Characterization of Model-based Visual Tracking Techniques for MOEMS using a New Block Set for MATLAB/Simulink

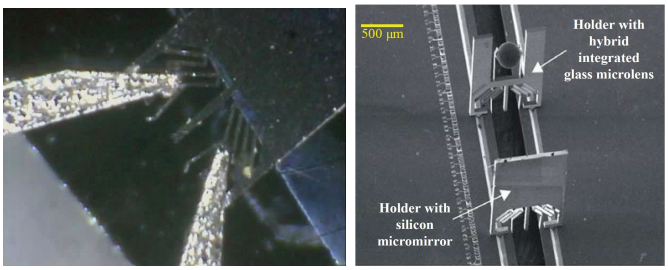
Andrey V. Kudryavtsev, Guillaume J. Laurent, Cédric Clévy, Brahim Tamadazte and Philippe Lutz
FEMTO-ST, AS2M, Université de Franche-Comté/CNRS/ENSMM, Besançon, France.
Email: Andrey.Kudryavtsev@femto-st.fr, Guillaume.Laurent@ens2m.fr, Cedric.Clevy@femto-st.fr

***Abstract*—Microassembly is an innovative alternative to the microfabrication process of MOEMS which is quite complex. It usually implies the usage of microrobots controlled by an operator. The reliability of this approach has been already confirmed for the micro-optical technologies. However, the characterization of assemblies has shown that the operator is the main source of inaccuracies in the teleoperated microassembly, so there is a great interest in automating the microassembly process. One of the constraints of automation in microscale is the lack of high precision sensors capable to provide the full information about the object position. Thus, the usage of visual-based feedback represents a very promising approach allowing to automate the microassembly process. The purpose of this paper is to characterize the techniques of object position estimation based on the visual data, i.e. visual tracking techniques from the ViSP library. These algorithms allows to get the 3D object pose using a single view of the scene and the CAD model of the object. The performance of three main types of model-based trackers is analyzed and quantified: edge-based, texture-based and hybrid tracker. The problems of visual tracking in microscale are discussed. The control of the micromanipulation station used in the framework of our project is performed using a new Simulink block set. Experimental results are shown and demonstrate the possibility to obtain the repeatability below 1 μm.**

I. Introduction

The needs for highly integrated, complex and smart MOEMS keep increasing especially for instrumentation and biomedical tools application fields [1]–[4]. The main lock for these developments relies in microfabrication process that are quite complex. To tackle this lock, several innovative alternatives are emerging, among them the use of microrobots to achieve robotized microassembling of elementary blocks.

Several works already established the viability of this approach and notably show that the key feature relies in the capability of the system to achieve modular and highly accurate assembling, i.e. typically smaller than 5 μm (maximum acceptable error) [5], [6]. Several concepts of microoptical benches to be assembled have also been proposed [7]–[9]. In [10], it is also established that a positioning accuracy smaller than 1 μm can be achieved in teleoperated mode. The operator being the main source of inaccuracies [10], there is a great interest in automating the micro-assembly process: increasing the yields throughput but also quantifying the main sources of inaccuracies which is of great interest for the design of MOEMS blocks and microrobots, the clean room fabrication and assembly strategies.



(a) (b)

Fig. 1: a) Microcomponent picking using the microgripper; b) example of assembled microoptical bench [10].

The main objective of the described works is to achieve automated assemblies of MOEMS through visual servoing approach and than study and characterize its precision, repeatability, etc. To achieve this goal, a strategy based on high level closed-loop vision control will be implemented. The studied methods are model-based visual tracking algorithms from the ViSP library, which is able to directly provide the 3D object pose using a single view of the scene [11].

The components to be assembled are displayed in Fig. 1. The assembly consists in insertion of holder in the V-groove guiding rails of silicon baseplate using a micromanipulation station that will be described in Section II.

Previous works on the field of automatic microassembly demonstrate the possibility of model-based visual trackers application [12]. However, these strategies can not be employed directly in our case, because of several constraints. Firstly, all objects of the scene are made of silicon which causes the reflections. Secondly, the object contains the deformable parts. Finally, the ratio between length and thickness is very high. Thus, the goal of this paper is to implement CAD-based visual tracking techniques and to quantify its performances for further automation of microassembly. For this, first of all, we should discriminate the noises : related to the algorithm, to the environmental factors and to the robot structure. Than, we will quantify the tracker precision by using a 3D trajectory. To simplify the analysis of microassemblies, the software parts (micromanipulation control, visual tracking) will be developed in the form of Simulink blocks in C++ language.

The remainder of this paper is organized as follows :

TABLE I: Characteristics of the stages comprised in the robot used in a micromanipulation station.

	Product reference	Specifications
Translation stages : XY	M-111-DG *PI Mercury*	Stroke : 15 mm Backlash : 2 µm Min. Inc. Motion : 0.05 µm Unidir. repeatability : 0.1 µm
Z	M-121-DG *PI Mercury*	Stroke : 25 mm Backlash : 2 µm Min. Inc. Motion : 0.05 µm Unidir. repeatability : 0.1 µm
Rotation stage : Θ	SR3610S *SmarAct*	Stroke : 360° Resolution : $< 10\mu$ °

TABLE II: Characteristics of the vision system.

	Product reference	Specifications
Camera : IDS uEye	UI-3480CP *Aptina*	CMOS Rolling Shutter Pixelpitch : 2.2 µm Pixel Class : 5 Megapixel Resolution (h x v) : 2560x1920
Objectif	CVO GM10HR35028	Class : high resolution Focal distance : 50 mm

Section II presents the used equipment, micromanipulation station in particular. In Section III, we present the software architecture allowing to interface different equipment such as translation and rotation stages, camera, etc. Sections IV to VI of this article describe the steps of trackers analysis and performances quantification. Finally, conclusions and prospects are discussed at the end.

II. Experimental setup

For the precise control of position and the alignment of optical path, we use the 3D microassembly station that comprises a serial robot of 4 degrees of freedom (XYZΘ) with a microgripper and a vision system (Fig. 2). The whole system is mounted on an antivibration table. The characteristics of the robot and vision system are represented in Tables I and II respectively.

One of the most important parameters, that was considered during the choice of the vision system, is a depth of field, because it is indispensable to have a focused image in the whole area of object displacement. The depth of field of presented vision system is about 2 mm.

The position of the camera in relation to the robot is represented in Fig. 3.

III. Simulink blocks development

Simulink is a data flow graphical programming language tool for modeling, simulating and analyzing multidomain dynamic systems. Its primary interface is a graphical block diagramming tool and a customizable set of block libraries. It is widely used in automatic control, aerospace engineering, signal processing and computer engineering applications [13].

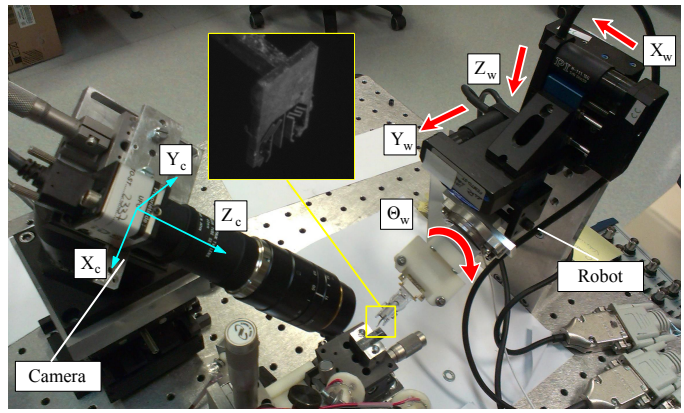
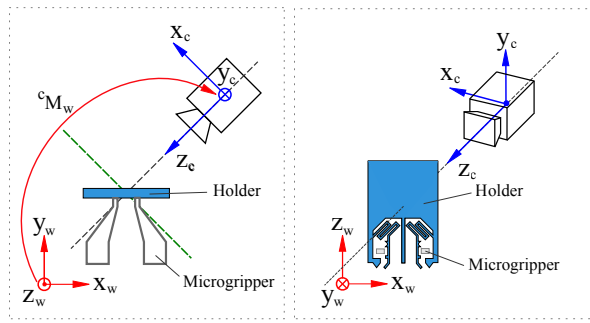


Fig. 2: Configuration of the micromanipulation station with four DOF.



(a) Top view (b) Side view

Fig. 3: Camera position related to the object; x_c, y_c and z_c refer to the coordinates in the camera frame; x_w, y_w and z_w refer to the world frame;

In many fields, MATLAB/Simulink has already become the number one simulation language. Furthermore, the creation of custom Simulink blocks is possible using a MATLAB S-function. Custom code may be written in M, in C or in Fortran languages.

The development of software in the form of Simulink blocks has several advantages. Firstly, it allows creating one block which exercises one function, thus, the system becomes modular. Secondly, the fact of working in a Simulink environment allows the usage of all already existing blocks, such as mathematical functions, scopes, advanced control laws, filters, etc. Thirdly, it became possible to use the dedicated functions such as "Robotic toolbox" developed by Peter Corke; it is a software package that allows a MATLAB user to readily create and manipulate data types fundamental to robotics such as homogeneous transformations, quaternions and trajectories [14]. Finally, as MATLAB/Simulink is widely used in a field of engineering, the developed blocks can be easily integrated in other projects. For all these reasons, we have chosen to develop the Simulink blocks for robot control and vision.

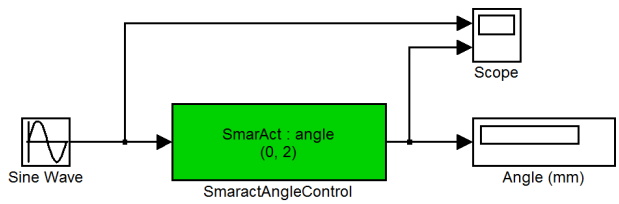


Fig. 4: Usage example of a block of rotation stage angle control

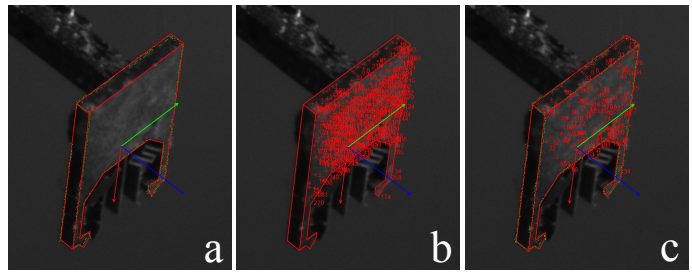


Fig. 6: Different types of tracking : a) edge-based; b) texture-based; c) hybrid.

A. C++ wrapper for S-function

The development of Simulink blocks using standard C++ language had become possible owing to the EasyLink interface. It was developed in the FEMTO-ST institute and it allows to use C++ compiler in order to create Simulink blocks. The fact of working with C++ language simplifies the usage of external libraries.

B. Blocks for Robot Control

In order to integrate the possibility of robot control in MATLAB/Simulink, a custom block set has been developed. It contains the blocks allowing the entire control of SmarAct and PI Mercury stages [15] [16]:

- two types of blocks for a stage : one block for position control and one block of speed control. There is a possibility to impose the displacement limits for every movement which allows to secure the station, in particular fragile components such as microgripper;
- one block for microgripper control. It allows to control the two fingers of microgripper either simultaneously or separately;
- one block which provides access to a joystick. It provides a possibility of teleoperation control of the station. The code of this block is based on SDL 2.0 (Simple DirectMedia Layer) library [17].

Fig. 4 represents an example of control of SmarAct stage.

The block set contains also the blocks of 3D model-based tracking that will be described in the following section. All of the blocks of the library are represented in Fig. 5.

C. Blocks for Tracking

In order to implement the tracking techniques, the C++ ViSP library was used. ViSP standing for Visual Servoing Platform is a modular cross platform library created by Inria Lagadic team that allows prototyping and developing applications using visual tracking and visual servoing techniques [11].

The tracking techniques studied in this article use a 3D model of the tracked object to provide 3D information of the object position from a monocular image. They can be divided into three main classes described in [18]:

- Edge-based tracking. It relies on the high spatial gradients outlining the contour of the object or some geometrical features;
- Texture-based tracking, which uses texture information for object tracking;
- Hybrid tracking. It is appropriate to track textured objects with visible edges.

For every type of tracker one block was created, thus, three blocks.

IV. Trackers analysis in a still frame

The first stage of pose estimation analysis consisted in tracking of the object in a still frame (Fig. 6). It should be noted that even with a still frame, neither the edge-based tracking nor the hybrid method are not perfectly stable. The tracking results obtained using the blocks displayed in previous section are represented in the form of standard deviation in Table III.

TABLE III: Standard deviation of position measurement for different trackers in a still frame for 1000 iterations in the camera frame.

Coordinates	Edge-based	Texture-based	Hybrid
X	0.031 µm	0 µm	0.117 µm
Y	0.059 µm	0 µm	0.010 µm
Z	3.204 µm	0 µm	0.760 µm
roll	0.0071°	0°	0.0194°
pitch	0.0068°	0°	0.0152°
yaw	0.0014°	0°	0.0252°

The noise of X and Y translation coordinates does not exceed 120 nm which represents an error of less than 30% of the image pixel size. This noise level represents a maximal limit of tracker performance for our application. The predominant noise on the Z coordinate can be explained by the fact that the focal distance is considerably higher than the size of the sensor, and the pinhole projection model becomes close to a parallel one, thus, it is more difficult to estimate the depth coordinate. The problem of depth coordinate estimation with long focal distance has also been established for SEM environments where assumption of parallel beams is close to reality [19] [20].

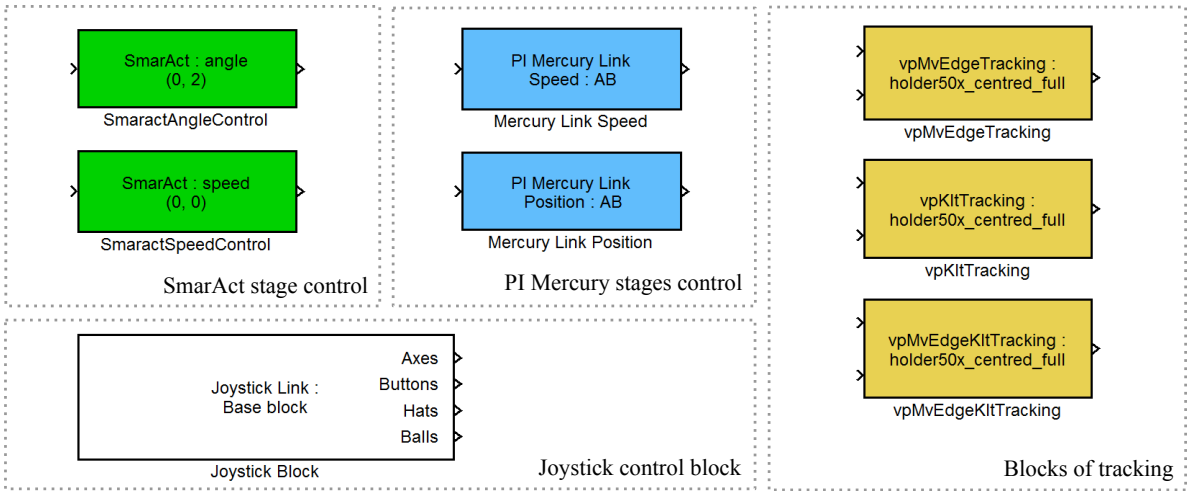


Fig. 5: Overview of the blocks of simulink library developed to control the micromanipulation station in teleoperated or automated mode

V. Trackers analysis with a static object

In this section, the object pose is recorded while the robot is not moving. This experiment allows to quantify the influence of the environmental factors such as luminosity and temperature change during the experiment, vibrations of axes of robot and of camera support, human presence in the room, etc.

The results (Table IV) were obtained from the video stream with the object fixed in the microgripper and the pose is recorded over 1000 images with 30 frames per second.

TABLE IV: Standard deviation of position measurement for different trackers with a static object in the camera frame.

Coordinates	Edge-based	Texture-based	Hybrid
X	0.2245 µm	0.7861 µm	0.2551 µm
Y	0.7026 µm	0.8650 µm	0.6207 µm
Z	24.3304 µm	44.1736 µm	14.8640 µm
roll	0.0859°	0.1607°	0.0610°
pitch	0.0577°	0.1254°	0.0539°
yaw	0.0608°	0.0984°	0.0461°

The results confirm the observations made before concerning the imprecision of the measurements along the Z axis of camera frame. The standard deviation of measured coordinates X and Y has become bigger comparing to the results with a still frame. However, this level of noise is not suprising. The order of magnitude is similar to the results of noise characterization in millimeter sized micromanipulation systems presented in [21]: the vibration at the free end of the cantilever of 30 mm subject to the environmental noise on an anti-vibration table and with human activity reaches 123.7 nm. Thus, since the distance between the microgripper tip and the point of robot mount is about 20 cm in our system, it may explain that the noise level can reach 700 nm.

VI. Trajectory tracking

Next step of tracking analysis consists in comparing the measurements between proprioceptive robot sensors and the visual sensor. In this section we will use the hybrid tracker to estimate the position of the object, which gave the best results during the previous experiments. The robot control is performed by applying a position set-point to the robotic axis and in order to simplify the analysis, the rotations are not taken into account. The comparison is performed by using the trajectory represented in Fig. 7, where the curves are represented in the world frame.

Thereafter, we use the following notations :

- i - image number;
- $\mathcal{R}_c$ - camera frame;
- $\mathcal{R}_w$ - world frame;
- $\mathbf{t}_c(i) = \{x_c\, y_c\, z_c\}^T$ - object translations in the camera frame obtained with the tracker;
- $\mathbf{t}_w(i)$ - object translations in the world frame (proprioceptive robot sensors);
- ${}^c\mathbf{t}_w(i) = \{x_w\ y_w\ z_w\}^T$ - object translations in the world frame transformed to the camera frame;

In order to be able to compare the results, it is necessary to transform object coordinates in the world frame $\mathbf{t}_w(i)$ to the camera frame ${}^c\mathbf{t}_w(i)$. This transformation can be represented by a matrix ${}^c\mathbf{M}_w$ which contains the information about frame rotations and translations:

$$ {}^c\mathbf{t}_w(i) = {}^c\mathbf{M}_w \mathbf{t}_w(i) \tag{1} $$

The ${}^c\mathbf{M}_w$ matrix, which represents the extrinsic parameters of the camera, is not known. In order to estimate it, an optimization algorithm was used. The goal of optimization consists in minimization of the distance $\mathbf{d}$ between the coordinates obtained from the tracker $\mathbf{t}_c$ and the sensor values of robot axis ${}^c\mathbf{t}_w$.

$$ \mathbf{d}(i) = \mathbf{t}_c(i) - {}^c\mathbf{M}_w \mathbf{t}_w(i) \tag{2} $$

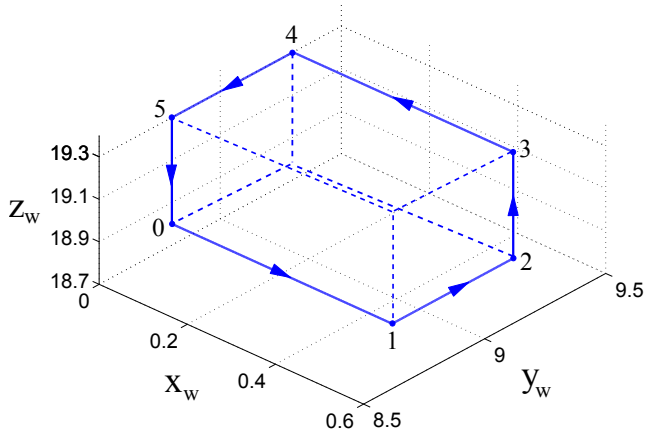


Fig. 7: 3D Trajectory in the world frame to be tracked (in mm).

TABLE V: Standard deviation error obtained between visual hybrid tracker and robot sensors in $\mathcal{R}_c$

Coordinates	Standard deviation error
Δx	2.8165 µm
Δy	4.0918 µm
Δz	215.8632 µm

The optimization criteria is then defined as a sum of squared distance between two curves:

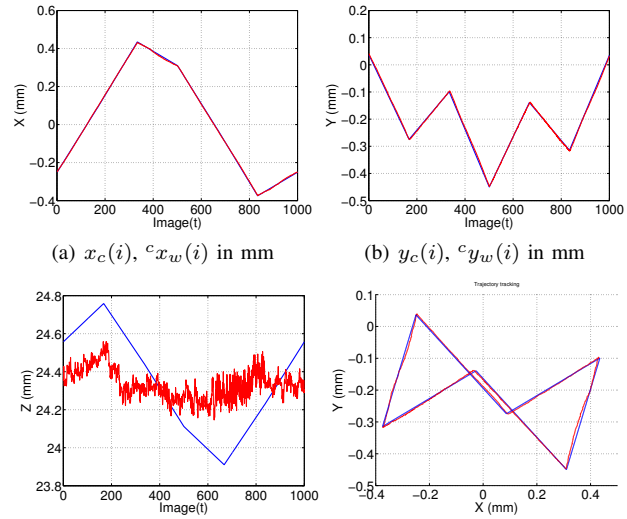
$$J = \sum_{i=1}^{n} \mathbf{d}^T(i)\, \mathbf{A}\, \mathbf{d}(i) \tag{3}$$

The previous analysis of obtained coordinates deviation between two curves allows to notice that Z_c coordinate estimated by the tracker is not only affected by noise but contains no real information about the object position, which means that any filter would be useless. So, in order to minimize its influence, the coefficient of 0.001 was applied for the Z coordinate.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.001 \end{pmatrix}$$

The used optimization algorithm is a Levenberg-Marquardt algorithm that is implemented in the Mathworks Optimization Toolbox [22] which can be used directly due to the fact that we are working on MATLAB/Simulink environment. It was programmed to take the best fit on 20 optimizations from random initial transforms. The results are represented in Fig. 8 and in Table V.

The standard deviations of errors between the tracker and the robot joint coordinates attain now 2.8 µm along X and 4 µm along Y axis. These deviations include intrinsic robot positioning errors which can typically reach several µm of amplitude. For example, M-111-DG translation stages have 150 µrad (Physik Instrumente company datasheet) yaw deviation which in our case (20 cm long robot end-effector) may induce an error of 3 µm. Tan et al. have also recently



(a) $x_c(i)$, ${}^c x_w(i)$ in mm (b) $y_c(i)$, ${}^c y_w(i)$ in mm

(c) $z_c(i)$, ${}^c z_w(i)$ in mm (d) Trajectory in a XY plane

Fig. 8: Experimental comparison between the results from the visual sensor $\mathbf{t}_c$ (red) and from robot sensors ${}^c\mathbf{t}_w$ (blue) in $\mathcal{R}_c$ using hybrid tracker;

shown that this kind of robot have typical positioning accuracy of a few µm and that the main sources of positioning errors are backlash, perpendicularity errors and yaw deviations [23].

VII. Conclusion

The creation of Simulink blocks in a standard programming language C++ allows to simplify the work with the equipment : program becomes graphic and modular and all already existing tools of MATLAB/Simulink can be used.

The problems of visual tracking in a micro-scale were studied through this paper. The analysis of trackers in different situations allowed to estimate the instability of tracking algorithm, environmental factors influence, precision of robot axis. In our setup, the performance of the algorithm can attain 0.12 µm, the noise level due the environment is about 0.7 µm. Finally, the robot axis imperfections add about 2 µm of uncertainty. Thus, the sum gives a value of 2.82 µm that corresponds to the results in Table V. Obtained results prove that it is possible to have a precision better than 1 µm for X and Y coordinates using the visual tracking techniques presented in this paper. It meets the specifications declared for assembly accuracy and represents a great interest for multi-Dof position measurements at the microscale. These vision tracking methods of object position estimation can be directly applied to many kinds of applications such as in the MEMS field. For example, they could be used to measure the position of the micromirror presented in [24] where sensor integration is strongly limited by available space constraints.

The trackers performance may be improved by using a texture or a periodic structure on the object sides.

Through experimental results, we also identified a non-validity of data obtained from the tracker concerning the Z_c axis of camera which is due to the fact that the focal distance is considerably higher than the size of a sensor. There are

several approaches allowing to compensate this effect. First of all, it is possible to use robot sensors and reconstruct the Z_c coordinate, however, the estimation will be limited by stages precision. Secondly, a second camera may be used. However, the stereoscopic vision setup must be calibrated and for three-dimensional reconstruction to be possible, the location and orientation of the cameras at different capture instants must be accurately known [25]. Finally, the system may be completed by unidirectional sensors capable to estimate Z_c coordinate such as laser sensors.

ACKNOWLEDGMENT

These works have been funded by the Franche-Comté region, partially supported by the Labex ACTION project (contract "ANR-11-LABX-01-01") and by the French RENATECH network through its FEMTO-ST technological facility.

REFERENCES

[1] D. Tolfree and M. J. Jackson, *Commercializing micro-nanotechnology products*. CRC Press, 2010.

[2] W. Noell, W. Sun, N. de Rooij, H. P. Herzig, O. Manzardo, and R. Dandliker, "Optical mems based on silicon-on-insulator (soi) for monolithic microoptics," vol. 2, pp. 580–581, 2002.

[3] F. S. Chau, Y. Du, and G. Zhou, "A micromachined stationary lamellar grating interferometer for fourier transform spectroscopy," *Journal of Micromechanics and Microengineering*, vol. 18, no. 2, p. 025023, 2008.

[4] R. Syms, H. Zou, and J. Stagg, "Micro-opto-electro-mechanical systems alignment stages with vernier latch mechanisms," *Journal of Optics A: Pure and Applied Optics*, vol. 8, no. 7, p. S305, 2006.

[5] N. Dechev, W. L. Cleghorn, and J. K. Mills, "Microassembly of 3-d microstructures using a compliant, passive microgripper," *IEEE Journal of Microelectromechanical Systems*, vol. 13, no. 2, pp. 176–189, 2004.

[6] J. Agnus, N. Chaillet, C. Clévy, S. Dembélé, M. Gauthier, Y. Haddab, G. Laurent, P. Lutz, N. Piat, K. Rabenorosoa *et al.*, "Robotic microassembly and micromanipulation at femto-st," *Journal of Micro-Bio Robotics*, vol. 8, no. 2, pp. 91–106, 2013.

[7] S. Bargiel, K. Rabenorosoa, C. Clevy, C. Gorecki, and P. Lutz, "Towards micro-assembly of hybrid moems components on a reconfigurable silicon free-space micro-optical bench," *Journal of Micromechanics and Microengineering*, vol. 20, no. 4, p. 045012, 2010.

[8] A. N. Das, J. Sin, D. O. Popa, and H. E. Stephanou, "On the precision alignment and hybrid assembly aspects in manufacturing of a microspectrometer," pp. 959–966, 2008.

[9] K. Aljasem, L. Froehly, A. Seifert, and H. Zappe, "Scanning and tunable micro-optics for endoscopic optical coherence tomography," *IEEE Journal of Microelectromechanical Systems*, vol. 20, no. 6, pp. 1462–1472, 2011.

[10] C. Clévy, I. Lungu, K. Rabenorosoa, and P. Lutz, "Positioning accuracy characterization of assembled microscale components for micro-optical benches," *Assembly Automation*, vol. 34, no. 1, pp. 69–77, 2014.

[11] E. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robotics & Automation Magazine*, vol. 12, no. 4, pp. 40–52, 2005.

[12] B. Tamadazte, E. Marchand, S. Dembélé, and N. Le Fort-Piat, "Cad model-based tracking and 3d visual-based control for mems microassembly," *The International Journal of Robotics Research*, 2010.

[13] D. Xue and Y. Chen, *System Simulation Techniques with MATLAB and Simulink*. John Wiley & Sons, 2013.

[14] P. Corke, "A robotics toolbox for matlab," *IEEE Robotics & Automation Magazine*, vol. 3, no. 1, pp. 24–32, 1996.

[15] "Smaract official website," 2014, http://www.smaract.de/.

[16] "Physik instrumente official website," 2014, http://www.physikinstrumente.com/.

[17] S. Lantinga, "Simple directmedia layer," 2014, http://www.libsdl.org/.

[18] M. Pressigout and E. Marchand, "Real-time hybrid tracking using edge and texture information," *The International Journal of Robotics Research*, vol. 26, no. 7, pp. 689–713, 2007.

[19] L. Cui and E. Marchand, "Calibration of scanning electron microscope using a multi-images non-linear minimization process," in *IEEE International Conference on Robotics and Automation*, 2014.

[20] L. Cui, E. Marchand, S. Haliyo, and S. Régnier, "6-dof automatic micropositioning using photometric information," in *IEEE International Conference on Advanced Intelligent Mechatronics*, 2014.

[21] M. Boudaoud, Y. Haddab, Y. Le Gorrec, and P. Lutz, "Noise characterization in millimeter sized micromanipulation systems," *Mechatronics*, vol. 21, no. 6, pp. 1087–1097, 2011.

[22] S. Prajna, A. Papachristodoulou, and P. A. Parrilo, "Sostools: sum of squares optimization toolbox for matlab–users guide," *Control and Dynamical Systems, California Institute of Technology*, vol. 91125, 2004.

[23] N. Tan, C. Clévy, G. J. Laurent, P. Sandoz, and N. Chaillet, "Characterization and compensation of xy micropositioning robots using vision and pseudo-periodic encoded patterns," *IEEE International Conference on Robotics and Automation*, 2014.

[24] A. Espinosa, X. Zhang, K. Rabenorosoa, C. Clévy, S. R. Samuelson, B. Komati, H. Xie, and P. Lutz, "Piston motion performance analysis of a 3dof electrothermal mems scanner for medical applications," *International Journal of Optomechatronics*, vol. 8, no. 3, 2014.

[25] G. Roth, R. Laganiere, and S. Gilbert, "Robust object pose estimation from feature-based stereo," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 4, 2006.