Highly precise micropositioning task using a direct visual servoing scheme.

Brahim Tamadazte, Guillaume Duceux, Nadine Le Fort-Piat, Eric Marchand

Highly Precise Micropositioning Task using a Direct Visual Servoing Scheme

B. Tamadazte, G. Duceux, N. Le-Fort Piat, and E. Marchand,

Abstract— This paper demonstrates a precise micropositioning scheme based on a direct visual servoing process. This technique uses only the pure image signal (photometric information) to design the control law. With respect to traditional visual servoing approaches that use geometric visual features (points, lines, ...), the visual features used in the control law is nothing but the pixel luminance. The proposed approach was tested in term of precision and robustness in several experimental conditions. The obtained results have demonstrated a good behavior of the control law and very good positioning precision. The obtained precisions are 89 nanometers, 14 nanometers, and 0.001 degrees in the x, y and θ axes of positioning platform, respectively.

I. OVERVIEW

The heterogeneous integration of high-performance electronic devices, microelectromechanical structures (MEMS), and optoelectronic devices onto the same substrate is important for the development of low-cost, high performance, and high-compact microsystems [9]. To set up intelligent and miniature systems, handle and assemble the various elements that constitute these microsystems, it is necessary to be able to manipulate the different individual micro-elements which compose the final MEMS. These operations (i.e. handle and assemble) must be done with high precision. In the last decade considerable researches have been performed on the development of robotic microassembly station, gripping systems, precise actuators, micromanipulation and microassembly strategies. Furthermore, a lot of works has been done concerning the development of control approaches to automatize the different micromanipulation and microassembly tasks such positioning, orientation, picking, placing and insertion of the different micro-objects [10], [8]. Most of these works consider the use of the vision sensor to control the behavior of the robotic structure of the microassembly station during the assembly process. Thus, the guidance of robots through real-time and continuous visual feedback is generally known as visual servoing [4], and the continuous observation of the objects of interest is referred to visual tracking [10], [11]. Visual tracking of an object involves the detection of some known object features in the acquired images and, using these features, the estimation of the object position and orientation.

A feature can be a distinctive part of the object and can exist naturally as a part of the geometry (i.e. a corner, an edge), or as a deliberately fabricated marking on the surface (i.e. markers). This tracking process is one of the bottleneck of the development of visual servoing techniques. Recently, it has been shown that these tracking and matching processes can be totally removed and that no other information than the image intensity [7] [2], the image entropy [3], and the image Laplacian can be considered to control a manipulator. In this paper, we will consider such a direct approach to control the microrobot motion with a precision. Only the image (as a whole) are used to design the control law. However, despite the fact that no complex image processing algorithms are consider, we will show that these new techniques are:

- robust to global light variations
- robust to partial occlusions
- robust to different perturbations such as the addition of others micro-objects during the manipulation process

Furthermore, as all the pixels of the image are considered, the proposed scheme is highly precise thanks to the redundant information.

In this paper, we use this photometric visual servoing scheme in an eye-to-hand configuration. The camera mounted on an optical microscope is motionless and observes a moving positioning platform. The developed methods are validated using a five degrees of freedom (dof) microassembly workcell. This development has been also validated using a multiple scale visual servoing which means the integration of the dynamic control of the optical microscope magnification in the control law.

Section II describes the experimental setup used to validate the proposed approach. Section III presents some definition about the principles of the eye-to-hand visual servoing. Section IV describes the new visual servoing approach without image processing using only the pixels intensity of the image as visual features, and the design of the new control law. Section V discusses the experimental results using the developed method in terms of the obtained precision quality during the different positioning and orientation tasks.

II. EXPERIMENTAL SETUP

The integration of the developed concept is done on the microassembly workcell illustrated in Fig. 1. This MEMS microassembly station had been developed in our laboratory. It includes a robotic system with five high accuracy dof (3-dof positioning platform: two linear stages i.e. xy and one rotating stage θ, and 2-dof micromanipulator: one vertical linear stage z and one rotating stage mounted at 45 degrees
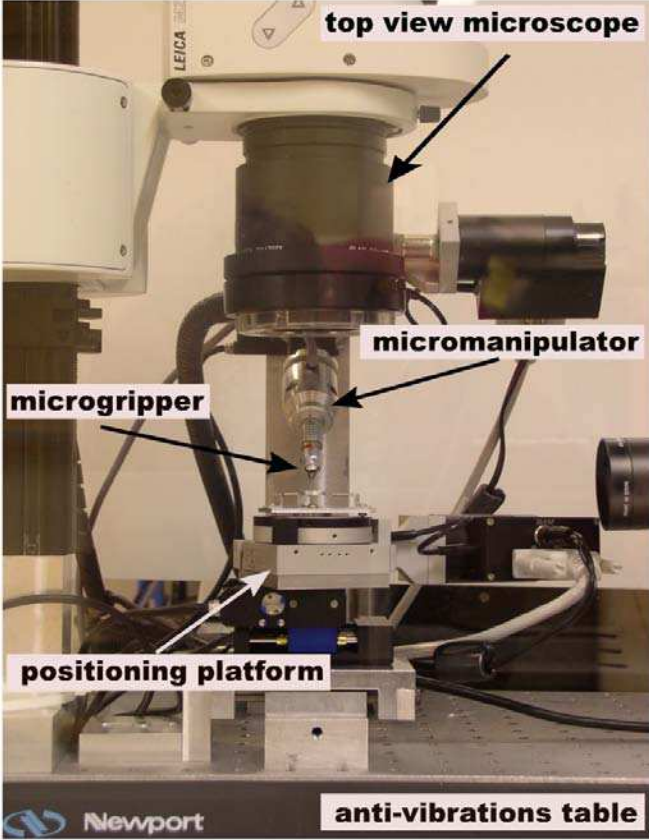
Fig. 1. Photography of the micromanipulation workcell.

from the vertical one), a microhandling system equipped 4-dof and allows open-and-close motions as well as up-and-down motions. It is based on piezoelectric actuators which consists of a two parallel piezoceramic PZT PIC 151 bimorphs. The imaging system is a video stereo microscope of the type LEICA MZ 16 A. It delivers a top view of the work scene. The zoom (and thus the magnification) and the focus are motorized and controlled by a PC. The field-of-view (FOV) varies from 700 μm $\times$ 900 μm (with a resolution of 1.4 μm/pixel) at the maximum magnification to 20 mm $\times$ 25 mm (with a resolution of 21 μm/pixel) at the minimum magnification. The depth-of-view (DOF) varies from 2.9 mm to 0.035 mm and the work distance is approximately 130 mm. Everything being placed on an anti-vibrations table and inside a controlled environment.

III. TRADITIONAL VISUAL SERVOING APPROACHES

Visual servoing techniques consist of using the data provided by one or several cameras in order to the control the motion of a robotic system [1], [6]. Whatever, the sensor configuration a set of visual features s has to be designed from the visual measurements $\mathbf{x}(t)$ ($\mathbf{s} = \mathbf{s}(\mathbf{x}(t))$), allowing control of the desired degrees of freedom. A control law has to be designed also so that these features s reach a desired value $\mathbf{s}^*$, defining a correct realization of the task. The aim of vision-based control schemes [6] is to minimize the error

e typically defined by:

$$\mathbf{e} = \mathbf{s} - \mathbf{s}^*. \tag{1}$$

The equation that links the variation $\dot{\mathbf{s}}$ of the visual feature s to the robot instantaneous camera velocity v is given by:

$$\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v} \tag{2}$$

where $\mathbf{L_s}$ represents the interaction matrix related to s and v the robot instantaneous velocity [1]. If we want to control the robot using the joint velocities, we have:

$$\dot{\mathbf{s}} = \mathbf{J_s}\dot{\mathbf{q}} \tag{3}$$

where $\mathbf{J_s}$ is the features Jacobian and $\dot{\mathbf{q}}$ the joint velocities. In this paper we mainly consider an eye-to-hand configuration. We then have:

$$\mathbf{J_s} = -\mathbf{L_s}{}^c\mathbf{V}_\mathcal{F}{}^\mathcal{F}\mathbf{J}_n(\mathbf{q}) \tag{4}$$

where

- ${}^\mathcal{F}\mathbf{J}_n(\mathbf{q})$ is the robot Jacobian expressed in the robot reference frame $R_\mathcal{F}$. In our case since only the manipulator has 3-dof ${}^\mathcal{F}\mathbf{J}_n$ is a 6×3 matrix ;
- ${}^c\mathbf{V}_\mathcal{F}$ allows to transform the velocity screw between coordinate frames (here the camera frame R_c and the robot reference frame $R_\mathcal{F}$).

Thereafter, the control law is designed as follows:

$$\dot{\mathbf{q}} = -\lambda \mathbf{J_s^+}\mathbf{e} \tag{5}$$

where λ is the proportional coefficient involved in the exponential convergence of the error, and $\mathbf{J_s^+}$ the pseudo-inverse of $\mathbf{J_s}$.

IV. VISUAL SERVOING WITHOUT IMAGE PROCESSING

In traditional visual servoing techniques, the visual features s are built from image measurements. s are mainly geometric features such as points or lines coordinates, image moments or even 3D features. These classical approaches require detection, tracking or matching steps and an estimation in each frame of the value of the visual features. This tracking or matching process has been, to date, considered as a necessary step and is also one of the bottlenecks of the expansion of visual servoing.

In this section, we shall consider a direct visual control scheme. The proposed method uses the pure image signal to design the vision-based control law. Rather than computing the visual from some image measurement, we consider that the vector of visual feature is the image itself. The visual feature s is replaced by the luminance $\mathbf{I}$ of each pixel of the image. Therefore, we can write:

$$\mathbf{s} = \mathbf{I} \tag{6}$$

where

$$\mathbf{I} = (\mathbf{I}_{1\bullet}, \mathbf{I}_{2\bullet}, \cdots, \mathbf{I}_{\mathbf{N}\bullet}) \tag{7}$$

where $\mathbf{I}_{k\bullet}$ is nothing but the k-th line of the image and $\mathbf{N}$ is the number of the line of the image. To build the control, we have to consider the interaction matrix that links the variation of the image luminance to the camera velocity [2], [7].

Considering the optical flow constraint equation (OFCE) hypothesis [5], we can compute the interaction matrix that links the variation of a pixel intensity to camera motion. The OFCE states that the intensity $\mathbf{I}(\mathbf{x}, t)$ of each projected physical point in the image remains the same during a short time interval dt. We have:

$$\mathbf{I}(\mathbf{x}, t) = \mathbf{I}(\mathbf{x} + \mathbf{dx}, t + dt) \tag{8}$$

.

A first order Taylor expansion of the equation (8) gives:

$$\frac{\partial \mathbf{I}}{\partial x} dx + \frac{\partial \mathbf{I}}{\partial y} dy + \frac{\partial \mathbf{I}}{\partial t} dt = 0 \tag{9}$$

which can be written as follows:

$$\dot{\mathbf{I}} = -\nabla \mathbf{I}_x \dot{x} - \nabla \mathbf{I}_y \dot{y} \tag{10}$$

with

$$\nabla \mathbf{I}_x = \frac{\partial \mathbf{I}}{\partial x} \tag{11}$$

and

$$\nabla \mathbf{I}_y = \frac{\partial \mathbf{I}}{\partial y} \tag{12}$$

Now, the temporal variations of the image must be linked to the camera displacements. For this, we introduce the interaction matrix of a point of the image which links the point velocity in the image to the camera velocity [1]. It is given by:

$$\dot{x} = \mathbf{L}_x \mathbf{v} \tag{13}$$

and

$$\dot{y} = \mathbf{L}_y \mathbf{v} \tag{14}$$

where $\mathbf{L}_x$ and $\mathbf{L}_y$ are the interaction related to the point:

$$\mathbf{L}_x = \begin{pmatrix} -1/Z & 0 & x/Z & xy & -(1+x) & y \end{pmatrix}$$
$$\mathbf{L}_y = \begin{pmatrix} 0 & -1/Z & y/Z & 1+y & xy & -x \end{pmatrix}$$

Hence, introducing equations (13) and (14) in the equation (10), we obtain:

$$\dot{\mathbf{I}} = -\left(\nabla \mathbf{I}_x \mathbf{L}_x + \nabla \mathbf{I}_y \mathbf{L}_y\right) \mathbf{v} \tag{15}$$

or

$$\dot{\mathbf{I}} = \mathbf{L_I} \mathbf{v} \tag{16}$$

Knowing the interaction matrix $\mathbf{L_I}$, it is possible to design a control law. As in [2], we use a control law inspired from the Levenberg-Maquardt optimization algorithm. This provides an efficient numerical solution to the problem of minimizing the error function $\mathbf{I} - \mathbf{I}^*$ which is highly non-linear. It is the interpolation of the Gauss-Newton and gradient descent method. More stable than a simple gradient descent, it converges faster that the Gauss-Newton scheme corresponding to equation (5). Therefore, the platform velocity $\dot{\mathbf{q}}$ is given by:

$$\dot{\mathbf{q}} = -\lambda \left(\mathbf{H} + \mu.diag(\mathbf{H})\right)^{-1} \mathbf{J_I}^\top \left(\mathbf{I} - \mathbf{I}^*\right) \tag{17}$$

where $\mathbf{J_I}$ represents the Jacobian matrix computed from interaction matrix (16) and (4) computed at the desired position. The parameters λ and μ are positive gains and

$diag(\mathbf{H})$ is the matrix of diagonal terms of the combination matrix $\mathbf{H}$ which is given by:

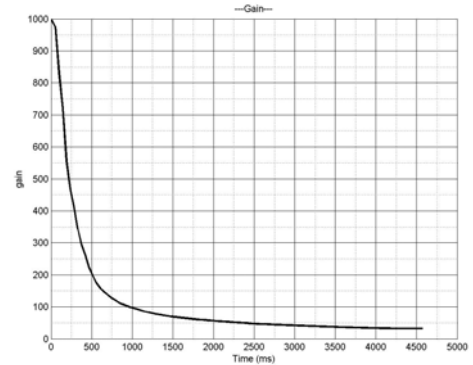$$\mathbf{H} = \mathbf{J_I}^\top \mathbf{J_I} \tag{18}$$

Fig. 2. Evolution of the adaptive gain during the control.

The parameter λ is modified during the control and is function of the variance between the current image $\mathbf{I}$ and the desired image $\mathbf{I}^*$ (see, Fig. 2). It is given by:

$$\lambda = \beta \parallel \mathbf{I} - \overline{\mathbf{I}^*} \parallel^2 \tag{19}$$

where β is a positive gain which allows to adjust the initial value of λ (in the presented experiments, β have a value of $10^2/M$, M being the number of pixels in the image). Figure 2 shows the evolution of the gain λ during a typical experiment.

V. EXPERIMENTAL RESULTS

A. Validation in Nominal Conditions

The methods described previously has been validated using the 5-dof microassembly workcell illustrated in Fig. 1. The task studied in this paper concerns the automatic positioning of silicon micro-object. Object's dimensions are of few hundreds micrometers. In the first experiments, only the 3 dof of the positioning platform are controlled (i.e. planar positioning). In a second time, the control of the optical microscope magnification (zoom) has been included in the control in order to achieve a multiple scale visual servoing. Various experimental situations were tested considering additional external disturbances (i.e. light changing, adding other micro-object, ...) during the positioning process.

The first experience concerns the positioning of a micro-object with dimensions of 400 μm $\times$ 400 μm $\times$ 100 μm. Figure 3 illustrates a sequence of images of the scene acquired during the positioning task. Image 3.(a) shows the initial position of the micro-object, and the images 3.(b) to 3.(k) represent the error between the acquired image (current image) and the desired position (reference image) i.e. $\mathbf{I} - \mathbf{I}^*$. The image 3.(l) shows the error at end of the positioning task.

Figure 4 shows velocities (on the 3-dof $xy\theta$) of the positioning platform during the micromanipulation process.
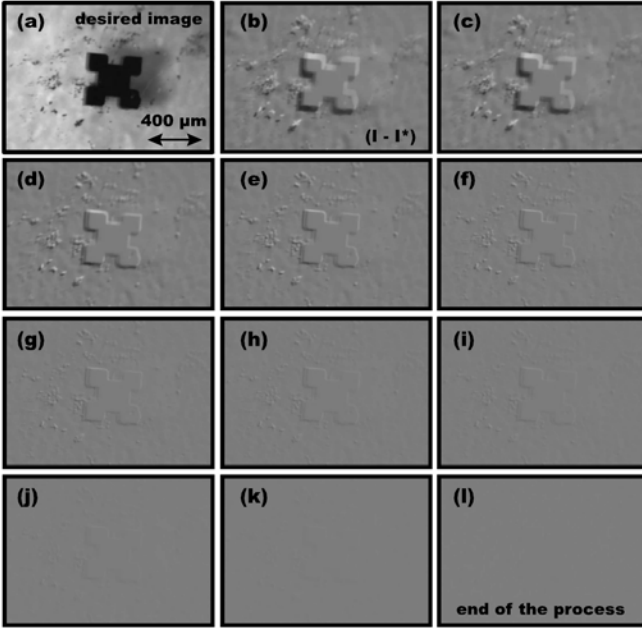
Fig. 3. Image (a) represents the initial position of the micro-object, images (b) to (k) show the error ($\mathbf{I}$ - $\mathbf{I}^*$) between the current image $\mathbf{I}$ and the desired image $\mathbf{I}^*$, and the image (i) shows the end of the visual servoing process where $\mathbf{I}$ - $\mathbf{I}^*$ = 0.
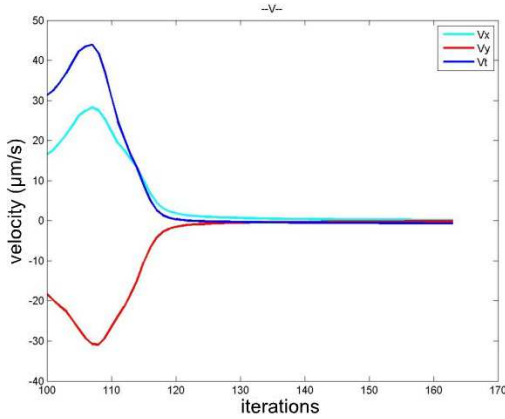

Fig. 4. Representation of the decreasing of the velocities $\dot{\mathbf{q}_x}$, $\dot{\mathbf{q}_y}$ and $\dot{\mathbf{q}_\theta}$ of positioning platform during the control process.

It can be seen the good convergence behavior of the proposed control law.

The precision obtained in the presented experience is estimated to **89 nanometers**, **14 nanometers**, and **0.001 degrees** in the x, y and θ axes, respectively. These precisions are measured using a double plane mirror SIOS SP 120 interferometer characterized by a high resolution of 0.1 nanometers. Otherwise, to reach these precisions, the image are filtered using a 3×3 Median filter in order to reduce the acquisition noise of the CCD sensor. This filter is given by:

$$M_I(p) = med\big(\mathbf{I}(q)|q \subset \mathbf{W}(p)\big) \tag{20}$$

where M_I is the image filtered such that for every pixel p, gray level $M_I(p)$ is the median of gray levels $\mathbf{I}(q)$ of q pixels in the window $\mathbf{W}(p)$.

Fig. 5. Illustration of the acquisition noise of the CCD sensor used in the experimental validations.

The acquisition noise (random image noise) can be shown in Figure 5. This noise image is reconstructed as follows:

$$\begin{aligned} \mathbf{I}(x,y) - \mathbf{I}^*(x,y) &= 0, \; the \; pixel \; is \; white, \\ \mathbf{I}(x,y) - \mathbf{I}^*(x,y) &\neq 0, \; the \; pixel \; is \; black, \; else \end{aligned} \tag{21}$$

This means that the images $\mathbf{I}$ and $\mathbf{I}^*$ represent the same image at the initial position without any displacements, and in stable conditions of illuminations. In an ideal case, Figure 5 should be uniformly white. However, as it can be noticed it is not the case. This is a source of imprecision in the repositioning process and the Median filter is a solution to this issue.

B. Validation with the addition of other Objects during the Process

The proposed vision-based control law has been tested in order to measure its robustness in terms of accuracy and convergence in non-perfect conditions. The first test consists in the addition of an other micro-object during the visual servoing achievement (positioning process). These micro-objects are not present in the desired image but in all images acquired during the positioning task. Figure 6 shows a sequence of images representing this test. Figure 6.(a) illustrates the image (i.e. desired image) of a single silicon micro-object, and figures 6.(b) to 6.(i) show the error $\mathbf{I}$ - $\mathbf{I}^*$ between the current image and the desired image. It can be noticed that in the current image there is a presence of additional micro-objects not present in the desired image. Beside, the changing of the external conditions during the process, it can be seen that the quality of the convergence and the precision remain good. Figure 7 shows the platform velocities during this experiment.

C. Validation using a Mask as a Desired Image

The third experiment is based on the use of an image mask $\mathbf{I}_m$ as desired image. Only the image of the object is considered and the background will not be considered in the desired image. This illustrates the fact that the object has not to be located at a specific position on the platform. The desired image is obtain as:

$$\begin{aligned} \mathbf{I}^*_m(x,y) &= I^*(x,y), \; if \; object, \\ \mathbf{I}^*_m(x,y) &= 255, \; else \end{aligned} \tag{22}$$
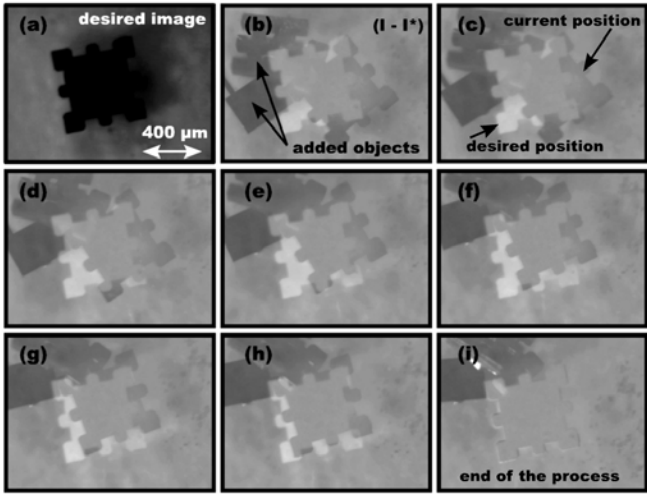
Fig. 6. Image (a) represents the initial position of the micro-object and the images (b) to (i) show the error ($\mathbf{I}$ - $\mathbf{I}^*$) between the current image $\mathbf{I}$ and the desired image $\mathbf{I}^*$.
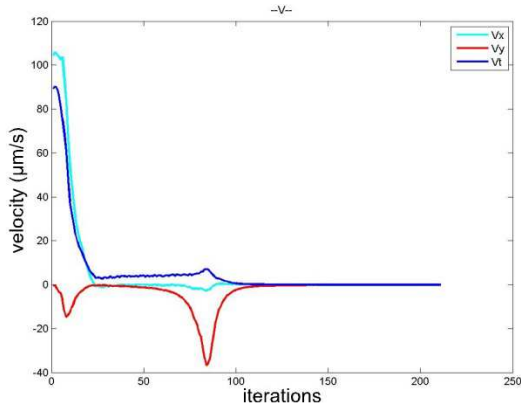

Fig. 7. Platform velocities q_x, q_y, and q_θ for experiment shown on Figure 6.

Figure 8 shows a sequence of images captured during the micromanipulation process. The figure 8.(a), represents the desired image which consists of the image of the micro-object inserted in a white image. Figures 8.(b) to 8.(g) represent the image error ($\mathbf{I}$ - $\mathbf{I}^*_m$) during the visual servoing task.

Despite the presence of more than half of white pixels (virtual pixels) in the desired image $\mathbf{I}^*_m$ which have not corresponding pixels in the current image, the proposed control law remains robust and converge to zero as shown in the figure 9. By analyzing the behavior of the control law, we remark that there is a presence of small oscillations at the beginning because of the large number of pixels in the current image that have no correspondences in the desired image.

D. Example of Application: Multiple Scale Visual Servoing

The last experience presented in this paper consists of the use of the proposed control law in multiple scale configuration. It means that the magnification (zoom) of the
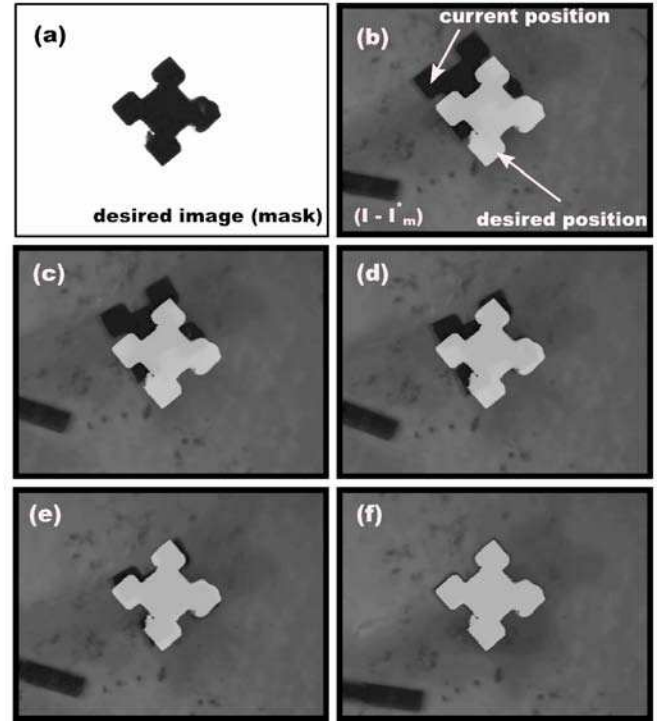

Fig. 8. Sequence image captured during the positioning process using a mask as a desired image (see, image (a)).
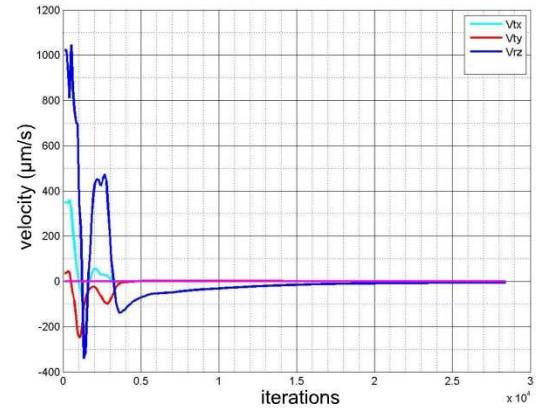

Fig. 9. Platform velocities $\dot{q}_x$, $\dot{q}_y$ and $\dot{q}_\theta$ for experiment shown on Figure 8.
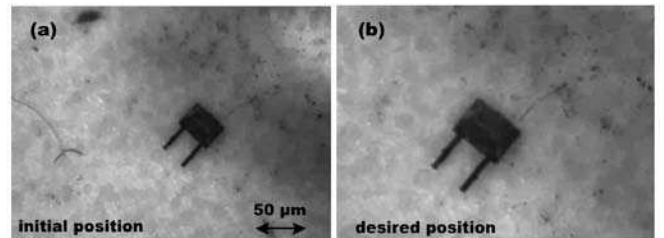

Fig. 10. Images (a) and (b) represent the initial position captured with a low magnification and the desired position acquired with a high magnification, respectively.

microscope is integrated in the control law. Thus, the idea is to use only a single view and then to resolve the problem

of the limiting FOV which characterizes the optical micro-scopes. To do this, the desired image is captured with a high magnification, i.e $\times$ 9 (low resolution/low FOV) (Fig. 10.b), and the visual servoing starts with a low magnification, i.e. $\times$ 2 (high resolution/high FOV) (see Fig. 10.a). After that, the magnification increases during the realization of the task until to attain the desired magnification in which the desired image is acquired.
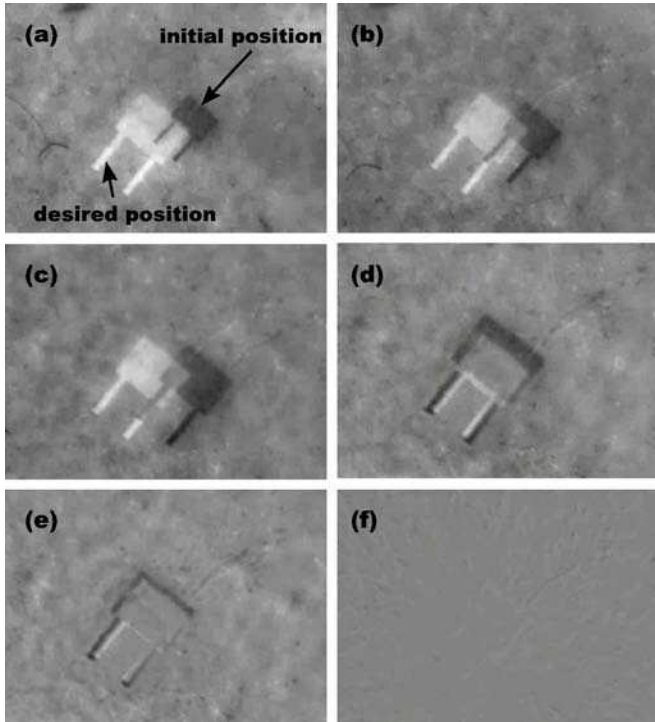

Fig. 11. Representation of the error ($\mathbf{I}$ - $\mathbf{I}^*$) during the positioning task and the magnification increasing in the same time.

Figure 11 illustrates the achievement of the multiple scale visual servoing. From the figure 11.(a) to figure 11.(d), it can be seen the error ($\mathbf{I}$ - $\mathbf{I}^*$) between the desired image and the current image as the evolution of the magnification factor during the positioning task (see, also Fig. 12).
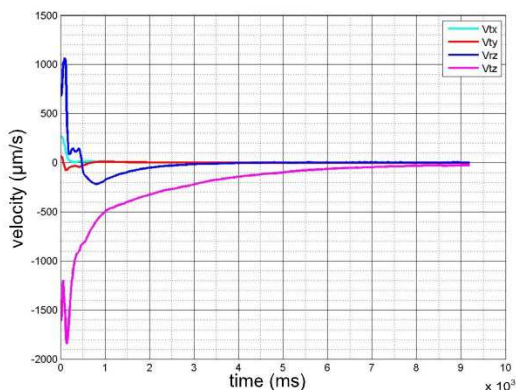

Fig. 12. Platform velocities $\dot{\mathbf{q}}_\mathbf{x}$, $\dot{\mathbf{q}}_\mathbf{y}$ and $\dot{\mathbf{q}}_\theta$ for experiment shown on Figure 11 along with the motion of the camera zoom velocity.

VI. CONCLUSION

The problem of the automation of micromanipulation and microassembly of MEMS using a direct vision-based control approach was studied in this paper. Classical visual servoing schemes use geometrical visual features (points, lines, edges, ...) extracted and tracked in the image sequence. This kind of approaches require the use of image processing algorithms to extract these features from the images. It is very time consuming and complex to set up and often not robust to external conditions modification. Therefore, we have proposed a new control law based on the use of the pure image signal (illumination signal). This photometric approach has been tested on our microassembly workcell. From the different validation tests of the proposed control law, it appears that it is robust to global light variations, to important occlusions and to different perturbations. The proposed approach allows highly precise positioning task precision (below 100 nanometers in translation and 0.001 degrees in rotation along the platform axis). This process have also been used to perform a multiple scale visual servoing. So, the first experimental results are promising in terms of precision and of the designed control law behavior. Future work will concern the use of these techniques to auto-mate the complete MEMS assembly process which includes, positioning, orientation, picking, placing, and insertion of the different elements of the MEMS.

REFERENCES

[1] F. Chaumette and S. Hutchinson, *Visual servo control, part 1: Basic approaches*, IEEE Rob. and Aut. Mag. **13** (2006), no. 4, 82–90.
[2] C. Collewet, E. Marchand, and F. Chaumette, *Visual servoing set free from image processing*, IEEE Int. Conf. on Rob. and Aut., (Pasadena, California), May 2008, pp. 81–86.
[3] A. Dame and E. Marchand, *Entropy-based visual servoing*, IEEE Int. Conf. on Rob. and Aut., (Kobe, Japan), 2009, pp. 707-713, .
[4] B. Espiau, F. Chaumette, and P. Rives, *A new approach to visual servoing in robotics*, IEEE Trans. on Rob. and Aut., **8**, (1992), no. 3, pp. 313–326.
[5] B.K.P. Horn and B.G. Schunck, *Determining optical flow*, Artificial Intelligence **17** (1981), no. 1-3, pp. 185–203.
[6] S. Hutchinson, G. Hager, and P. Corke, *A tutorial on visual servo control*, IEEE Trans. on Rob. and Aut., **12**, (1996), no. 5, pp. 651–670.
[7] E. Marchand, *Control camera and light source positions using image gradient information*, IEEE Int. Conf. on Rob. and Aut., (Rome, Italy), 2007, pp. 417-422.
[8] M. Probst, R. Borer, and B. J. Nelson, *A microassembly system for manufacturing hybrid mems*, IFTOMM World Congress (Besançon, France), June 2007.
[9] S. Shet, V. R. Mehta, A. F. Fiory, M. P. Lepselter, and N. M. Ravindra, *The magnetic field-assisted assembly of nanoscale semiconductor devices: A new technique*, Jo. of materials (JOM) **56(10)**, (2004), pp. 32–34.
[10] B. Tamadazte, E. Marchand, S. Dembélé, N. Le Fort-Piat. CAD model based tracking and 3D visual-based control for MEMS microassembly. The Int. Journal of Robotics Research, IJRR, doi: 10.1177/0278364910376033, 2010.
[11] K. Yesin and B. Nelson, *A cad model based tracking system for visually guided microassembly*, Robotica **23** (2005), pp. 409–418.