Simultaneous Calibration of Stereo Vision and 3D Optical Tracker for Robotic Microsurgery

S. Rodríguez Palma[1], B. C. Becker[2], C. N. Riviere[2]
[1]School of Industrial Engineering, University of Valladolid, Valladolid, Spain 47011
[2]The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA

***Abstract*—Visual servoing systems and 3-D optical trackers are important features in current robotic surgery research. Before performing any operation both the cameras and the tracker need to be calibrated. The procedure can be tedious and time-consuming, especially if the conditions change frequently. In this paper, a procedure is developed for the simultaneous calibration of a stereo vision system and a 3-D optical tracker. Key to the approach is joining the parameters of the two systems in a unique objective function. We show that this gives a simple method that requires less human time without increasing the calibration error significantly.**

I. INTRODUCTION

Visual servoing of an active handheld micromanipulator is under investigation for automation of certain surgical tasks [1]. However, the necessary calibration remains a challenging and time-consuming task. In recent years, camera calibration methods have been widely investigated [2]. Many different techniques for calibrating a pair of sensors including a vision device have been proposed [3, 4]. However, most such methods are designed for non-surgical contexts. There is a lack of methods to simultaneously calibrate optical sensors with micrometer accuracy, under the constraints of the microsurgical environment, and without using obtrusive calibration patterns like checkerboards, all of which are requisite for microsurgical use.

The purpose of this paper is to propose a new procedure to calibrate a stereo vision and an optical 3-D position measurement system in a manner suitable for microsurgery. This approach is applied for the calibration of the sensing system associated with Micron, a hand-held actively stabilized tool to enhance accuracy in microsurgery and other precision manipulations [1, 5]. The paper presents preliminary results of the calibration and a comparison with previously used methods.

II. BACKGROUND

Micron (Fig. 1) is a handheld micromanipulator with piezoelectric actuators built into the handle of the tool. The actuators can position the endpoint, or tip, within a roughly cylindrical workspace 500 µm in diameter [5].

Visual servoing is applied to guide the tip using visual feedback from cameras. Designed for microsurgical work, Micron is operated under a high-power Zeiss® OPMI® 1 microscope with a magnification often exceeding 25X and a visual workspace often only several millimeters in diameter. Two Point Grey Flea2 cameras are mounted to the microscope, providing a stereo view of the workspace. Each camera is approximately 2x3 mm with each pixel corresponding to 3.4 µm. Because the optical system involves high magnifications due to the microscope, the standard calibration techniques yield unsatisfactory results.

An external measurement system supplies Micron with the real-time position of its four LEDs: three attached to the tip, and one to the handle. The system is accurate to 4±2 µm [6].

Current calibration requires three consecutive steps. The first is the "offset" calibration: the surgeon should fix the tip and move the handle, recording different positions. This enables computation of the position offset of the tip from the LEDs. The second calibration routine involves the operator moving the tip randomly through the workspace, including up and down. As the 3-D tip position is already known, the two camera matrices are the result of this second step. The third and last calibration obtains the kinematic parameters of the manipulator. This paper describes a method for performing the first two calibrations simultaneously.

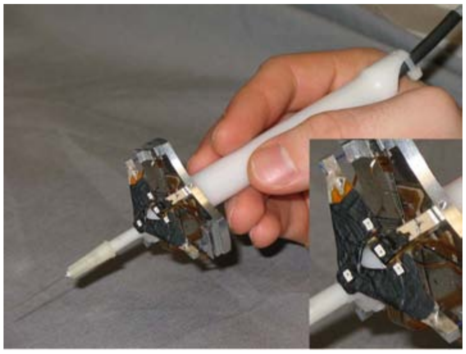

Fig. 1. Micron handpiece micromanipulator. Position sensor LEDs.

III. APPROACH

A. Problem description

Consider the problem of performing the two first steps of the routine calibration simultaneously, compute the offset and obtain the camera matrices of the stereo vision. Some issues that complicate the problem: the cameras are not perfectly affine, voluminous patterns (checkerboards) cannot be used, the movement of the cameras is restricted to the vertical axis, and accuracy to microns is required.

This work was supported by NIH grants R01EB000526 and R01EB007969.

B. *Calibration Routine*

To calibrate, the operator moves the tip through the workspace. The tip should reach the four corners of a square whose sides are known (Fig. 2). A typical 45-s calibration yields approximately 1500 data points, from which outliers are automatically removed via a simple distance metric before the calculations are performed.

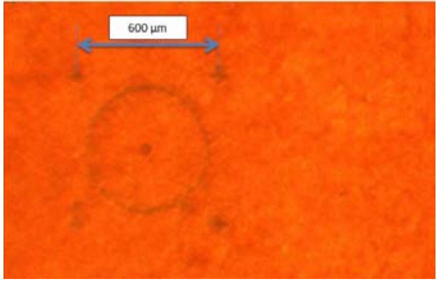

Fig. 2. 600 µm square pattern used for calibration.

C. *Algorithm*

The algorithm is divided in three sections. First of all we compute the 3x3 rotation matrices R_i which relate the position of the 3 LEDs between different points. These matrices and an initial iterant of the offset provide a first value for the 3-D tip position. After that, we apply the direct linear transform (DLT) to calculate the camera matrices using image points and 3-D points [7]. Lastly and the key of the approach is posing a single calibration optimization function objective (1). This is minimized improving the initial values of the offset, O, and the camera matrices, M_1 and M_2, using Lebenverg-Marquardt optimization. The function is built using the data set $\{u_i^A, u_i^B, L_i^1, L_i^2, L_i^3\}_{i=1..n}$, where u are the corresponding pixels of the tip in the two images (given by tracking), L are the 3-D positions of the 3 LEDs attached to the tip (given by the external measurement system for n points), Ψ is a function that makes the 3D reconstruction of a point using two camera matrices and the corresponding pair of image points, and C is the centroid of the 3 tip LEDs for each position.

$$Calibrate(M_1, M_2, O) = \sum_i^n \left\| \Psi(M_1, M_{2,} u_i^A, u_i^B) - C_i - R_i \cdot O \right\|^2 \tag{1}$$

In order to improve the camera calibration, we apply six distance preservation constraints. This is done using the pixels that match with the four corners of the square and forcing the 3-D reconstructions of these points to be separated by a fixed distance d_{ij} (side or diagonal of the square). Equation (2) shows the sort of the terms that are added to (1).

$$Const_{ij}(M_1, M_2, O) = \sigma^2 \cdot \left(\left\| \Psi(M_1, M_{2,} u_i^A, u_i^B) - \Psi(M_1, M_{2,} u_j^A, u_j^B) - d_{ij} \right\| \right)^2 \tag{2}$$

The algorithm is repeated until the maximum error between the image projection of the estimated 3-D points and the real pixel points are lower than a fixed value.

IV. RESULTS

For a total of 1380 data points, after removing outliers, Fig. 3 shows the real image points of the tip given by the tracking vs. the estimated points after the calibration (for one of the cameras). The procedure takes 45 s, whereas the old one took 90 s. Fig. 4 presents the histogram of the errors. Maximum errors are, respectively, 28.0 and 28.3 µm for each camera. The maximum and average errors of a typical run of the previously used method are about 18 µm and 7.5 µm.

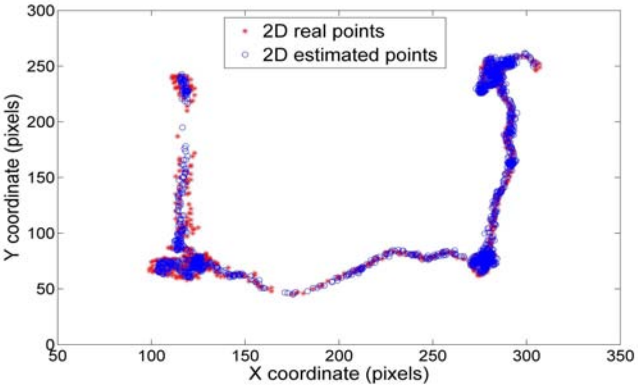

Fig. 3. Projection tip position in camera image (real VS estimated)

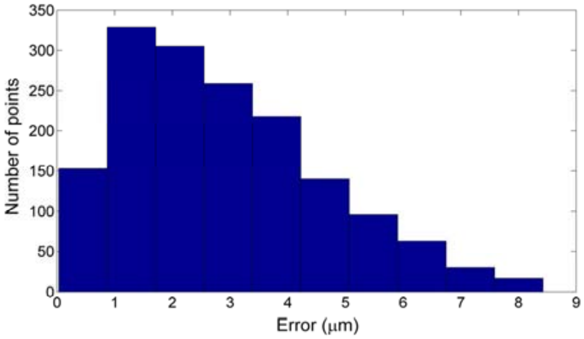

Fig. 4. Histogram of errors in micrometers.

V. DISCUSSION

This method allows calibration in one step, reducing the human time required by half. The results reflect an error of about 28 µm at the worst points, and an average error of 8.3 µm, which is similar to our previous method. The offset is computed correctly without investing extra time on its calibration. Future work will involve the incorporation of the kinematics in the simultaneous calibration.

REFERENCES

[1] B. C. Becker, S. Voros, R. A. MacLachlan, G. D. Hager, and C. N. Riviere, “Active guidance of a handheld micromanipulator using visual servoing,” in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 339-344, 2009.

[2] E. E. Hemayed, “A survey of camera self-calibration,” in *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, 2003, pp. 351-357.

[3] H. Zhuang, K. Wang, and Z. S. Roth, “Simultaneous calibration of a robot and a hand-mounted camera,” *IEEE Trans. Robot. Autom.*, vol. 11, no. 5, pp. 649-660, 1995.

[4] Q. V. Le and A. Y. Ng, “Joint calibration of multiple sensors,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 3651-3658, 2009.

[5] R. A. MacLachlan, B. C. Becker, G. W. Podnar, J. Cuevas Tabarés, L. A. Lobes, and C. N. Riviere, “Micron: an actively stabilized handheld tool for microsurgery,” *IEEE Trans. Robotics,* vol. 28, no. 1, pp. 195-212, 2012.

[6] R. A. MacLachlan and C. N. Riviere, “High-speed microscale optical tracking using digital frecuency-domain multiplexing,” *IEEE Trans. Instrum. Meas.*, vol. 58, no. 6, pp. 1991-2001, 2009.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2003.