

A Comparative Study of Conventional Visual Servoing Schemes in Microsystem Applications

H. Bilen, M. Hocaoglu, E. Ozgur, M. Unel, A. Sabanovic

Abstract—This paper presents an experimental comparison of conventional (calibrated and uncalibrated) image based visual servoing methods in various microsystem applications. Both visual servoing techniques were tested on a microassembly workstation, and their regulation and tracking performances are evaluated. Calibrated visual servoing demands the optical system calibration for the image Jacobian estimation and if a precise optical system calibration is done, it ensures a better accuracy, precision and settling time compared with the uncalibrated approach. On the other hand, in the uncalibrated approach, optical system calibration is not required and since the Jacobian is estimated dynamically, it is more flexible.

I. INTRODUCTION

The assembly of microsystems differs from the macroassembly applications due to the high precision requirements and mechanics of microassembly. Although a precision of a few hundred microns is typical for a robotic manipulator in the macro domain, for the applications in the micro domain, submicron precision is required and this degree of precision is beyond the capability of the assembly devices used in the industry. In addition to the different precision requirements, the mechanics of object interactions is different for macro and micro assembly. In macro world, the mechanics of manipulation are predictable to a degree since the forces due to gravity are dominant. However, in the micro world, due to the scaling effects, forces that are not significant in the macro world become dominant [1], [2]. For example, when the parts to be handled are less than one millimeter in size, adhesive forces between gripper and object can be significant compared to gravitational forces. Such issues can be resolved by utilizing real-time visual feedback. In most of the microassembly applications that utilize visual feedback, calibrated visual servoing approach is employed [3], [11].

In this paper, regulation and tracking performances of calibrated and uncalibrated visual servoing are experimentally compared on a microassembly workstation. In the calibrated visual servoing the image Jacobian matrix that 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, it has to be calibrated in order to compute the image Jacobian matrix for the control design. However, as Nelson et al. [8] point out the unique characteristics of the optical microscope introduces new challenges for the calibration. Thus, a different

calibration approach is required to estimate the parameters of the optical system. On the other hand, in model-free or so called uncalibrated visual servoing there is not a requirement of a priori information of the (robot + optical) system since the composite Jacobian, i.e. product of robot and image Jacobians, is estimated dynamically [12]. Since model-free visual servoing does not require the model of the system and adapts itself to the changes in the system configuration, it may provide more flexibility in performing visual tasks in microsystems.

Section II defines image based calibrated and uncalibrated visual servoing along with controller synthesis. Section III presents experimental results and discussions. Finally, Section IV concludes the paper with some remarks.

II. CONVENTIONAL IMAGE BASED VISUAL SERVOING SCHEMES

Image based visual servoing approaches employ the following differential relation

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

where f 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.

The Jacobian matrix can be computed analytically via calibrating the optical system or be estimated dynamically using an adaptive model.

A. Calibrated Visual Servoing

1) *Calibration of the Optical Microscope*: Several calibration methods exist in the literature that are mostly used in macro scale vision applications [4], [5], [6]. However, these methods cannot directly be employed to calibrate an optical microscope coupled with a CCD camera due to the unique characteristics of the optical system. Large numerical apertures and high optical magnifications, and thus very small depth-of-field property of optical microscopes restricts the calibration to a single parallel plane. Modifications to Tsai's and Zhang's algorithms have resulted in several camera calibration algorithms ([7], [8], [9]) for optical microscope and camera systems.

Zhou and Nelson's parametric calibration method [8] was preferred in this work since it was validated by successful experiments in our system. In this method, the complex combination of the image forming elements in the optical pathway is modeled via the objective focal length (f), the tube length (T_{op}), and the distance between the calibration

H. Bilen, M. Hocaoglu, E. Ozgur, M. Unel and A. Sabanovic are with the Faculty of Engineering and Natural Sciences, Sabanci University, 34956 Istanbul, Turkey {hakanbil, muhammet, erol}@su.sabanciuniv.edu, {munel, asif}@sabanciuniv.edu

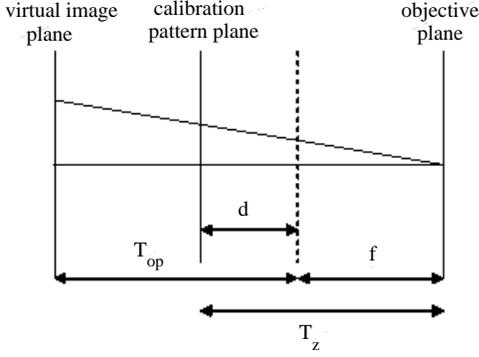


Fig. 1. Ray Diagram of the Optical Model

pattern plane and the front focal plane (d), as shown in Fig. 1.

The first step of the algorithm [8] employs the Radial Alignment Constraint (RAC) from Tsai's algorithm. The RAC gives the three rotation angles (α , β , γ), and T_x and T_y components of the translation vector, T , from the world coordinate frame to the objective coordinate frame. In the second step, a parallel plane assumption is made to obtain an initial estimate of the total magnification (M) of the system and the radial distortion coefficient (κ_1) for performing non-linear optimization. With a manufacturer specified objective focal length (f), the intrinsic parameters T_{op} , f , d , and κ_1 can be determined. The near parallel assumption between the calibration pattern plane and the virtual image plane provides the translation along the optical axis as $T_z = f + d$.

2) *Derivation of the Image Jacobian*: Let (X, Y, Z) denote the objective frame coordinates of an observed feature point P . 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

$$x'_s = MX, y'_s = MY, \quad (2)$$

where $M = \frac{T_{op}+f}{f+d}$ is the total magnification of the optical system.

Neglecting the lens radial distortion parameter (κ_1), the distorted image coordinates (x_s, y_s) in pixels can be written as

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

where s_x and s_y are the effective pixel sizes.

Differentiation of (3) with respect to time implies

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

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

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

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

$$\begin{pmatrix} \dot{x}_s \\ \dot{y}_s \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{M}{s_x} & 0 & 0 & 0 & \frac{M}{s_x} Z & -\frac{s_y}{s_x} y_s \\ 0 & \frac{M}{s_y} & 0 & -\frac{M}{s_y} Z & 0 & \frac{s_x}{s_y} x_s \end{pmatrix}}_{\triangleq J} \begin{pmatrix} V_x \\ V_y \\ V_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \quad (6)$$

where J is the Jacobian for a point feature.

B. Uncalibrated Visual Servoing

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

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

where $f^*(t)$ and $f(\theta)$ 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 [12], 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,

$$\varepsilon_k = \sum_{i=0}^{k-1} \lambda^{k-i-1} \|\Delta m_{ki}\|^2 \quad (7)$$

where

$$\Delta m_{ki} = m_k(\theta_i, t_i) - m_i(\theta_i, t_i) \quad (8)$$

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:

$$m_k(\theta, t) = e(\theta_k, t_k) + \hat{J}_k(\theta - \theta_k) + \frac{\partial e_k}{\partial t}(t - t_k) \quad (9)$$

In light of (9), (8) becomes

$$\Delta m_{ki} = e(\theta_k, t_k) - e(\theta_i, t_i) - \frac{\partial e_k}{\partial t}(t_k - t_i) - \hat{J}_k h_{ki}, \quad (10)$$

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:

$$\hat{J}_k = \hat{J}_{k-1} + (\Delta e - \hat{J}_{k-1} h_\theta - \frac{\partial e_k}{\partial t} h_t) (\lambda + h_\theta^T P_{k-1} h_\theta)^{-1} h_\theta^T P_{k-1} \quad (11)$$

where

$$P_k = \frac{1}{\lambda} (P_{k-1} - P_{k-1} h_\theta (\lambda + h_\theta^T P_{k-1} h_\theta)^{-1} h_\theta^T P_{k-1}) \quad (12)$$

and $h_\theta = \theta_k - \theta_{k-1}$, $h_t = t_k - t_{k-1}$, $\Delta e = e_k - e_{k-1}$, and $e_k = f_k - f_k^*$, 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:

$$\frac{\partial e_k}{\partial t} \approx -\frac{f_k^* - f_{k-1}^*}{h_t} \quad (13)$$

C. Optimal Visual Controller Synthesis

Equation (1) can be written in discrete time as

$$f(k+1) = f(k) + TJ(k)u(k) \quad (14)$$

where $f \in \mathbb{R}^{2N}$ 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.

The aim of the visual servoing tasks in the experiments is to locate the end effector to a constant or time varying desired target $f^*(k)$ by controlling its velocity. A cost function as in [11] is introduced to penalize the pixelized position errors and the control energy as

$$E(k+1) = (f(k+1) - f^*(k+1))^T Q (f(k+1) - f^*(k+1)) + u^T(k) L u(k) \quad (15)$$

The resulting optimal control input $u(k)$ can be derived as

$$u(k) = -(TJ^T(k)QTJ(k) + L)^{-1} TJ^T(k)Q(f(k) - f^*(k+1)) \quad (16)$$

The weighting matrices Q and L can be adjusted to ensure desired response. We should remark that the same optimal control input is used both in calibrated and uncalibrated approaches.

III. EXPERIMENTAL RESULTS

A. Hardware Setup

The experiments were conducted with the microassembly workstation 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 Zyvox 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. Two calibration patterns, Edmund Optics IAM-1 with 50 μm and 200 μm square sizes and Mvtec calibration grid with 70 μm radius circles (Fig. 3) were employed to calibrate the optical system.

B. Calibration Results

Before visual servoing tasks were performed, a sub-micron accurate calibration of the optical system was accomplished through a parametric model [8]. Two different types of calibration patterns were used to establish the correspondence between the world and image coordinates under 1X and 4X zoom levels, hence implying ~ 1.6 and ~ 6.4 magnifications as can be verified from Table I. For the square one, a Sobel

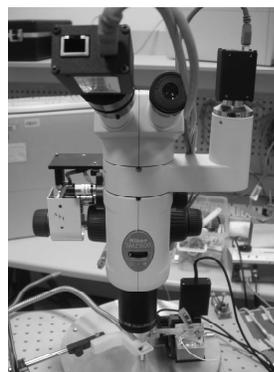


Fig. 2. Microassembly Workstation

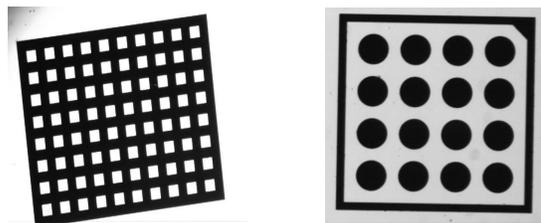


Fig. 3. Square and Circular Calibration Patterns

edge operator, edge linking and then a line fitting algorithm were applied to obtain every edge line of the squares. Corners of the squares -intersections of the calculated edge lines- were taken as the calibration points. For the round calibration grid, the center coordinates of the circles were calculated through a least square solution. Calibration results are tabulated in Table I.

It can be observed from this table 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.

In the experiments it was observed that the circle grids give more accurate calibration results. Due to imperfect illumination, lens aberration, systematic and random sensor errors, the image might be blurred by a point spread function (PSF) and the features might not be extracted very accurately. Flusser and Zitova [10] claim that most of the PSF are circularly symmetric and circular shapes are invariant to this type of PSF. Thus, our experimental results are in accordance with their interpretation as shown in Table II.

C. Real-Time Feature Tracking

Visual servoing algorithms necessitate real-time measurement of the image features in an efficient, accurate and robust manner. Both Kalman filtering and the efficient second-order minimization (ESM) algorithm [13], which is based on the minimization of the sum-of-squared-differences (SSD) between the reference template and the current image using parametric models, were employed in our experiments. The ESM algorithm has high convergence rate like the Newton

TABLE I
COMPUTED INTRINSIC AND EXTRINSIC PARAMETERS USING CIRCULAR
AND SQUARE PATTERNS

	Circular		Square	
	1X	4X	1X	4X
M	1.5893	6.3859	1.6236	6.444
T_{op} (μm)	200490	200610	199690	200880
f (μm)	126150	31415	122990	31174
d (μm)	78750	4955.5	75441	4833.5
κ_1 (μm^{-2})	$-8.4e-10$	$1.5e-11$	$2.1e-10$	$1.5e-10$
α (deg)	90.7144	88.9825	87.2897	95.4143
β (deg)	-2.7912	2.6331	-1.6248	-1.9407
γ (deg)	175.9179	0.9088	177.6637	178.2925
T_x (μm)	-781.4	76.755	-1792.7	-1653.1
T_y (μm)	-55.002	-156.58	-1210.3	-1194.5
T_z (μm)	204900	36370	198430	203610

TABLE II
3D REPROJECTION ERRORS OF CIRCULAR AND SQUARE PATTERNS FOR
1X AND 4X

	Circular		Square	
	1X	4X	1X	4X
Mean Error (μm)	0.2202	0.0639	0.4618	0.0920
Standard Deviation (μm)	0.3869	0.1321	2.9101	0.5316
Maximum Error (μm)	1.7203	0.5843	12.0898	2.3988

method, however, it can track more frames per second than the other tracking algorithms. The ESM algorithm accomplished to track a 50×50 window up to 250 pixels/sec velocity at 33 Hz in the experiments.

D. Visual Servoing Results

In the experiments, micropositioning and trajectory following tasks were performed at 1X and 4X zoom levels to compare the performances of calibrated and uncalibrated visual servoing (VS) algorithms. For the optimal control design, Q and L matrices in (16) were chosen as diagonal matrices with diagonal entries (0.9,0.9) and (0.025, 0.05) respectively. Micropositioning VS results are plotted in Figs. 4-7, and the trajectory following results for circular and square trajectories are depicted in Figs. 8-11.

For the micropositioning task, regulation performances of both approaches for a step input in terms of settling time (t_s), accuracy and precision are tabulated in Table III. For the trajectory following task, tracking performances of both approaches for different trajectories (square, circular and sinusoidal) are presented in Tables IV-V.

Both of the visual servoing approaches guarantee convergence to the desired targets with sub-micron error when time considerations are not so vital. When the time performance has priority for the task, the calibrated approach performs better than uncalibrated one in terms of settling time, accuracy and precision (Table III). Moreover, the tracking performance of the calibrated approach is more accurate and precise than the uncalibrated 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 $M = 6.3859$ 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 uncalibrated approach

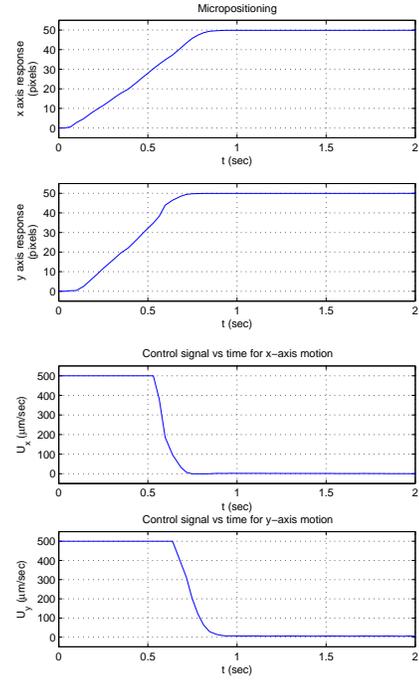


Fig. 4. Step responses and control signals of calibrated VS at 1X

in the regulation and tracking problems are also acceptable, and the difference between two approaches are not that significant. Furthermore, the uncalibrated method provides a more flexible servoing since the calibration of the optical system is a tedious and error prone process as explained in earlier sections and recalibration is required at each focusing level of the optical system.

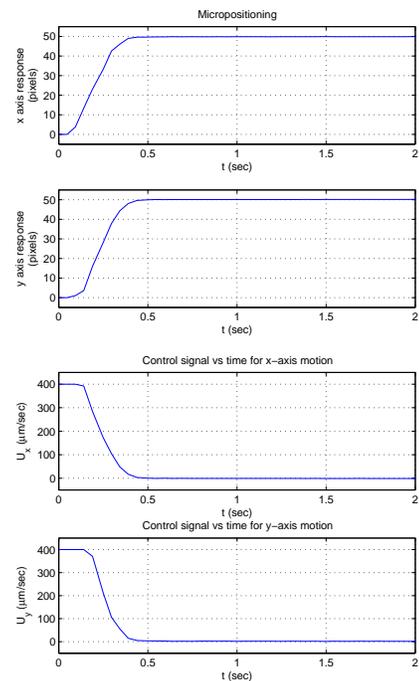


Fig. 5. Step responses and control signals of calibrated VS at 4X

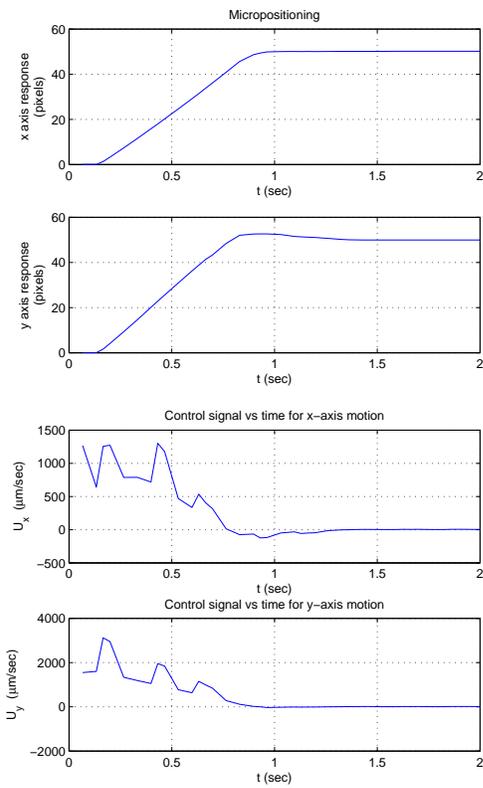


Fig. 6. Step responses and control signals of uncalibrated VS at 1X

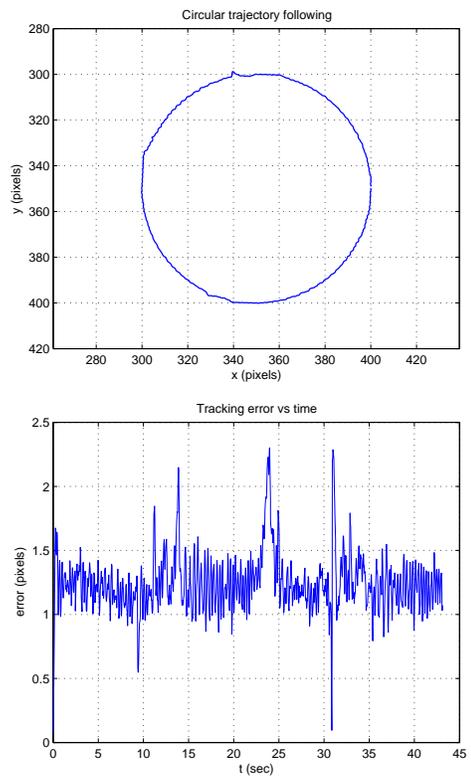


Fig. 8. Circular trajectory and tracking error in calibrated VS at 1X

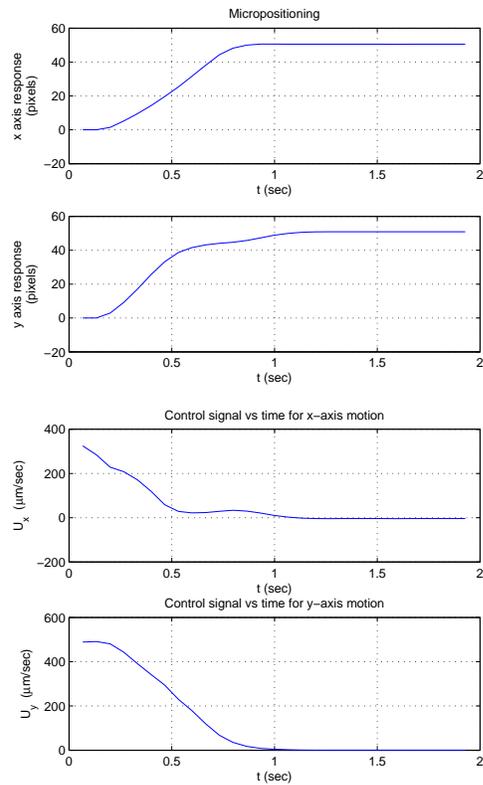


Fig. 7. Step responses and control signals of uncalibrated VS at 4X

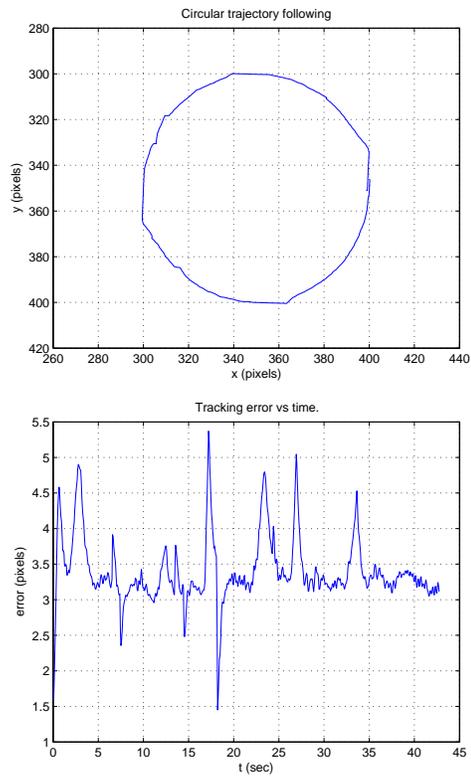


Fig. 9. Circular trajectory and tracking error in uncalibrated VS at 1X

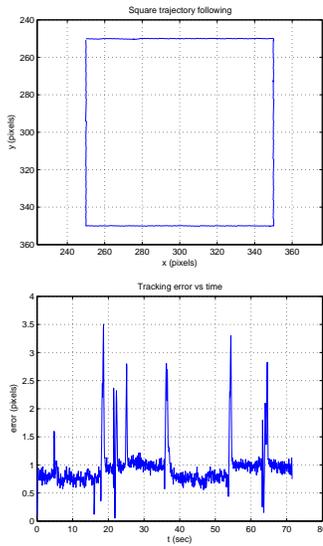


Fig. 10. Square trajectory and tracking error in calibrated VS at 1X

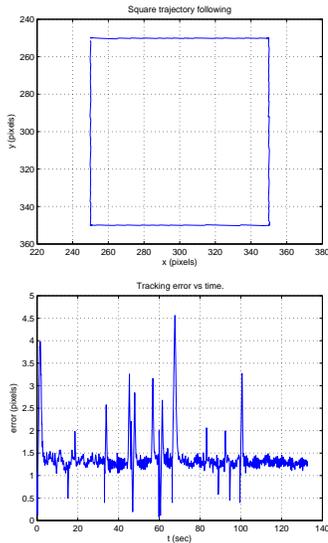


Fig. 11. Square trajectory and tracking error in uncalibrated VS at 1X

TABLE III
MICROPOSITIONING FOR CALIBRATED & UNCALIBRATED VS

	Step (pixels)	Calibrated			Uncalibrated		
		t_s (sec)	Acc. (μm)	Prec. (μm)	t_s (sec)	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

TABLE IV
TRAJECTORY TRACKING FOR CALIBRATED VS

	Square		Circular		Sinusoidal	
	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)
1x	5.93	2.28	7.72	1.40	4.79	2.37
4x	1.47	1.19	1.57	0.95	1.12	1.31

TABLE V
TRAJECTORY TRACKING FOR UNCALIBRATED VS

	Square		Circular		Sinusoidal	
	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)	Acc. (μm)	Prec. (μm)
1x	8.65	2.70	21.05	2.90	6.14	2.74
4x	1.64	1.12	3.30	1.17	1.17	0.57

IV. CONCLUSION

We have now presented an experimental comparison of conventional visual servoing schemes in certain microsystem applications. The results of these experiments have shown that the performance of the calibrated approach in terms of accuracy, precision and settling time is better than the uncalibrated approach, however, this difference does not necessarily imply a superiority for a coarse manipulation strategy. Moreover, the uncalibrated visual servoing has the advantages of carrying out a task without requiring a model of the system and adapting itself to different operating modes through a dynamic estimation of the composite Jacobian.

V. ACKNOWLEDGEMENT

This work is supported by SU Internal Grant No. IACF06 – 00417.

REFERENCES

- [1] R.S. Fearing, "A Planar Milli-Robot System on an Air Bearing", *Robotics Research the 7th International Symposium*, edited by G. Giralt and G. Hirzinger, London, Springer-Verlag, 1996, pp. 570-581.
- [2] I. Shimoyama, "Scaling in microrobotics", *IEEE/RSJ Int. Workshop on Intelligent Robots and Systems (IROS)*, Pittsburgh, PA, 1995, pp. 208-211.
- [3] Y. Mezouar, P. K. Allen, "Visual Servoed Micropositioning for Protein Manipulation Tasks", *IEEE Conf. on Intelligent Robots and Systems*, Switzerland, 2002, pp. 1766-1771.
- [4] R.Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *RA*, vol. 3, 1987, pp. 323-344.
- [5] Z.Y. Zhang, "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations", *IEEE International Conference on Computer Vision*, 1999, pp. 666-673.
- [6] B. Caprile and V. Torre, "Using Vanishing Points for Camera Calibration", *International Journal on Computer Vision*, Vol. 4, 1990, pp.127-140.
- [7] H. Zhuang and W. C. Wu, "Camera Calibration with a Near-Parallel Calibration Board Configuration", *IEEE Transactions on Robotics and Automation*, vol. 12, 1996, pp. 918-921.
- [8] Y. Zhou and B.J. Nelson, "Calibration of a parametric model of an optical microscope", *Optical Engineering*, vol.38, 1999, pp. 1989-1995.
- [9] M. Ammi, V. Fremont, A. Ferreira, "Flexible Microscope Calibration using Virtual Pattern for 3-D Telemicromanipulation", *IEEE Transactions on Robotics and Automation*, 2005, pp. 3888-3893.
- [10] J. Flusser, B. Zitova, "Invariants to Convolution with Circularly Symmetric PSF", *IEEE International Conference on Pattern Recognition*, 2004, pp. 11-14.
- [11] S.J. Ralis, B. Vikramaditya, and B.J. Nelson, "Micropositioning of A Weakly Calibrated Microassembly System Using Coarse-to-Fine Visual Servoing Strategies", *IEEE Trans. On Electronics Packing Manufacturing*, vol. 23, 2000, pp. 123-131.
- [12] J.A. Piepmeyer, "Uncalibrated Eye-in-Hand Visual Servoing", *The International Journal of Robotics Research*, vol. 22, 2003, pp. 805-819.
- [13] 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, pp. 943-948.