

MODEL-BASED VS. MODEL-FREE VISUAL SERVOING: A PERFORMANCE EVALUATION IN MICROSYSTEMS

Muhammet A. Hocaoglu, Hakan Bilen, Erol Ozgur, Mustafa Unel
Faculty of Engineering and Natural Sciences
Sabanci University
Orhanli Tuzla 34956 Istanbul, Turkey
email: {muhammet, hakanbil, erol}@su.sabanciuniv.edu, munel@sabanciuniv.edu

ABSTRACT

In this paper, model-based and model-free image based visual servoing (VS) approaches are implemented on a microassembly workstation, and their regulation and tracking performances are evaluated. A precise image based VS relies on computation of the image jacobian. In the model-based visual servoing, the image Jacobian is computed via calibrating the optical system. Precisely calibrated model based VS promises better positioning and tracking performance than the model-free approach. However, in the model-free approach, optical system calibration is not required due to the dynamic Jacobian estimation, thus it has the advantage of adapting to the different operating modes.

KEY WORDS

Visual servoing, Visual tracking, Micropositioning

1 Introduction

Visual servoing is one of the effective methods to compensate the uncertainties in the calibration of systems, manipulators and workspaces. Over the past years, intense research effort in this area has resulted in a number of successful applications. Two major approaches are presented in the visual servoing (VS) literature, position-based and image-based VS [1]-[5]. The first approach is based on reconstruction of 3D model of the object and a calibrated camera to provide feedback in the cartesian space. In the second one, control values are defined in terms of image coordinates and no estimation of robot pose is required. The complex geometry of the observed micro-objects and high numerical apertures of optical microscope which results in small depth of field lead to a challenging 3D construction and pose estimation problem. Therefore, an image based approach is preferred in our micro visual servoing experiments since it does not require an inverse perspective projection.

In this paper, model-based and model-free visual servoing approaches are experimentally tested in point-to-point positioning and trajectory following tasks. Since the accuracy of image based VS depends on the computation of the image Jacobian matrix, which relates the changes in the cartesian pose to the corresponding changes in the visual

features, includes the intrinsic and extrinsic parameters of the microscope-camera system. Thus, the calibration information is vital for computation of the image Jacobian matrix and thus the control design. On the other hand, model-free visual servoing does not require a priori information of the (robot + optical) system since the composite Jacobian, i.e. product of robot and image Jacobians, is estimated dynamically [6]. Thus, model-free visual servoing approach eliminates the dependence to the system parameters.

The paper is organized as follows: Section 2 defines image based model-free and model-based visual servoing along with controller synthesis. Section 3 introduces hardware setup and real-time tracking algorithm, and presents experimental results and discussions. Finally, Section 4 concludes the paper with some remarks.

2 Image Based Visual Servoing

Image based visual servoing approaches employ the following differential relation

$$\dot{s} = J\dot{r} \quad (1)$$

where s is a vector of visual features, J is the image Jacobian matrix which is a function of the visual features and intrinsic/extrinsic parameters of the visual sensor, and \dot{r} is a velocity screw in the task space.

Depending on the computation of the Jacobian matrix, one can talk about model-based or model-free visual servoing strategies. In the sequel, we will review these approaches.

2.1 Model Based Visual Servoing

Model based visual servoing implies analytical computation of the Jacobian matrix through the calibration of the optical system.

To develop an analytical model of the Jacobian for calibration purposes, let the objective frame coordinates of an observed feature point be $P_o = (X_o, Y_o, Z_o)$. Locating the image coordinate frame at the center of the CCD array and assuming weak perspective projection, the undistorted image coordinates (x'_s, y'_s) in objective frame are given as

$$x'_s = MX_o, y'_s = MY_o \quad (2)$$

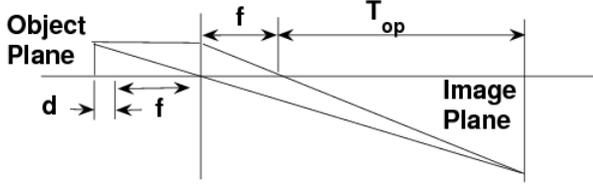


Figure 1. Ray Diagram of the Optical Model

where $M = \frac{T_{op}+f}{f+d}$ is the total magnification of the optical system, f is the objective focal length, T_{op} is the tube length, and d is the working distance, as shown in Fig. 1. Since the lens radial distortion parameter (κ_1) is very small, the distorted image coordinates (x_s, y_s) in pixels can be written as

$$x_s \approx x'_s = \frac{M}{s_x} X_o, \quad y_s \approx y'_s = \frac{M}{s_y} Y_o \quad (3)$$

where s_x and s_y are the effective pixel sizes.

The optical flow equations can be obtained by differentiating (3) with respect to time

$$\dot{x}_s = \frac{M}{s_x} \dot{X}_o, \quad \dot{y}_s = \frac{M}{s_y} \dot{Y}_o \quad (4)$$

Assume that the point P is rigidly attached to the end effector of the manipulator and moves with an angular velocity $\Omega_o = (\omega_x, \omega_y, \omega_z)$ and a translational velocity $V_o = (V_x, V_y, V_z)$. The motion in the objective frame is given by

$$\begin{pmatrix} \dot{X}_o \\ \dot{Y}_o \\ \dot{Z}_o \end{pmatrix} = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} + \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} X_o \\ Y_o \\ Z_o \end{pmatrix} \quad (5)$$

Substituting (5) into (4) and using (3) implies

$$\dot{x}_s = \frac{M}{s_x} V_x + \frac{M}{s_x} Z_o \omega_y - \frac{s_y}{s_x} y_s \omega_z \quad (6)$$

and

$$\dot{y}_s = \frac{M}{s_y} V_y - \frac{M}{s_y} Z_o \omega_x + \frac{s_x}{s_y} x_s \omega_z \quad (7)$$

In light of (6) and (7), the Jacobian matrix is obtained as

$$J = \begin{pmatrix} \frac{M}{s_x} & 0 & 0 & 0 & \frac{M}{s_x} Z_o & -\frac{s_y}{s_x} y_s \\ 0 & \frac{M}{s_y} & 0 & -\frac{M}{s_y} Z_o & 0 & \frac{s_x}{s_y} x_s \end{pmatrix} \quad (8)$$

2.2 Model-Free Visual Servoing

Let θ denote the vector of joint variables of the robot. The error function in the image plane is defined as

$$e(\theta, t) = s(\theta) - s^*(t)$$

where $s^*(t)$ and $s(\theta)$ denote the positions of a moving target and the end-effector at time t , respectively.

Since the system (robot+optical microscope) model is assumed to be unknown, a recursive least-squares (RLS) algorithm [6], main steps of which are briefly summarized below, is used to estimate the composite Jacobian $J = J_I J_R$, where J_I and J_R are the image and the robot Jacobians. Jacobian estimation is accomplished by minimizing the following cost function, which is a weighted sum of the changes in the affine model over time,

$$\varepsilon_k = \sum_{i=0}^{k-1} \lambda^{k-i-1} \|\Delta m_{ki}\|^2 \quad (9)$$

where

$$\Delta m_{ki} = m_k(\theta_i, t_i) - m_i(\theta_i, t_i) \quad (10)$$

where $m_k(\theta, t)$ is an expansion of $m(\theta, t)$, which is the affine model of the error function $e(\theta, t)$, about the k^{th} data point as follows:

$$m_k(\theta, t) = e(\theta_k, t_k) + \hat{J}_k(\theta - \theta_k) + \frac{\partial e_k}{\partial t}(t - t_k) \quad (11)$$

In light of (11), (10) becomes

$$\Delta m_{ki} = e(\theta_k, t_k) - e(\theta_i, t_i) - \frac{\partial e_k}{\partial t}(t_k - t_i) - \hat{J}_k h_{ki}, \quad (12)$$

where $h_{ki} = \theta_k - \theta_i$, the weighting factor λ satisfies $0 < \lambda < 1$, and the unknown variables are the elements of \hat{J}_k .

Solution of the minimization problem yields the following recursive update rule for the composite Jacobian:

$$\hat{J}_k = \hat{J}_{k-1} + (\Delta e - \hat{J}_{k-1} h_\theta - \frac{\partial e_k}{\partial t} h_t) (\lambda + h_\theta^T P_{k-1} h_\theta)^{-1} h_\theta^T P_{k-1} \quad (13)$$

where

$$P_k = \frac{1}{\lambda} (P_{k-1} - P_{k-1} h_\theta (\lambda + h_\theta^T P_{k-1} h_\theta)^{-1} h_\theta^T P_{k-1}) \quad (14)$$

and $h_\theta = \theta_k - \theta_{k-1}$, $h_t = t_k - t_{k-1}$, $\Delta e = e_k - e_{k-1}$, and $e_k = s_k - s_k^*$, which is the difference between the end-effector position and the target position at k^{th} iteration. The term $\frac{\partial e_k}{\partial t}$ predicts the change in the error function for the next iteration, and in the case of a static camera it can directly be estimated from the target image feature vector with a first-order difference.

2.3 Visual Controller Design

Discrete-time equivalent of equation (1) can be written as

$$s(k+1) = s(k) + T J(k) u(k) \quad (15)$$

where $s \in \mathbb{R}^{2N}$ is the vector of image features being tracked, N is the number of the features, T is the sampling time of the vision sensor, and $u(k)$ is the velocity vector of the end effector.

Controller synthesis in this paper is done by optimizing the following cost function

$$E(k+1) = (s(k+1) - s^*(k+1))^T Q (f(k+1) - s^*(k+1))$$

$$+u^T(k)Lu(k) \quad (16)$$

whose solution yields the following control input

$$u(k) = -(TJ^T(k)QTJ(k) + L)^{-1}TJ^T(k)Q(s(k) - s^*(k+1)) \quad (17)$$

where Q and L are adjustable weighting matrices.

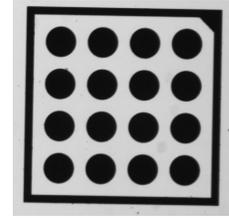


Figure 3. Circular Calibration Pattern

3 Experimental Results and Discussion

The Microassembly Workstation is shown in Fig. 2. It consists of PI M-111.1 high-resolution micro-translation stages with 50 nm incremental motion in x , y and z positioning axes, and is controlled by a dSpace ds1005 motion control board. A Zyvx microgripper, with a 100 μm opening gap is rigidly attached to the translational stage to grasp and pick objects.

Nikon SMZ 1500 stereomicroscope coupled with a Basler A602fc camera, orthogonal to XY plane with $9.9 \mu\text{m} \times 9.9 \mu\text{m}$ cell sizes was utilized to provide visual feedback. The microscope has 1.6X objective and additional zoom. Zoom levels can be varied between 0.75X – 11.25X, implying 15 : 1 zoom ratio.

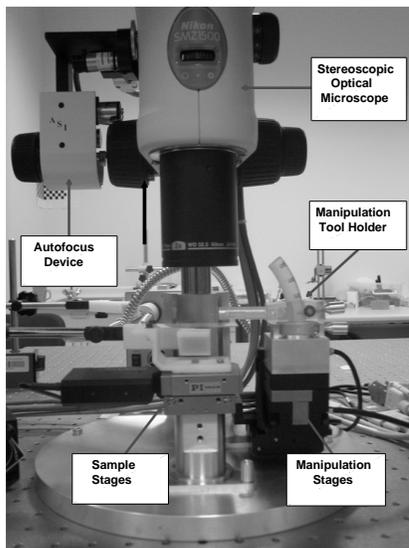


Figure 2. Microassembly Workstation

3.1 Calibration Results

For model-based visual servoing, an accurate calibration of the optical system is required and it was accomplished through a parametric model [7]. A round calibration pattern (Fig. 3) is used to establish the correspondence between the world and image coordinates under 1X and 4X zoom levels. The center coordinates of the circles are calculated through a least square solution.

Computed extrinsic parameters (rotation angles α , β , γ ; components of the translation vector, T_x , T_y , T_z) and intrinsic parameters (total magnification M , objective focal

length f , tube length T_{op} , working distance d and radial distortion coefficient κ_1), and the 3D reprojection errors for the calibration are tabulated in Table 1 and Table 2 respectively. It can be observed from Table 1 that the radial distortion coefficient is very small. This proves that the microscope lenses are machined very precisely. Moreover, β and γ angles have non-zero values which can be resulted from a mechanical tilt of the microscope stage or from an inaccurate design of the calibration pattern.

Table 1. Computed Extrinsic and Intrinsic Parameters

	1X	4X
α (degrees)	90.7144	88.9825
β (degrees)	-2.7912	2.6331
γ (degrees)	175.9179	0.9088
T_x (μm)	-781.4	76.755
T_y (μm)	-55.002	-156.58
T_z (μm)	204900	36370
M	1.5893	6.3859
d (μm)	78750	4955.5
f (μm)	126150	31415
T_{op} (μm)	200490	200610
κ_1 (μm^{-2})	-8.4408×10^{-10}	1.5399×10^{-11}

Table 2. 3D Reprojection Errors for 1X and 4X Zoom

	1X	4X
Mean Error (μm)	0.2202	0.0639
Standard Deviation (μm)	0.3869	0.1321
Maximum Error (μm)	1.7203	0.5843

3.2 Visual Servoing Results

In order to implement visual servoing algorithms real-time measurement of the image features are needed. This is achieved by the ESM algorithm [8], which is based on the minimization of the sum-of-squared-differences (SSD) between the reference template and the current image using parametric models.

Model-based and model-free visual servoing (VS) algorithms were experimentally compared in microposition-

ing and trajectory following tasks at 1X and 4X zoom levels. Micropositioning VS results are plotted in Figs. 4-7, and the trajectory following results for sinusoidal and square trajectories are depicted in Figs. 8-11.

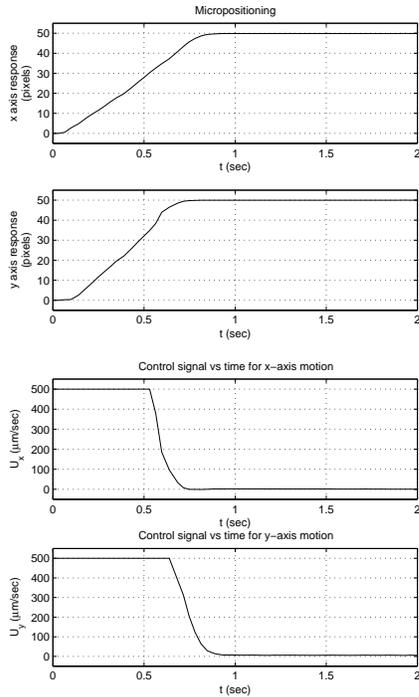


Figure 4. Step responses and control signals of model-based VS at 1X

Regulation performances of both approaches for micropositioning tasks in terms of settling time (t_s), accuracy and precision are tabulated in Table 3. In the trajectory following task, tracking performances of both approaches for square and sinusoidal trajectories are presented in Tables 4-5.

The experimental results illustrate that both of the visual servoing approaches ensure convergence to the desired targets with sub-micron error when time considerations are not primarily important. When the time performance has priority for the task, the model-based, so called calibrated approach performs better than model-free one in terms of settling time, accuracy and precision (Table 3). Moreover, the tracking performance of the calibrated approach is more

Table 3. Micropositioning for model-based and model-free VS

	Step (pix)	Model-based			Model-free		
		t_s (s)	Acc. (μm)	Prec. (μm)	t_s (s)	Acc. (μm)	Prec. (μm)
1x	50	0.80	9.86	2.71	1.6	8.60	3.65
4x	50	0.45	1.35	0.57	1.6	4.74	1.92

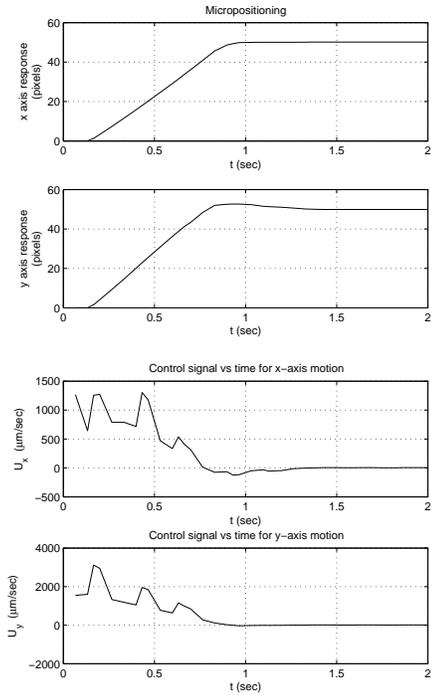


Figure 5. Step responses and control signals of model-free VS at 1X

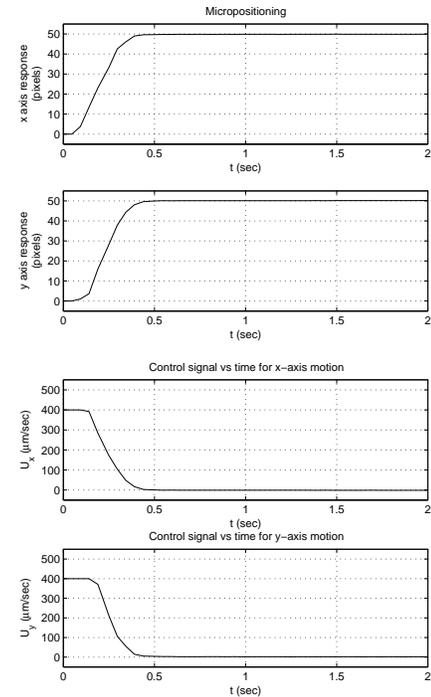


Figure 6. Step responses and control signals of model-based VS at 4X

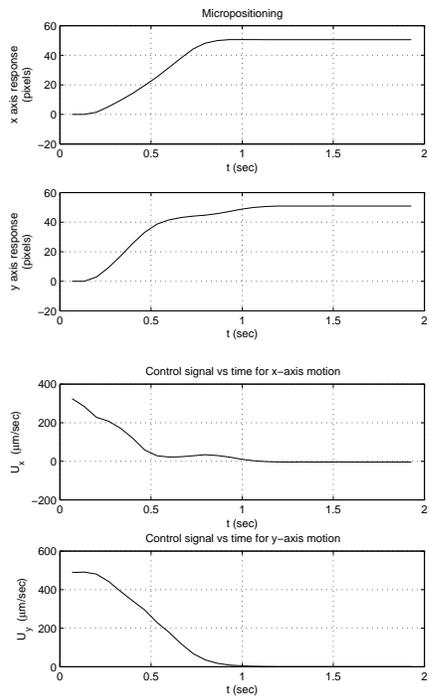


Figure 7. Step responses and control signals of model-free VS at 4X

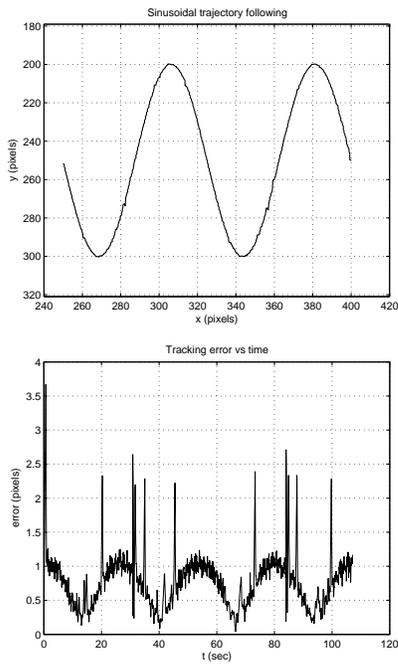


Figure 8. Actual sinusoidal trajectory and resulting tracking error in model-based VS at 1X

accurate and precise than the model-free one. Thus, the calibrated method is more preferable, when accurate and precise manipulation are strongly demanded in a limited time. However, at small magnifications such as $M = 1.5893$ and

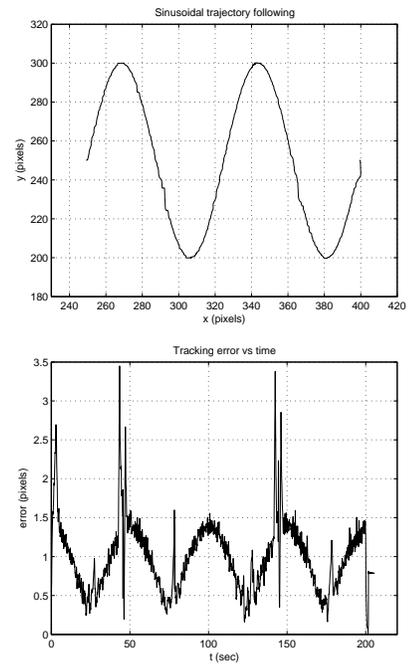


Figure 9. Actual sinusoidal trajectory and resulting tracking error in model-free VS at 1X

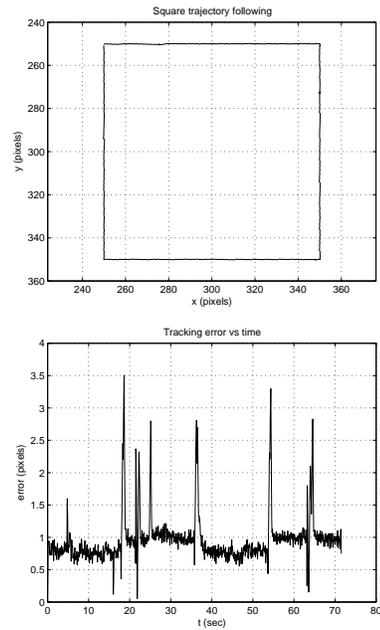


Figure 10. Actual square trajectory and resulting tracking error in model-based VS at 1X

$M = 6.3859$ over a large workspace ($4 \times 3 \text{ mm}^2$), only a coarse microvisual servoing task could be assumed. Therefore, the accuracy and precision of the model-free approach in the regulation and tracking problems are also acceptable, and the difference between two approaches are not that significant.

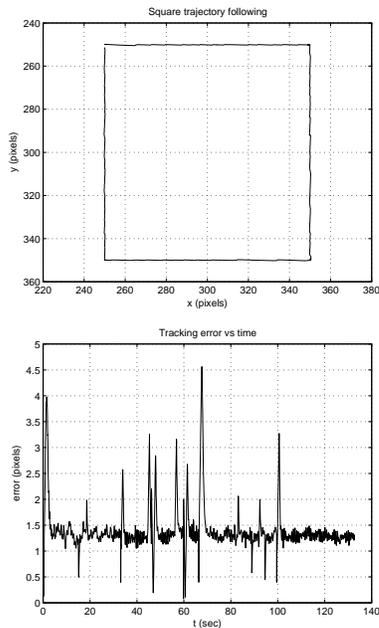


Figure 11. Actual square trajectory and resulting tracking error in model-free VS at 1X

Table 4. Trajectory tracking for model-based VS

	Square		Sinusoidal	
	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)
1x	5.93	2.28	4.79	2.37
4x	1.47	1.19	1.12	1.31

Table 5. Trajectory tracking for model-free VS

	Square		Sinusoidal	
	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)
1x	8.65	2.70	6.14	2.74
4x	1.64	1.12	1.17	0.57

4 Conclusion

Model-based and model-free visual servoing were experimentally evaluated in micropositioning and trajectory tracking tasks. In these experiments, model-based approach performed better in terms of accuracy, precision and settling time than the model-free approach, however, this difference does not necessarily imply a superiority for a coarse manipulation strategy. In addition, the model-free visual servoing is advantageous due to the fact that there is not a requirement of the system model in the implementation of the tasks and it can be adapted to different operating modes through a dynamic estimation of the composite Jacobian.

5 Acknowledgement

Authors gratefully acknowledge the support provided by SU Internal Grant No. IACF06 – 00417.

References

- [1] S. Hutchinson, G. Hager, and P. Corke, A Tutorial on Visual Servo Control, *IEEE Trans. Robotics and Automation*, vol. 12, no. 5, 1996, 651-670.
- [2] F. Chaumette and S. Hutchinson, Visual Servo Control Part 1: Basic Approaches, *IEEE Robotics and Automation Magazine*, December 2006, 82-90.
- [3] D. Kragic and H. I. Christensen, Survey on visual servoing for manipulation. Technical report, *Computational Vision and Active Perception Laboratory, Royal Institute of Technology*, 2002.
- [4] W. J. Wilson, C. C.W. Hulls, and G. S. Bell, Relative end-effector control using Cartesian position-based visual servoing, *IEEE Trans. Robot. Automat.*, vol. 12, 684696, 1996.
- [5] P.I. Corke and S. A. Hutchinson, A new partitioned approach to image-based visual servo control, *Proceedings of the International Symposium on Robotics*, 2000.
- [6] J.A. Piepmeier, Uncalibrated eye-in-hand visual servoing, *The International Journal of Robotics Research*, vol. 22, 2003, 805-819.
- [7] Y. Zhou and B.J. Nelson, Calibration of a parametric model of an optical microscope, *Optical Engineering*, vol.38, 1999, 1989-1995.
- [8] S. Benhimane and E. Malis, Real-time image-based tracking of planes using efficient second order minimization, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, 2004, 943-948.