experimentally on a 6-DoF parallel-kinematics positioning stage and a digital microscope. The manuscript is organized as follows: Section II briefly recalls the classic visual servoing and introduces the general approach. Section

¹Le Cui and Eric Marchand are with Université de Rennes 1, Lagadic group, IRISA, Rennes, France le.cui@univ-rennes1.fr, eric.marchand@irisa.fr

²Sinan Haliyo and Stéphane Régnier are with Sorbonne Universités, UPMC Univ Paris 06, UMR 7222, ISIR, Paris, France sinan.haliyo@upmc.fr, stephane.regnier@upmc.fr

III describes selected visual features, image intensity and gradient. The hybrid control laws constructed with these features are presented in Section III C. Experimental results obtained on a 6-DoF parallel micro-assembly workcell are shown in Section IV.

II. CLASSICAL VISUAL SERVOING

Two distinct robot-camera relation cases exist in classical visual servoing [18]: eye-in-hand case, in which the camera is installed in the end-effector, and alternatively, the eye-to-hand case, where the camera is fixed and look toward the end-effector. In micro/nano-robotics, the eye-to-hand case is generally considered since the sensor (microscope) is usually motionless.

In classical visual servoing, the objective is to minimize the error \mathbf{e} between the current visual feature $\mathbf{s}(\mathbf{q})$ and the desired one \mathbf{s}^* :

$$\mathbf{e}(\mathbf{q}) = \mathbf{s}(\mathbf{q}) - \mathbf{s}^* \quad (1)$$

The relation between the time derivative $\dot{\mathbf{s}}$ and the robot joint velocity $\dot{\mathbf{q}}$ is given by:

$$\dot{\mathbf{s}} = \mathbf{J}_s \dot{\mathbf{q}} \quad (2)$$

where \mathbf{J}_s represents the visual feature Jacobian.

With an exponential decrease of the error $\dot{\mathbf{e}} = -\lambda \mathbf{e}$, with (1) and (2), the control law can be expressed as:

$$\dot{\mathbf{q}} = -\lambda \mathbf{J}_s^+ \mathbf{e} \quad (3)$$

where λ is the proportional coefficient and \mathbf{J}_s^+ is the pseudo-inverse of \mathbf{J}_s .

Considering the eye-to-hand visual servoing context, the Jacobian \mathbf{J}_s can be expressed as:

$$\mathbf{J}_s = -\mathbf{L}_s {}^c \mathbf{V}_F {}^F \mathbf{J}_n(\mathbf{q}) \quad (4)$$

where \mathbf{L}_s represents the interaction matrix, which links the relative camera instantaneous velocity \mathbf{v} and the feature motion $\dot{\mathbf{s}}$, ${}^c \mathbf{V}_F$ is the motion transform matrix which transforms velocity expressed in camera reference frame onto the robot frame, ${}^F \mathbf{J}_n(\mathbf{q})$ is the robot Jacobian in the robot reference frame.

To improve the robustness of algorithm, a Levenberg-Marquardt-like method is considered:

$$\dot{\mathbf{q}} = -\lambda (\mathbf{H} + \mu \cdot \text{diag}(\mathbf{H}))^{-1} \mathbf{J}_s^T \mathbf{e}(\mathbf{q}) \quad (5)$$

where μ is a coefficient whose typical value ranges from 0.001 to 0.0001. $\text{diag}(\mathbf{H})$ represents a diagonal matrix of the matrix $\mathbf{H} = \mathbf{J}_s^T \mathbf{J}_s$.

III. PURE IMAGE INFORMATION BASED VISUAL FEATURES

In visual servoing, one or more feature information, such as geometric measurements (e.g. position and orientation of interesting points) or direct image information including image gradient [19],[20] and image entropy [21] can be extracted as visual features. In [9],[10], the photometric information is shown to be efficient for 6-DoF visual servoing. However, the accuracy of translation along z axis is much

inferior to others due to the lack of observation on z axis in micro/nano imaging: At high magnifications, existing models assume an orthogonal projection hence the output image varies slowly over the object/camera distance. In this paper a hybrid approach with features extracted from the pure image information is proposed for 6-DoF visual servoing. The general idea is to use image gradient information for controlling the motion on z axis and to use image intensity information to control the other degrees of freedom.

A. Image intensity as a visual feature

First considering the intensity of all the pixels I from the pure image as the main visual feature \mathbf{s} , the cost function is defined as:

$$\mathbf{e}_I(\mathbf{q}) = \mathbf{I}(\mathbf{q}) - \mathbf{I}^* \quad (6)$$

For a pixel $\mathbf{x} = (x, y)$ on image plane, the time deviation of \mathbf{x} can be expressed by

$$\dot{\mathbf{x}} = \mathbf{L}_x \mathbf{v}. \quad (7)$$

where $\mathbf{v} = (v, \mathbf{w})$ contains the relative camera instantaneous linear velocity v and angular velocity \mathbf{w} , \mathbf{L}_x is the interaction matrix:

$$\mathbf{L}_x = \begin{bmatrix} -\frac{1}{z} & 0 & xy & -(1+x^2) & y \\ 0 & -\frac{1}{z} & 1+y^2 & -xy & -x \end{bmatrix}. \quad (8)$$

Let $I(\mathbf{x}, t)$ be the intensity of the pixel \mathbf{x} at time t , then

$$\nabla I = \begin{bmatrix} \frac{\partial I}{\partial x} & 0 \\ 0 & \frac{\partial I}{\partial y} \end{bmatrix}, \quad (9)$$

the total deviation of the intensity $I(\mathbf{x}, t)$ can be written as

$$\dot{I}(\mathbf{x}, t) = \nabla I \dot{\mathbf{x}} + \dot{I}, \quad (10)$$

where $\dot{I} = \frac{\partial I}{\partial t}$ represents the time variation of I . According to [22] based on the temporal luminance constancy hypothesis, $\dot{I}(\mathbf{x}, t) = 0$. In this case,

$$\dot{I} = -\nabla I \mathbf{L}_x \mathbf{v} = \mathbf{L}_I \mathbf{v}. \quad (11)$$

Considering the entire image, $\mathbf{I} = (I_{00}, I_{01}, \dots, I_{MN})$, where M, N represent the image size:

$$\dot{\mathbf{I}} = \begin{pmatrix} \mathbf{L}_{I_{00}} \\ \vdots \\ \mathbf{L}_{I_{MN}} \end{pmatrix} \mathbf{v} = \mathbf{L}_I \mathbf{v} \quad (12)$$

where $\dot{\mathbf{I}}$ is the variation of the whole image intensity. \mathbf{L}_I is a $MN \times 5$ matrix that theoretically allows a control law able to compute the 5 DoFs.

B. Image gradient as a visual feature

As mentioned before, one difficulty in micro/nano vision at a high magnification is the lack of observation on z axis. Actually, for a sensor with small depth of field, it is evident that the image sharpness changes when z position changes. That leads the motion on z can be controlled according to the image sharpness. Hence, the image gradient is chosen as a visual feature to control the motion along z axis. An approach employing image gradient information in visual

servoing is proposed in this part. The general idea is to consider that the image gradient varies when object position on z axis changes, where the focal length of the sensor is always a constant. Practically, for a sensor with a small depth of field, the focal length of sensor can be adjusted in order to acquire a sharp image at the desired pose. In this case, the aim is to minimize the error of image gradient between the image at current pose and the image at desired pose to move the object to the target pose.

The linear image formation model, which is commonly used in the case of optical sensors [13] is employed here. Let Z be the current position on z axis, the defocus image $\mathbf{I}(x, y, Z)$ at the position Z can be expressed as the convolution of a sharp image $\mathbf{I}^*(x, y, Z^*)$ at the desired pose Z^* and a defocus kernel $f(x, y)$:

$$\mathbf{I}(x, y, Z) = \mathbf{I}^*(x, y, Z^*) * f(x, y) \quad (13)$$

In previous studies, such as [23], the Gaussian kernel is widely used as an approximation of defocus model by many authors. Its probability density function (PDF) can be expressed by

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (14)$$

For a small displacement dZ on z axis, consider a proportional relation between $d\sigma$ and dZ :

$$d\sigma = mdZ. \quad (15)$$

where m is a constant coefficient.

For an image $\mathbf{I}(x, y, Z)$ at Z on z axis, the square of the norm of the image gradient at a point (x, y) on the image plane is

$$\begin{aligned} g(x, y, Z) &= \|\nabla\mathbf{I}(x, y, Z)\|^2 \\ &= \nabla I_x^2(x, y, Z) + \nabla I_y^2(x, y, Z) \end{aligned} \quad (16)$$

Considering the square of the norm of the image gradient for the whole image as the visual feature:

$$\begin{aligned} G(Z) &= \sum_{x=0}^M \sum_{y=0}^N g(x, y, Z) \\ &= \sum_{x=0}^M \sum_{y=0}^N (\nabla I_x^2(x, y, Z) + \nabla I_y^2(x, y, Z)) \end{aligned} \quad (17)$$

One goal is then to minimize the error of the current image gradient $G(Z)$ and the desired image gradient $G^*(Z^*)$. In this case, the cost function is defined as:

$$e_G(Z) = G(Z) - G^*(Z^*) \quad (18)$$

The relation between the relative camera instantaneous linear velocity v_z along z axis and the time variation of image gradient G is

$$\dot{G} = L_G v_z \quad (19)$$

where L_G is the Jacobian (hence a scalar) which can be expressed by:

$$L_G = \frac{\partial G}{\partial \sigma} \frac{\partial \sigma}{\partial Z} \quad (20)$$

From (15), it leads to

$$L_G = m \frac{\partial G}{\partial \sigma}. \quad (21)$$

where $\frac{\partial G}{\partial \sigma}$ can be expressed by

$$\begin{aligned} \frac{\partial G}{\partial \sigma} &= \sum_{x=0}^M \sum_{y=0}^N 2(\nabla I_x(x, y) \frac{\partial \nabla I_x(x, y)}{\partial \sigma} \\ &\quad + \nabla I_y(x, y) \frac{\partial \nabla I_y(x, y)}{\partial \sigma}) \end{aligned} \quad (22)$$

In (13), the convolution can also be written as:

$$\mathbf{I}(x, y) = \sum_u \sum_v \mathbf{I}^*(x-u, y-v) f(u, v). \quad (23)$$

From (14), compute the derivative

$$\frac{\partial f(u, v)}{\partial \sigma} = \frac{1}{2\pi} (u^2 + v^2 - 2\sigma^2) \sigma^{-5} e^{-\frac{u^2+v^2}{2\sigma^2}}. \quad (24)$$

According to (23) and (24):

$$\begin{aligned} \frac{\partial \nabla I_x(x, y)}{\partial \sigma} &= \sum_u \sum_v \nabla (I_x^*(x-u, y-v) \\ &\quad \cdot \frac{1}{2\pi} (u^2 + v^2 - 2\sigma^2) \sigma^{-5} e^{-\frac{u^2+v^2}{2\sigma^2}}) \end{aligned} \quad (25)$$

and

$$\begin{aligned} \frac{\partial \nabla I_y(x, y)}{\partial \sigma} &= \sum_u \sum_v \nabla (I_y^*(x-u, y-v) \\ &\quad \cdot \frac{1}{2\pi} (u^2 + v^2 - 2\sigma^2) \sigma^{-5} e^{-\frac{u^2+v^2}{2\sigma^2}}) \end{aligned} \quad (26)$$

Considering (25) and (26) in (22), L_G can be finally computed.

C. Control law for hybrid visual servoing

For the hybrid visual servoing, both image intensity and image gradient are considered as visual features $\mathbf{s} = (\mathbf{I}(\mathbf{q}), G(Z))^T$. Using image intensity as a visual feature, the velocities (the linear and angular velocities on x, y axes, and the angular velocity around z axis) of the end-effector can be computed from (4), where the Jacobian is

$$\mathbf{J}_I = -\mathbf{L}_I^c \mathbf{V}_F^F \mathbf{J}_n(\mathbf{q}). \quad (27)$$

Similarly, using image gradient as a visual feature, the linear velocity along z axis is:

$$\dot{Z} = -\lambda_z L_G^{-1} e_G(Z) \quad (28)$$

where λ_z is an exponential coefficient. During the visual servoing process, the control laws for motion along z axis and the other 5 DoFs are computed respectively and simultaneously.

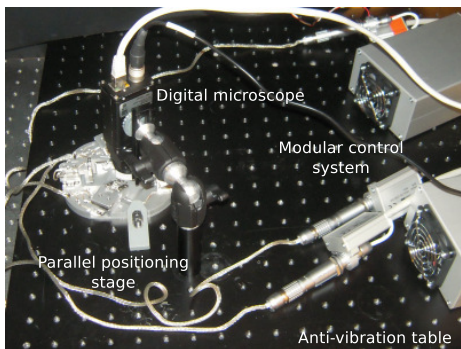


Fig. 1. The micropositioning workcell

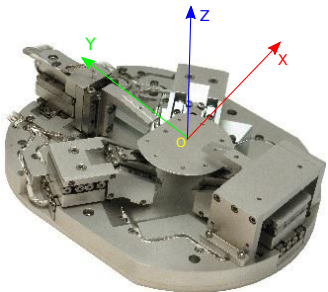


Fig. 2. Parallel positioning stage

IV. EXPERIMENTAL RESULTS

A. Experimental setup

Experiments are accomplished on a micropositioning workcell installed on an anti-vibration table shown in Fig. 1. It contains a 6-DoF positioning-kinematics micro-stage¹ as well as its modular control system and a digital microscope² with an aperture-adjustable lens towards the top-plate of positioning stage. Experiments are realized on an optical magnification of $60\times$.

TABLE I
POSITIONING STAGE SPECIFICATIONS

	Travel range	Closed-loop resolution
X	+/-6 mm	1 nm
Y	+/-6 mm	1 nm
Z	+/-3 mm	1 nm
θ_X	+/-10°	1 μ rad
θ_Y	+/-10°	1 μ rad
θ_Z	+/-20°	1 μ rad

The SmarPod positioning stage is a parallel robot (hexapod) that provides three positioners supporting a top-plate. By the motion of its positioners, the top-plate can be moved in three directions and rotated around three axes. The hexapod and the reference frame are shown in Fig. 2. Table I describes its specifications.

The specimen is a microchip which measures $10\text{ mm}\times 5\text{ mm}$, with 0.5 mm in thickness. The resolution of acquired

¹SmarPod 70.42-S-HV made by SmarAct and its positioner SLC 17.20-S-HV

²Basler acA1600-60gm

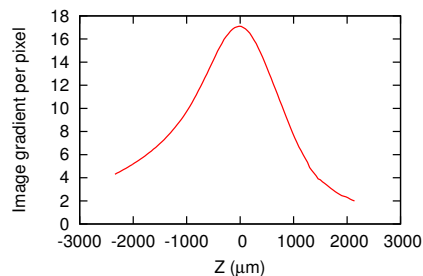


Fig. 3. Image gradient per pixel with respect to z position

image in our experiments from the digital microscope is 659×494 pixels.

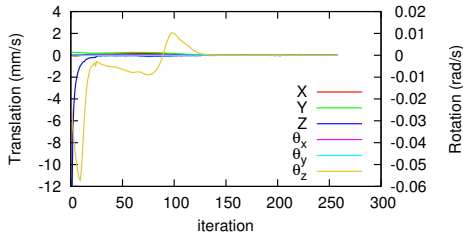
B. Validation of the method

First, the positioning stage is moved from -2.3 mm to 2.1 mm along the z axis to evaluate the variation of image gradient with respect to z position. Images are acquired at each $40\text{ }\mu\text{m}$ step. By computing the image gradient per pixel for each image, the relation between the image gradient and z position is shown in Fig. 3. It can be seen from the figure that the depth of field is small enough for an accurate positioning and a single optimum is found in image gradient.

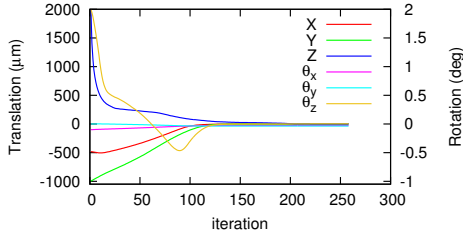
In positioning experiments, the stage is first set to an initial pose then moved to a predefined desired pose iteratively by comparing the image at the desired pose with the image at the current pose. The focus of microscope is adjusted so that the image is focused at the desired pose.

To validate the method, the initial pose of positioning stage is set to $500\text{ }\mu\text{m}$ in x and 1 mm in y axes, 2 mm in z axis; 0.1° around x axis, 2° around z axes away from the desired pose to test the performance of the proposed method. The initial image and desired image after image processing is shown in Fig. 5(a) and Fig. 5(b), respectively. Fig. 5(c) to Fig. 5(d) shows the evolution of image intensity error $\mathbf{e}_I(\mathbf{q}) = \mathbf{I}(\mathbf{q}) - \mathbf{I}^*$ until the end of visual servoing procedure. The velocities converge fast to 0. As a consequence of optimization, the error image is almost null.

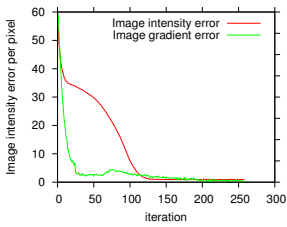
The experimental results are shown in Fig. 4. Since visual servoing is robust to calibration errors [24], the positioning task without explicit calibration performs also quite well. The image intensity error as well as image gradient error per pixel decrease to negligible values when the velocities converge. The object pose errors between the final pose and the desired pose reach $0.65\text{ }\mu\text{m}$, $0.47\text{ }\mu\text{m}$ and $0.17\text{ }\mu\text{m}$ in translation along x, y, z axes; 0.027° , 0.036° and 0.003° in rotation around x, y, z axes, respectively. It can be mentioned that in Fig. 4(c), the image gradient error increases around the 70th iteration. It is mainly because the sample is too large that cannot be presented in the whole image. When the positioning stage is moving, details of the sample on the image vary, which causes the computation of image gradient to be slightly disturbed. In experiments, the proposed visual servoing scheme shows robustness to such situations.



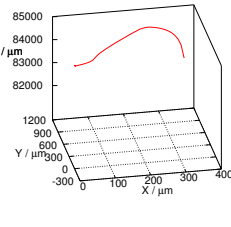
(a)



(b)



(c)



(d)

Fig. 4. 6-DoF positioning using hybrid visual servoing (a) Evolution of joint velocity (in mm/s and rad/s). (b) Evolution of object pose error (in $\mu\text{m/s}$ and degree). (c) Evolution of image intensity error and image gradient error per pixel. (d) Object trajectory in camera frame

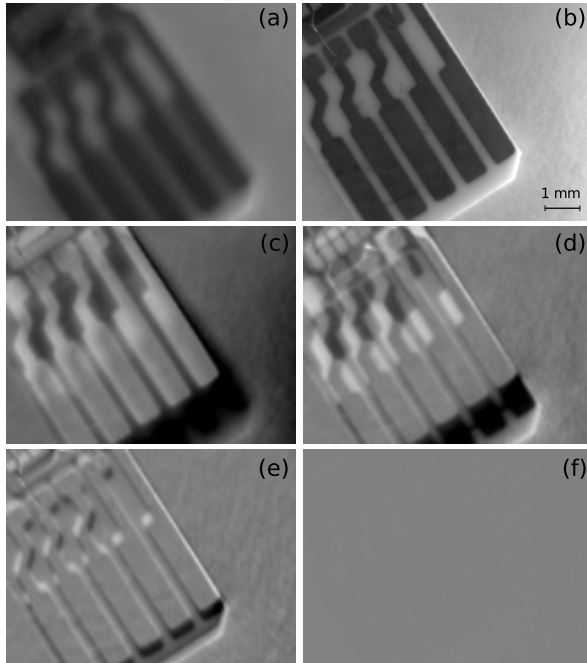


Fig. 5. Progress of 6-DoF positioning using hybrid visual servoing (a) Initial image, (b) desired image, (c) to (f) show image error $e_I(\mathbf{q})$ at 1st, 16th, 82th and last iteration.

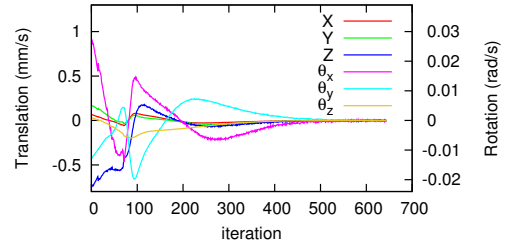


Fig. 6. Positioning using only image intensity based visual servoing: Evolution of joint velocity (in mm/s and rad/s)

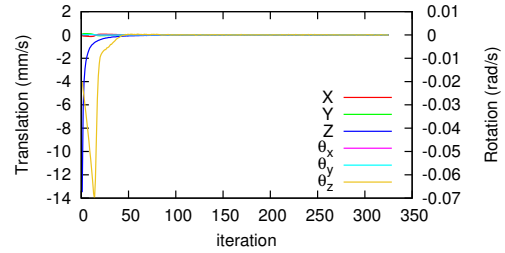


Fig. 7. Positioning using hybrid visual servoing: Evolution of joint velocity (in mm/s and rad/s)

C. Hybrid visual servoing vs visual servoing using image intensity

Experiments have been achieved to evaluate the proposed hybrid visual servoing by comparing to an approach using only image intensity [10] in the same conditions. The initial pose of the positioning stage is set to be 2 mm on z axis and 2° around z axis away from the desired pose. The positioning task based on the proposed hybrid visual servoing and image intensity based visual servoing are accomplished respectively. The evolution of joint velocity are illustrated in Fig. 6 and Fig. 7. The positioning error on translation along z axis with hybrid visual servoing is $0.26 \mu\text{m}$, which is smaller than with the image intensity based visual servoing ($1.66 \mu\text{m}$), where also more iterations are needed to converge. On other DoFs, both these two methods perform equivalent, where the pose errors are less than $0.6 \mu\text{m}$ in translation along x and y axes, less than 0.01° in rotation around x , y and z axes. Furthermore, because of the limited travel range of the positioning stage on z axis, the initial pose on z cannot be extremely far away from the desired pose. Indeed, in that case, the image-intensity-only method fails to converge because little details can be extracted from the initial blur image. However, the hybrid visual servoing performs well since the motion on z axis can be conducted even the image is heavily blurred.

D. Robustness to light variation

As the proposed method uses photometric information, the sensibility to variable light conditions is indeed an important issue. Therefore, the robustness to light variation of the proposed method is tested. The initial pose of the stage is also set to be 2 mm on z axis and 2° around z axis. Fig. 8 shows the evolution of joint velocity. The luminance

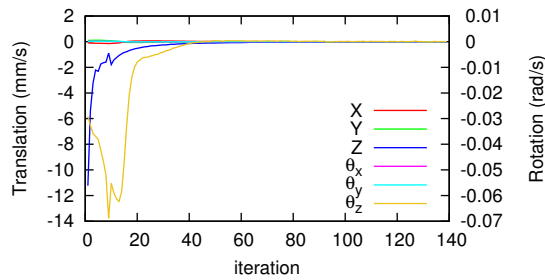


Fig. 8. Positioning using hybrid visual servoing with lightning perturbation: Evolution of joint velocities (in mm/s and rad/s)

of the environment light is changed suddenly at the 8th iteration. Oscillations in velocities appear, caused by the lighting changes. However, the convergence and the accuracy are unaffected in spite of the changing light. The system keeps stable for a small perturbation occurring during the positioning task.

V. CONCLUSION

In this paper, a hybrid visual servoing scheme is proposed for an accurate and robust 6-DoF micropositioning. Differently from traditional visual tracking and object localization approach, only pure image photometric information is needed in this method. The image intensity information is employed to control the linear motion along x, y axes and the angular motion around x, y, z axes. To improve the precision of positioning on z axis, image gradient is introduced as a new visual feature according to the variation of image sharpness due to the motion along z axis. This method is validated by experiments on a 6-DoF parallel positioning stage and a digital microscope. Although the initial pose is far away from the desired pose (2 mm in translation and 2° in rotation), the results are accurate (pose errors are below $1 \mu\text{m}$ in translation, below 0.04° in rotation around x, y and 0.003° in rotation around z). Future work will be to modify the sensor projection model and to validate this approach with a scanning electron microscope.

ACKNOWLEDGMENT

The authors would like to thank Soukeyna Bouchebout, Mokrane Boudaoud and Jean Ochin Abrahamians for their help during the experiments at ISIR, Université Pierre et Marie Curie. The positioning stage is supported by French Robotex platform. This work was realized in the context of the French ANR P2N Nanorobust project (ANR-11-NANO-0006).

REFERENCES

[1] C. Pawashe and M. Sitti, "Two-dimensional vision-based autonomous microparticle manipulation using a nanoprobe," *J. Micromechatronics*, vol. 3, pp. 285–306, July 2006.

[2] S. Fatikow, C. Dahmen, T. Wortmann, and R. Tunnell, "Visual feedback methods for nanohandling automation," *I. J. Information Acquisition*, vol. 6, no. 3, pp. 159–169, 2009.

[3] S. Devasia, E. Eleftheriou, and S. Moheimani, "A survey of control issues in nanopositioning," *Control Systems Technology, IEEE Transactions on*, vol. 15, no. 5, pp. 802–823, 2007.

[4] N. Ouarti, B. Sauvete, S. Haliyo, and S. Régnier, "Robposit, a robust pose estimator for operator controlled nanomanipulation," *Journal of Micro-Bio Robotics*, vol. 8, no. 2, pp. 73–82, 2013.

[5] Y. Sun, M. Greminger, D. Potasek, and B. Nelson, "A visually servoed mems manipulator," in *Experimental Robotics VIII*. Springer, 2003, pp. 255–264.

[6] N. Ogawa, H. Oku, K. Hashimoto, and M. Ishikawa, "Microrobotic visual control of motile cells using high-speed tracking system," *Robotics, IEEE Transactions on*, vol. 21, no. 4, pp. 704–712, 2005.

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

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

[9] B. Tamadazte, N. Le-Fort Piat, and E. Marchand, "A direct visual servoing scheme for automatic nanopositioning," *Mechatronics, IEEE/ASME Transactions on*, vol. 17, no. 4, pp. 728–736, 2012.

[10] L. Cui, E. Marchand, S. Haliyo, and S. Régnier, "6-dof automatic micropositioning using photometric information," in *IEEE/ASME Int Conf. on Advanced Intelligent Mechatronics, AIM'14*, 2014.

[11] L. Cui and E. Marchand, "Calibration of scanning electron microscope using a multi-images non-linear minimization process," in *IEEE Int. Conf. on Robotics and Automation, ICRA'14*, 2014.

[12] J.-L. Pouchou, D. Boivin, P. Beauchêne, G. L. Besnerais, and F. Vignon, "3d reconstruction of rough surfaces by sem stereo imaging," *Microchimica Acta*, vol. 139, no. 1-4, pp. 135–144, 2002.

[13] S. K. Nayar and Y. Nakagawa, "Shape from focus," *Pattern analysis and machine intelligence, IEEE Transactions on*, vol. 16, no. 8, pp. 824–831, 1994.

[14] A. P. Pentland, "A new sense for depth of field," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 4, pp. 523–531, 1987.

[15] P. Favaro, A. Mennucci, and S. Soatto, "Observing shape from defocused images," *Internat. J. Comput. Vision.*, vol. 52, no. 1, pp. 25–43, 2003.

[16] S. Bhasin and S. Chaudhuri, "Depth from defocus in presence of partial self occlusion," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, 2001, pp. 488–493 vol.1.

[17] D. Ziou and F. Deschenes, "Depth from defocus estimation in spatial domain," *Computer Vision and Image Understanding*, vol. 81, no. 2, pp. 143 – 165, 2001.

[18] F. Chaumette and S. Hutchinson, "Visual servo control, Part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, December 2006.

[19] E. Marchand and C. Collewet, "Using image gradient as a visual feature for visual servoing," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'10*, Taipei, Taiwan, October 2010.

[20] E. Marchand and N. Courty, "Controlling a camera in a virtual environment: Visual servoing in computer animation," *The Visual Computer Journal*, vol. 18, no. 1, pp. 1–19, Feb. 2002.

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

[22] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185–203, Aug. 1981.

[23] J. Ens and P. Lawrence, "An investigation of methods for determining depth from focus," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 2, pp. 97–108, 1993.

[24] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," in *Int. Symposium on experimental Robotics, ISER'93*, Kyoto, 1993.