[13] L. Kavraki, P.Švestka, J.-C. Latombe, and M. H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Trans. Robot. Automat.*, vol. 12, 1996.
[14] F. Lamiraux, S. Sekhavat, and J.-P. Laumond, "Motion planning and control for Hilare pulling a trailer," *IEEE Trans. Robot. Automat.*, vol. 15, 1999.
[15] J. C. Latombe, "A fast path planner for a car-like indoor mobile robot," in *9th Nat. Conf. Artif. Intell., AAAI*, Anaheim, CA, July 1991, pp. 659–665.
[16] ——, *Robot Motion Planning*. Boston, MA: Kluwer Academic, 1991.
[17] J.-P. Laumond, P. Jacobs, M. Taïx, and R. Murray, "A motion planner for non holonomic mobile robots," *IEEE Trans. Robot. Automat.*, vol. 10, pp. 577–593, 1994.
[18] J.-P. Laumond, "Feasible trajectories for mobile robots with kinematic and environment constraints," in *Int. Conf. Intell. Autonomous Syst.*, Amsterdam, The Netherlands, 1986.
[19] ——, "Controllability of a multibody mobile robot," *IEEE Trans. Robot. Automat.*, vol. 9, pp. 755–763, 1993.
[20] J.-P. Laumond, S. Sekhavat, and F. Lamiraux, "Guidelines in nonholonomic motion planning," in *Robot Motion Planning and Control*, J.-P. Laumond, Ed. New York: Springer-Verlag, 1998.
[21] B. Mirtich and J. Canny, "Using skeletons for nonholonomic motion planning among obstacles," in *IEEE Int. Conf. Robot. Automat.*, Nice, France, 1992, pp. 2533–2540.
[22] R. M. Murray and S. Sastry, "Steering nonholonomic systems using sinusoids," in *IEEE Int. Conf. Decision Contr.*, 1990, pp. 2097–2101.
[23] J. A. Reeds and R. A. Shepp, "Optimal paths for a car that goes both forward and backward," *Pacific J. Math.*, vol. 145, no. 2, pp. 367–393, 1990.
[24] A. Scheuer and T. Fraichard, "Collision-free and continuous-curvature path planning for car-like robots," in *IEEE Int. Conf. Robot. Automat.*, Albuquerque, NM, 1997.
[25] S. Sekhavat and J.-P. Laumond, "Topological property for collision-free nonholonomic motion planning: The case of sinusoidal inputs for chained form systems," *IEEE Trans. Robot. Automat.*, vol. 14, 1998.
[26] J. Sellen *et al.*, "Approximation and decision algorithms for curvature-constrained path planning: A state-space approach," in *Robotics: The Algorithmic Perspective*, P. K. Agarwal *et al.*, Eds: A.K. Peters, 1998.
[27] T. Siméon, J.-P. Laumond, and F. Lamiraux, "Move3D: A generic platform for path planning," in *Int. Symp. Assembly and Task Planning*, Fukuoka, Japan, May 2001.
[28] P. Souères and J.-D. Boissonnat, "Optimal trajectories for nonholonomic mobile robots," in *Robot Motion Planning and Control*, J.-P. Laumond, Ed. New York: Springer-Verlag, 1998.
[29] P. Svestka and M. Overmars, "Coordinated motion planning for multiple car-like robots using probabilistic roadmaps," in *IEEE Int. Conf. Robot. Automat.*, Nagoya, Aichi, Japan, 1995.
[30] D. Tilbury, R. Murray, and S. Sastry, "Trajectory generation for the n-trailer problem using goursat normal form," *IEEE Trans. Automat. Contr.*, vol. 40, 1995.

Automatic Camera Calibration for a Multiple-Sensor Integrated Coordinate Measurement System

Tzung-Sz Shen and Chia-Hsiang Menq

***Abstract*—An automatic camera calibration scheme that utilizes a coordinate measuring machine (CMM) and a novel camera calibration algorithm is presented in this paper for a multiple-sensor integrated coordinate measurement system. In the proposed calibration scheme, the touch probe tip carried by the CMM is employed to automatically generate high-precision calibration target points for camera calibration and sensor integration. A novel camera calibration algorithm with analytical formulations is developed to calibrate camera parameters in three stages without nonlinear minimization procedures. Simulations and experiments were performed to verify the proposed camera calibration algorithm. The precision of the automatic camera calibration scheme is also evaluated.**

***Index Terms*—Camera calibration, coordinate measuring machine, multiple-sensor integration.**

I. Introduction

Among three-dimensional (3-D) digitizing sensors, vision systems can simultaneously acquire thousands of data points over a large spatial range; therefore, capturing the global surface information of an object in real time is possible. In order to simultaneously achieve high-speed and high-precision coordinate acquisition, a multiple-sensor integrated coordinate measurement system can employ the vision sensor to rapidly acquire the initial global surface information [10]. The obtained information can be subsequently used to automatically guide the touch probe sensor or laser scanner for coordinate measurement with higher precision. For reliable and precise surface information acquisition, the accuracy of the vision system is essential.

When a vision system is used for 3-D coordinate measurement, camera calibration is usually required. The accuracy of the camera parameters, influenced by the calibration hardware as well as the calibration algorithms, has direct impact on the performance of the whole vision system. In this paper, an automatic camera calibration scheme that utilizes the coordinate measuring machine (CMM) and a novel calibration algorithm is developed for the multiple-sensor integrated coordinate measurement system. The automatic camera calibration scheme adopts the tip of the CMM probe to provide high-precision 3-D coordinates for camera calibration and to establish a common coordinate system for sensor integration. On the other hand, camera calibration techniques can be classified into: 1) linear methods (e.g., [1], [10]); 2) direct nonlinear minimization methods (e.g., [6]); and 3) multiple-stage methods (e.g., [2], [3], [11], [12]). Multiple-stage methods have become more popular because they are compromises between linear and direct nonlinear minimization methods. In the early stage, most camera parameters are obtained using closed-form solutions, and then iterative approaches are used to numerically solve remaining parameters and/or refine the parameters.

In multiple-stage methods, Tsai [11] first developed a two-stage technique under a radial alignment constraint (RAC). When the

Manuscript received April 7, 2000; revised November 27, 2000. This paper was recommended for publication by Associate Editor J. Ponce and Editor S. Hutchinson upon evaluation of the reviewers' comments. This work was supported by the National Science Foundation under Grant DMI-9713768.

The authors are with the Coordinate Metrology and Measurement Laboratory, Department of Mechanical Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: shen.45@osu.edu; menq.1@osu.edu).

Publisher Item Identifier S 1042-296X(01)08621-9.

lens distortion is considered, the focal length, z translation, and radial distortion are optimized using nonlinear optimization. However, as conventional linear methods, the orthonormality constraints are not considered [12]. Weng *et al.* [12] developed a two-step calibration procedure. Unlike Tsai's method that optimizes only three parameters, they iteratively compute 10 parameters with nonlinear minimization and improve all parameters. Bacakoglu and Kamel [2] proposed a three-step method with similar procedures in [12], but all parameters are fine-tuned by nonlinear optimization. Chatterjee *et al.* [3] developed a nonlinear calibration method. The calibration parameters are reduced to three (the image center and the scaling factor) in the nonlinear minimization. An interesting question is whether it is possible to avoid the nonlinear minimization procedure in camera calibration without scarifying the calibration accuracy. The bottleneck is in the calibration of the image center, scaling factor, and lens distortions.

The camera calibration method proposed in this paper is based on the property and physical meaning of the intrinsic parameters. Analytical formulations are used to calibrate camera parameters in three stages without nonlinear minimization procedures. Given the image center, the rotation matrix, x translation, y translation, and the scaling factor can be obtained using an analytical method in the first stage. The effect of the scaling factor on the orthonormality of the rotation matrix can be revealed from the analytical formulations. The effective focal length, z translation, and the distortion coefficient are obtained in the second stage. The image center offset can then be either numerically or analytically estimated in the third stage to recursively update the image center. The convergence of the image center calibration takes only two to three steps. Distinct from multiple-stage methods in [2], [3], and [12], the proposed method requires neither additional initial guess procedure nor nonlinear minimization process, the calibration of camera parameters have analytical formulations, and the distortion coefficient can be solved in closed form.

II. Calibration Target Generation Using CMM

In the multiple-sensor integrated system, camera calibration targets must be measurable by the touch probe to establish a common coordinate system or to obtain the coordinate transformation between the vision system and the CMM for sensor integration. In prior work, calibration target points were either generated by paper printing [10], photolithography processes [3], [12], or associated with a high-precision engineered artifact [11]. However, the touch probe cannot measure calibration patterns created by printing or photolithography processes, and the precision of engineered artifacts is limited to that of machining processes employed. Furthermore, the use of moving mechanisms for producing noncoplanar calibration targets or performing error analysis complicates the calibration procedures and swells the system setup.

In the automatic camera calibration scheme, the CMM is employed to automatically generate high-precision target points and provide flexible calibration target patterns through programming. The touch probe tip is very close to a perfect sphere and its image is a circular object at any viewing angle. Therefore, the tip center can be easily estimated and adopted as the camera calibration target. A tip locating algorithm with subpixel accuracy in [8] is employed to locate the tip center.

III. Camera Model

Perfect pinhole camera model is employed in this paper. An object point $\boldsymbol{P}(x_w,\ y_w,\ z_w)$ is first transformed from the world coordinate system to the camera coordinate system $(x_m,\ y_m,\ z_m)$, whose z axis and origin coincide with the optical axis and optical center of the camera lens, respectively. This rigid body transformation is formulated by

$$\begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} = \boldsymbol{R}^T \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \boldsymbol{T} \tag{1}$$

where $\boldsymbol{R} = [\hat{\boldsymbol{n}}_x\ \ \hat{\boldsymbol{n}}_y\ \ \hat{\boldsymbol{n}}_z]$ and $\boldsymbol{T} = [T_x\ \ T_y\ \ T_z]^T$ are the rotation matrix and the translation vector. Through perspective projection, the undistorted image coordinate $(x_u,\ y_u)$ is

$$\begin{cases} x_u = f\dfrac{x_m}{z_m} = f\dfrac{\hat{\boldsymbol{n}}_x^T\boldsymbol{P} + T_x}{\hat{\boldsymbol{n}}_z^T\boldsymbol{P} + T_z} \\ y_u = f\dfrac{y_m}{z_m} = f\dfrac{\hat{\boldsymbol{n}}_y^T\boldsymbol{P} + T_y}{\hat{\boldsymbol{n}}_z^T\boldsymbol{P} + T_z} \end{cases} \tag{2}$$

where f is the effective focal length.

When there exist lens distortion and frequency mismatch between the charge coupled device (CCD) camera and the image grabber board, lens distortion coefficients and a scaling factor must be considered. For lenses suitable for machine vision, the radial distortion is more significant than other types of lens distortion. Therefore, only one term of radial distortion coefficient (k_1) is included. According to the distortion model in [12], the distorted image coordinate can be expressed as

$$\begin{cases} x_d = \left(1 + k_1\left(x_u^2 + y_u^2\right)\right)x_u = \left(1 + k_1 r_u^2\right)x_u \\ y_d = \left(1 + k_1\left(x_u^2 + y_u^2\right)\right)y_u = \left(1 + k_1 r_u^2\right)y_u. \end{cases} \tag{3}$$

However, because the exact $(x_u,\ y_u)$ cannot be obtained, it is reasonable to replace $(x_u,\ y_u)$ by $(x_d,\ y_d)$ [12]. Therefore, r_u in (3) is replaced by $r(x_d,\ y_d) = \sqrt{x_d^2 + y_d^2}$

$$\begin{cases} x_d = \left(1 + k_1 r^2\right)x_u \\ y_d = \left(1 + k_1 r^2\right)y_u. \end{cases} \tag{4}$$

Note that the distorted image coordinate expressed in (4) is slightly different from those in [10], [11]. The sampled 2-D image frame coordinate of the object point $(x_f,\ y_f)$ is

$$\begin{cases} x_f = C_x - S_x d_x^{-1} x_d \\ y_f = C_y - d_y^{-1} y_d \end{cases} \tag{5}$$

where $(C_x,\ C_y)$ represents the 2-D image frame coordinate of the image center (O_f), S_x is the scaling factor, and d_x and d_y are the pixel sizes of the CCD camera, which are usually available from the manufacturer. In the camera model described above, the intrinsic parameters include image center $(C_x,\ C_y)$, scaling factor (S_x), radial distortion coefficient (k_1), and effective focal length (f). Extrinsic parameters are rotation angles $(\alpha,\ \beta,\ \gamma)$ and translations $(T_x,\ T_y, T_z)$.

IV. Camera Calibration Algorithm

Due to the perspective projection, the extrinsic and intrinsic parameters of the camera are coupled together. It results in a nonlinear problem for camera calibration. According to the sensitivity analysis in [4], the influence of the image center offset on the rigid body transformation is less significant than that of the scaling factor. Consequently, in order to estimate the rigid body transformation, it is reasonable to first ignore the image center offset and use the center of the image frame as the initial image center. The proposed camera calibration algorithm is illustrated in Fig. 1.

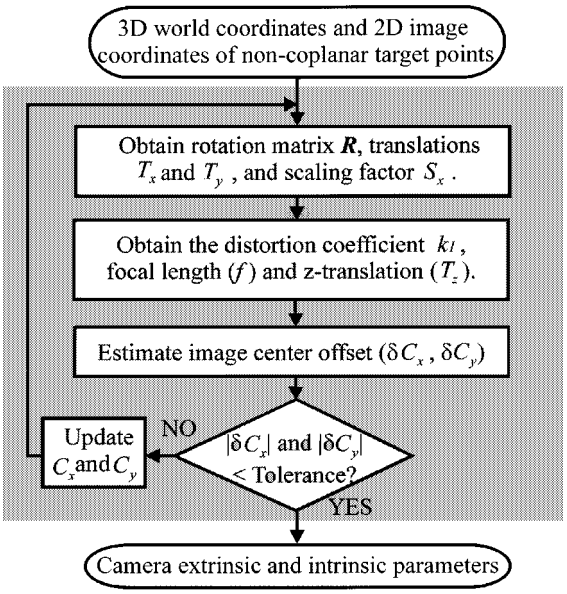

Fig. 1. Camera calibration algorithm.

A. Calibration of Rotation Matrix, X/Y Translations, and Scaling Factor

When only the radial distortion is considered, the RAC constraint in [11] can be employed to separate the camera parameters. For the ith target point, dividing x_{di} by y_{di} leads to

$$\frac{x_{di}}{y_{di}} = \frac{x_{ui}}{y_{ui}} = \frac{\hat{\boldsymbol{n}}_x^T \boldsymbol{P}_i + T_x}{\hat{\boldsymbol{n}}_y^T \boldsymbol{P}_i + T_y}. \tag{6}$$

From (5), $(x_{di},\ y_{di})$ can be expressed in terms of $(x_{fi},\ y_{fi})$ and (6) can be rewritten as

$$\frac{x_{di}}{y_{di}} = \frac{S_x^{-1} d_x (C_x - x_{fi})}{d_y (C_y - y_{fi})} = \frac{\hat{\boldsymbol{n}}_x^T \boldsymbol{P}_i + T_x}{\hat{\boldsymbol{n}}_y^T \boldsymbol{P}_i + T_y}. \tag{7}$$

When there are measurement errors in the image frame coordinates, the residual error can be expressed as

$$\begin{aligned} \varepsilon_{Ri} &= \hat{\boldsymbol{n}}_x^T \boldsymbol{P}_i y_{di} - \hat{\boldsymbol{n}}_y^T \boldsymbol{P}_i x_{di} + T_x y_{di} - T_y x_{di} \\ &= \sqrt{1 + S_x^{-2}}\, \hat{\boldsymbol{n}}_R^T {}^6\boldsymbol{p}_i + \boldsymbol{t}_R^T {}^2\boldsymbol{p}_i \end{aligned} \tag{8}$$

where

$$\begin{aligned} {}^6\boldsymbol{p}_i &= [\, \boldsymbol{P}_i^T d_y (C_y - y_{fi}) \quad -\boldsymbol{P}_i^T d_x (C_x - x_{fi})\,]^T \\ {}^2\boldsymbol{p}_i &= [\, d_y (C_y - y_{fi}) \quad -d_x (C_x - x_{fi})\,]^T \\ \boldsymbol{t}_R &= [\, T_x \quad S_x^{-1} T_y\,]^T \end{aligned}$$

and

$$\hat{\boldsymbol{n}}_R = 1/\sqrt{1 + S_x^{-2}}\, [\, \hat{\boldsymbol{n}}_x^T \quad S_x^{-1} \hat{\boldsymbol{n}}_y^T\,]^T.$$

Given n noncoplanar target points, an objective function can be defined as

$$\begin{aligned} J_R(\hat{\boldsymbol{n}}_R,\ \boldsymbol{t}_R,\ S_x) &= \sum_{i=1}^{n} \varepsilon_{Ri}^2 \\ &= \left(1 + S_x^{-2}\right) \hat{\boldsymbol{n}}_R^T {}^6\boldsymbol{H}_6 \hat{\boldsymbol{n}}_R \\ &\quad + \boldsymbol{t}_R^T {}^2\boldsymbol{H}_2 \boldsymbol{t}_R + 2\sqrt{1 + S_x^{-2}}\, \hat{\boldsymbol{n}}_R^T {}^6\boldsymbol{H}_2 \boldsymbol{t}_R \end{aligned} \tag{9}$$

where

$${}^6\boldsymbol{H}_6 = \sum_{i=1}^{n} {}^6\boldsymbol{p}_i\, {}^6\boldsymbol{p}_i^T$$

$${}^2\boldsymbol{H}_2 = \sum_{i=1}^{n} {}^2\boldsymbol{p}_i\, {}^2\boldsymbol{p}_i^T$$

and

$${}^6\boldsymbol{H}_2 = \sum_{i=1}^{n} {}^6\boldsymbol{p}_i\, {}^2\boldsymbol{p}_i^T.$$

From (9), $\nabla_{\boldsymbol{t}_R} J_R = 0$ gives

$$\boldsymbol{t}_R = \begin{bmatrix} T_x \\ S_x^{-1} T_y \end{bmatrix} = -\sqrt{1 + S_x^{-2}} \quad {}^2\boldsymbol{H}_2^{-1} \quad {}^6\boldsymbol{H}_2^T \hat{\boldsymbol{n}}_R. \tag{10}$$

Substituting (10) into (9) simplifies the objective function as

$$J_R'(\hat{\boldsymbol{n}}_R,\ S_x) = \left(1 + S_x^{-2}\right) \hat{\boldsymbol{n}}_R^T \boldsymbol{H} \hat{\boldsymbol{n}}_R \tag{11}$$

where $\boldsymbol{H} = {}^6\boldsymbol{H}_6 - {}^6\boldsymbol{H}_2 \quad {}^2\boldsymbol{H}_2^{-16} \quad \boldsymbol{H}_2^T$. J_R' can be minimized if

$$\hat{\boldsymbol{n}}_R = \begin{bmatrix} \dfrac{1}{\sqrt{1 + S_x^{-2}}} \hat{\boldsymbol{n}}_x \\ \dfrac{S_x^{-1}}{\sqrt{1 + S_x^{-2}}} \hat{\boldsymbol{n}}_y \end{bmatrix} = \hat{\boldsymbol{u}}_{H6} = \begin{bmatrix} \boldsymbol{u}_{H6x} \\ \boldsymbol{u}_{H6y} \end{bmatrix} \tag{12}$$

where $\hat{\boldsymbol{u}}_{H6}$ is the singular vector associated with the smallest singular value of $\boldsymbol{H}$. It is revealed from (12) that S_x can influence the normality of the rotation matrix, but not the orthogonality. The scaling factor can be obtained by

$$S_x = \frac{\|\boldsymbol{u}_{H6x}\|}{\|\boldsymbol{u}_{H6y}\|}. \tag{13}$$

Likewise, $\hat{\boldsymbol{n}}_x$ and $\hat{\boldsymbol{n}}_y$ can be determined from (12), and their cross product determines $\hat{\boldsymbol{n}}_z$. T_x and T_y can be obtained from (10). If the actual directions of $\hat{\boldsymbol{n}}_x$ and $\hat{\boldsymbol{n}}_y$ are opposite to those of $\boldsymbol{u}_{H6x}$ and $\boldsymbol{u}_{H6y}$, they can be corrected according to the sign of the effective focal length. Since the scaling factor is calibrated, the resulting rotation matrix will be very close to orthonormal. In addition, as the image center is recursively calibrated and converges to the real location, the rotation matrix will be finally orthonormal.

B. Calibration of Distortion, Focal Length, and Z Translation

From (5), $(x_{di},\ y_{di})$ can be calculated with $(x_{fi},\ y_{fi})$, the calibrated scaling factor and current image center. On the other hand, from (2) and (4), $(x_{di},\ y_{di})$ can be expressed as

$$\begin{cases} x_{di} = \left(1 + k_1 r_i^2\right) f \dfrac{x_{mi}}{z_{mi}} = \left(1 + k_1 r_i^2\right) f \dfrac{\hat{\boldsymbol{n}}_x^T \boldsymbol{P}_i + T_x}{\hat{\boldsymbol{n}}_z^T \boldsymbol{P}_i + T_z} \\ y_{di} = \left(1 + k_1 r_i^2\right) f \dfrac{y_{mi}}{z_{mi}} = \left(1 + k_1 r_i^2\right) f \dfrac{\hat{\boldsymbol{n}}_y^T \boldsymbol{P}_i + T_y}{\hat{\boldsymbol{n}}_z^T \boldsymbol{P}_i + T_z}. \end{cases} \tag{14}$$

The following residual errors can be defined:

$$\begin{cases} \varepsilon_{xi} = T_z x_{di} - f\left(1 + k_1 r_i^2\right) x_{mi} + \hat{\boldsymbol{n}}_z^T \boldsymbol{P}_i x_{di} \\ \varepsilon_{yi} = T_z y_{di} - f\left(1 + k_1 r_i^2\right) y_{mi} + \hat{\boldsymbol{n}}_z^T \boldsymbol{P}_i y_{di}. \end{cases} \tag{15}$$

It can be observed from (15) that k_1 is coupled with f and T_z; therefore, these three parameters were numerically determined by nonlinear minimization in [11]. Nevertheless, with careful arrangements, k_1 can be separated from f and T_z. The objective function can be derived as

$$J_K(k_1,\ \boldsymbol{b}) = \sum_{i=1}^{n} \left(\varepsilon_{xi}^2 + \varepsilon_{yi}^2\right) = \boldsymbol{b}^T \boldsymbol{K} \boldsymbol{b} + 2\boldsymbol{\kappa}^T \boldsymbol{b} + c \tag{16}$$

where

$$\boldsymbol{K} = \begin{bmatrix} s_1 & -s_2 - s_3 k_1 \\ -s_2 - s_3 k_1 & s_4 + 2 s_5 k_1 + s_6 k_1^2 \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} T_z \\ f \end{bmatrix}$$

$$\boldsymbol{\kappa} = \begin{bmatrix} s_7 \\ -s_8 - s_9 k_1 \end{bmatrix}$$

$$c = \hat{\boldsymbol{n}}_z^T \left[\sum_{i=1}^{n} \left(r_i^2 \boldsymbol{P}_i \boldsymbol{P}_i^T \right) \right] \hat{\boldsymbol{n}}_z$$

and $s_1 - s_9$ are variables depending on $\boldsymbol{P}_i$, $(x_{di},\ y_{di})$, and the parameters estimated in Section IV-A. From (16), $\nabla_{\boldsymbol{b}} J_K = 0$ gives

$$\boldsymbol{b} = \begin{bmatrix} T_z \\ f \end{bmatrix} = -\boldsymbol{K}^{-1} \boldsymbol{\kappa}. \tag{17}$$

Substituting (17) into (16) leads to the following objective function that is only a function of k_1:

$$J_K'(k_1) = c - \boldsymbol{\kappa}^T \boldsymbol{K}^{-1} \boldsymbol{\kappa}. \tag{18}$$

From $\nabla_{k_1} J_K' = 0$, a quadratic equation of k_1 can then be easily derived as

$$c_2 k_1^2 + c_1 k_1 + c_0 = 0 \tag{19}$$

where $c_j' s,\ j = 0, 1,$ and 2, are functions of $(x_{di},\ y_{di})$, $\boldsymbol{P}_i$, and parameters estimated in Section IV-A. Derivations of s_1–s_9 and c_0–c_2 can be found in [8]. The solution of k_1 is the one that minimizes J_K'. f and T_z can be calculated using (17). If the obtained f is negative, the directions of $\hat{\boldsymbol{n}}_x$ and $\hat{\boldsymbol{n}}_y$ are opposite to their actual ones. Thus, the signs of $\hat{\boldsymbol{n}}_x$, $\hat{\boldsymbol{n}}_y$, T_x and T_y must be corrected. The closed-form solution of k_1 eliminates nonlinear iterations for distortion calibration.

C. Calibration of Image Center

Under the current image center, the camera parameters calibrated in previous stages are optimal solutions according to the defined objective functions. These parameters form a 9×1 vector $\tilde{\boldsymbol{t}}$, $\tilde{\boldsymbol{t}} = [\alpha\ \beta\ \gamma\ T_x\ T_y\ T_z\ f\ k_1\ S_x]^T$, which is a function of the image center $\boldsymbol{C}$, $\boldsymbol{C} = [C_x\ \ C_y]^T$. For the ith target point, the distorted image coordinates $\boldsymbol{P}_{di}(\tilde{\boldsymbol{t}})$, $\boldsymbol{P}_{di}(\tilde{\boldsymbol{t}}) = [S_x x_{di}\ y_{di}]^T$, can be calculated by projecting 3-D world coordinates to a 2-D image frame with the calibrated $\tilde{\boldsymbol{t}}$. The calculated distorted image coordinates can then be compared with the sampled image frame coordinates $\boldsymbol{P}_{fi}(\boldsymbol{C})$, $\boldsymbol{P}_{fi} = [d_x(C_x - x_{fi})\ d_y(C_y - y_{fi})]^T$ to determine the image center. Two approaches can be employed. In the first approach, a nonlinear minimization procedure can be used to find the image center by minimizing the objective function

$$J_C(\boldsymbol{C}) = \sum_{i=1}^{n} \left[\boldsymbol{P}_{di}\left(\tilde{\boldsymbol{t}}(\boldsymbol{C})\right) - \boldsymbol{P}_{fi}(\boldsymbol{C}) \right]^2. \tag{20}$$

Compared to approaches in [2], [3], and [12], in terms of the number of parameters involved in the nonlinear minimization, there are only two parameters, C_x and C_y, to be searched.

In the second approach, a linearized formula based on Taylor's series expansion is used to recursively update the image center. Let $\boldsymbol{C}_0$ be the current estimated image center. Then the linearized relation can be derived as follows:

$$\begin{aligned} \tilde{\varepsilon}_i &= \boldsymbol{P}_{di}\left(\tilde{\boldsymbol{t}}(\boldsymbol{C}_0)\right) - \boldsymbol{P}_{fi}(\boldsymbol{C}_0) \\ &\cong \left(\left(\frac{\partial \boldsymbol{P}_{di}}{\partial \tilde{\boldsymbol{t}}} \right)^T \frac{\partial \tilde{\boldsymbol{t}}}{\partial \boldsymbol{C}} - \frac{\partial \boldsymbol{P}_{fi}}{\partial \boldsymbol{C}} \right) \delta \boldsymbol{C}. \end{aligned} \tag{21}$$

Given n points, the image center offset can be obtained by

$$\delta \boldsymbol{C} = \begin{bmatrix} \delta C_x \\ \delta C_y \end{bmatrix} = ({}^m\boldsymbol{M}_C - {}^m\boldsymbol{D}_C)^+ \tilde{\varepsilon}_E \tag{22}$$

where

$$\tilde{\varepsilon}_E = [\ \tilde{\varepsilon}_1^T \quad \cdots \quad \tilde{\varepsilon}_n^T\ ]^T$$

$${}^m\boldsymbol{M}_C = \left[\left(\frac{\partial \boldsymbol{P}_{d1}}{\partial \tilde{\boldsymbol{t}}} \right) \quad \cdots \quad \left(\frac{\partial \boldsymbol{P}_{dn}}{\partial \tilde{\boldsymbol{t}}} \right) \right]^T \frac{\partial \tilde{\boldsymbol{t}}}{\partial \boldsymbol{C}}$$

$${}^m\boldsymbol{D}_C = \left[\left(\frac{\partial \boldsymbol{P}_{f1}}{\partial \boldsymbol{C}} \right)^T \quad \cdots \quad \left(\frac{\partial \boldsymbol{P}_{fn}}{\partial \boldsymbol{C}} \right)^T \right]^T$$

and "+" indicates pseudoinverse of a nonsquare matrix. ${}^m\boldsymbol{M}_C$ defines the sensitivity of image coordinate errors to the image center offset; while ${}^m\boldsymbol{D}_C$ maps image center offset from pixel to metric domain.

The matrix ${}^m\boldsymbol{M}_C$ can be obtained analytically or numerically. Through the sensitivity analysis on $\tilde{\boldsymbol{t}}$ and $\boldsymbol{C}$, ${}^m\boldsymbol{M}_C$ can be obtained with analytical formulations. The derivation of $(\partial \boldsymbol{P}_{di}/\partial \tilde{\boldsymbol{t}})$ in ${}^m\boldsymbol{M}_C$ is straightforward, but deriving $(\partial \tilde{\boldsymbol{t}}/\partial \boldsymbol{C})$ involves the eigensensitivity analysis on the $\boldsymbol{H}$ matrix in (11) that is too lengthy to be presented in this paper. Interested readers are referred to [7]. Alternatively, a numerical approach is adopted in this paper to estimate ${}^m\boldsymbol{M}_C$ by perturbing C_x and C_y and recalibrating the camera parameters to calculate image coordinate variations. The step size of the perturbation is initially one pixel and decreases by an order in the next step.

Using (22), the image center offset can be compensated to minimize the objective function in (20). With this physically meaningful formulation, the nonlinear minimization procedure is avoided; therefore, the second image center calibration approach is adopted in this research. Compared to the experimental approach in [5] and nonlinear approaches in [2], [3], and [12], the proposed image center calibration method is completely algorithmic without nonlinear minimization processes.

V. Camera Calibration Results

The developed automatic camera calibration scheme is evaluated using computer simulations and experimental data. In the experiment, the multiple-sensor integrated system developed in [9] is adopted. The resolution of the camera is 768 × 576 and the pixel size is 11 μm × 11 μm.

Three sets of tests were conducted: 1) camera calibration using synthetic data; 2) camera calibration using experimental data; and 3) precision evaluation of the overall system. The results obtained using the proposed calibration method are compared with linear and nonlinear methods. The linear method refers to the linear estimation procedure in [12], which ignores lens distortion. The nonlinear method is modified from the two-step estimation procedure in [12] and includes the distortion calibration in the nonlinear minimization. These calibration methods are implemented using C language and tested using a PC with a 266-MHz CPU and 128-MB memory.

A. Camera Calibration Using Synthetic Data

Using synthetic data generated with known camera parameters, parameter errors can be calculated to evaluate calibration accuracy. The simulated calibration target points are noncoplanar and lie on two parallel planes with a distance of 60 mm. Each plane contains target points in a 5 × 5 grid with distances of 31.5 mm in the vertical direction and 22.5 mm in the horizontal. The measurement error of the sampled image frame coordinates is simulated by random error with standard deviation of 0.081 65 pixels. Three configurations with β, the rotation angle about Y_w axis, set as 160°, 180°, and 200° are tested. For

TABLE I
CAMERA CALIBRATION WITH SYNTHETIC DATA

Calibration method	Linear	Nonlinear	Proposed
$\lvert\hat{n}_x - \hat{n}_x^*\rvert / \lvert\hat{n}_x^*\rvert$	0.009553	0.001028	0.001249
$\lvert\hat{n}_y - \hat{n}_y^*\rvert / \lvert\hat{n}_y^*\rvert$	0.000968	0.001051	0.001219
$\lvert\hat{n}_z - \hat{n}_z^*\rvert / \lvert\hat{n}_z^*\rvert$	0.009825	0.001598	0.001851
$\lvert T - T^* \rvert / \lvert T^* \rvert$	0.014693	0.001796	0.002014
$\lvert f - f^* \rvert / \lvert f^* \rvert$	0.004111	0.000669	0.000681
$\lvert Cx - Cx^* \rvert / \lvert Cx^* \rvert$	0.040931	0.004196	0.005139
$\lvert Cy - Cy^* \rvert / \lvert Cy^* \rvert$	0.005112	0.005212	0.006105
$\lvert Sx - Sx^* \rvert / \lvert Sx^* \rvert$	0.001100	0.000083	0.000160
$\lvert k1 - k1^* \rvert / \lvert k1^* \rvert$	-	0.014664	0.016438
$\overline{\lvert\delta X_w\rvert}$ (mm)	0.078957	0.017218	0.018185
$\overline{\lvert\delta Y_w\rvert}$ (mm)	0.059495	0.016656	0.017085
$\overline{\delta r_w}$ (mm)	0.106995	0.026698	0.027806
Image Error (pixel)	0.451439	0.107504	0.112366

each configuration, 10 random trials are repeated. The other parameters are $[T_x, T_y, T_z] = [63.0797, -45.5880, 399.9209]$, $\alpha = 0.1^\circ$, $\gamma = -0.1^\circ$, $f = 16.0$ mm, $k_1 = -8 \times 10^{-4}$, $S_x = 1.04$, and $(C_x, C_y) = (374, 278)$, which has an offset of $(-10, -10)$ away from image frame center. Two error measures, namely image error and position error, are defined to evaluate the camera calibration results. The image error is defined as

$$\text{Image Error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\left(x_{fi} - x'_{fi}\right)^2 + \left(y_{fi} - y'_{fi}\right)^2\right]} \qquad (23)$$

where n is the number of target points, (x_{fi}, y_{fi}) is the sampled 2-D image coordinate, and (x'_{fi}, y'_{fi}) is the 2-D image frame coordinate calculated by projecting known 3D coordinate of the target point to 2-D image frame with calibrated camera parameters. The image error should be close to the simulated random noise if calibrated camera parameters are close to optimal solutions. Since the depth cannot be determined with a monocular camera setup, the depth of a known target point is given to project its sampled 2-D image frame coordinate back to the 3-D coordinate. The deviations between projected and known 3-D world coordinates in X_w and Y_w directions are defined as the position errors, δX_w and δY_w. The radial position error δr_w is defined by $\sqrt{\delta X_w^2 + \delta Y_w^2}$.

The calibration results are listed in Table I. The superscript "*" indicates the true value of the camera's parameter. The simulation results show that the recursive calibration of the image center takes only two to three steps to converge. From Table I, it is evident that the proposed method can calibrate camera parameters more accurately than the linear method, and the error measures are less than one-fourth of those obtained by the linear method. The results of the proposed method are very close to those optimized using the nonlinear method.

B. Camera Calibration Using Experimental Data

In the experiments, high-precision noncoplanar calibration target points are automatically generated using the CMM. The configurations of calibration target points are the same as those described in Section V-A. The focal length of the camera lens is 16.0 mm. The two error measures defined in Section V-A are used as performance indicators of a calibration algorithm.

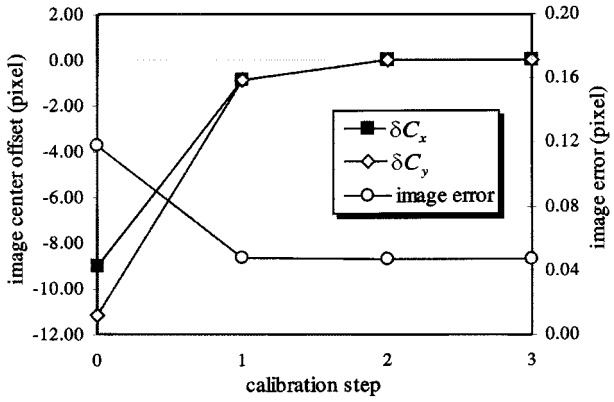

Fig. 2. Image center calibration.

TABLE II
CAMERA CALIBRATION WITH EXPERIMENTAL DATA

Calibration method	Linear	Nonlinear	Proposed
$\overline{\lvert\delta X_w\rvert}$ (μm)	37.7028	6.9137	8.9781
$\overline{\lvert\delta Y_w\rvert}$ (μm)	27.0504	9.3481	9.4550
$\overline{\delta r_w}$ (μm)	49.7493	12.9035	14.5264
Image Error (pixel)	0.1716	0.0439	0.0485

As shown in Fig. 2, the recursive image center calibration result indicates that the absolute value of the image center offset converges to be within 0.01 pixels at the second calibration step, and the image error is within 0.05 pixels. At the third step, the image center offset is less than 0.001 pixels.

The camera calibration was experimentally conducted 15 times at three different camera locations. As can be seen from Table II, the averages of the error measures using the proposed method and the nonlinear method are only 20%–35% of those obtained using the linear method. It is evident that the lens distortion is not negligible in this experiment, and ignoring orthonormality constraints of the rotation matrix will impose errors on calibration results. Using the proposed method, the error measures are very close to those obtained by the nonlinear method. Since the proposed method needs no nonlinear minimization, it requires less computation time. The average computation time of the proposed method is 0.13 s .The nonlinear method takes about 1.0 s, although it does not make much difference for off-line applications.

C. Precision Evaluation of the Camera Calibration Scheme

A stereo vision setup is used to evaluate the absolute precision of the overall system. With the stereo setup, the depth position error in the Z_w direction, δZ_w, can be obtained. The absolute position error, δD_w, is the distance between the known target point and measured point in the 3-D world coordinate system.

In this test, 50 object points in a 5×5 grid on two parallel planes inside the calibrated range are generated using the CMM. The distance between these two planes is 40 mm and object points are in an area of 126 mm $\times$ 90 mm. The measured 3-D data points are fitted with planes and compared with known plane parameters. A plane is represented by its normal vector $\hat{\boldsymbol{n}}_{\boldsymbol{P}}$ and the distance between the plane and the origin along plane normal direction L. The plane orientation error $|\delta\hat{\boldsymbol{n}}_{\boldsymbol{P}}|$ is defined as the norm of the plane normal deviation. The plane location error $|\delta L|$ is the deviations of the distance L. Therefore, position errors of individual points as well as parameter errors of planar features can be evaluated.

TABLE III
PRECISION EVALUATION USING STEREO VISION

Calibration method	Linear	Nonlinear	Proposed
$\overline{\lvert\delta X_w\rvert}$ (μm)	39.7000	5.6711	6.0003
$\overline{\lvert\delta Y_w\rvert}$ (μm)	55.6575	18.1429	18.5753
$\overline{\lvert\delta Z_w\rvert}$ (μm)	29.3909	15.2511	15.8119
$\overline{\delta D_w}$ (μm)	82.6229	27.4295	28.0719
$\lvert\delta N_p\rvert$	5.1279E-04	3.0102E-04	3.8576E-04
$\overline{\lvert\delta L\rvert}$ (μm)	53.1198	11.6246	10.9733

The measurements repeat 10 times, and the averages of the error measures are listed in Table III. As can be seen in Table III, the orientation errors of the planar features are all very small. The average absolute position error using the linear method is 82.62 μm, and the average location error of planar features is 53.12 μm. Using the proposed method, the average absolute position error is reduced to 28.07 μm, and the location error is less than 11.0 μm, which are comparable to those obtained using the nonlinear method.

VI. CONCLUSION

An automatic camera calibration scheme that utilizes a CMM and a novel camera calibration algorithm is presented for a multiple-sensor integrated coordinate measurement system. Distinct from other multiple-stage methods, the proposed camera calibration method requires neither particular initial guess procedure nor nonlinear minimization process. Synthetic and experimental tests have demonstrated that accurate camera calibration and precise coordinate measurements can be obtained using the proposed calibration scheme. With high-precision calibration targets generated automatically using CMM and the proposed camera calibration algorithm, a fully automated and accurate camera calibration process can be performed for a multiple-sensor integrated system.

REFERENCES

[1] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation into object space coordinates in close-range photogrammetry," in *Proc. Symp. on Close-Range Photogrammetry*, Urbana, IL, Jan. 1971, pp. 1–18.

[2] H. Bacakoglu and M. S. Kamel, "A three-step camera calibration method," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 1165–1172, 1997.

[3] C. Chatterjee, V. P. Roychowdhury, and E. K. P. Chong, "A nonlinear Gauss–Seidel algorithm for noncoplanar and coplanar camera calibration with convergence analysis," *Computer Vision and Image Understanding*, vol. 67, no. 1, pp. 58–80, 1997.

[4] J. Z. C. Lai, "On the sensitivity of camera calibration," *Image and Vision Computing*, vol. 11, pp. 656–664, 1993.

[5] R. K. Lenz and R. Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, pp. 713–720, 1988.

[6] American Society of Photogrammetry, *Manual of Photogrammetry*, 4th ed., 1980.

[7] R. B. Nelson, "Simplified calculation of eigenvector deviations," *AIAA J.*, vol. 14, pp. 1201–1205, 1976.

[8] T. S. Shen, "Multiple-sensor integration for rapid and high-precision coordinate metrology," Ph.D. dissertation, Dep. Mech. Eng., Ohio State Univ., Columbus, 2000.

[9] T. S. Shen, J. Huang, and C. H. Menq, "Multiple-sensor integration for rapid and high-precision coordinate metrology," *IEEE/ASME Trans. Mechatron.*, vol. 5, pp. 110–121, 2000.

[10] S. W. Shih, Y. P. Hung, and W. S. Lin, "Accurate linear technique for camera calibration considering lens distortion by solving an eigenvalue problem," *Opt. Eng.*, vol. 32, no. 1, pp. 138–149, 1993.

[11] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV camera and lenses," *IEEE Trans. Robot. Automat.*, vol. RA-3, pp. 323–344, Aug. 1987.

[12] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, pp. 965–980, Oct. 1992.

A New Partitioned Approach to Image-Based Visual Servo Control

Peter I. Corke and Seth A. Hutchinson

***Abstract*—In image-based visual servo control, where control is effected with respect to the image, there is no direct control over the Cartesian velocities of the robot end effector. As a result, the robot executes trajectories that are desirable in the image, but which can be indirect and seemingly contorted in Cartesian space. In this paper, we introduce a new partitioned approach to visual servo control that overcomes this problem. In particular, we decouple the z-axis rotational and translational components of the control from the remaining degrees of freedom. Then, to guarantee that all features remain in the image throughout the entire trajectory, we incorporate a potential function that repels feature points from the boundary of the image plane. We illustrate our new control scheme with a variety of results.**

***Index Terms*—Image-based visual servo, potential field, visual servo.**

I. INTRODUCTION

In image-based visual servo (IBVS) control, an error signal is measured in the image and mapped directly to actuator commands (see, e.g., [1] and [2]). This in contrast to posititon-based visual servo (PBVS) systems in which extracted features are used to compute a (partial) 3-D reconstruction of the environment or of the motion of a target object in the environment [3]. An error is then computed in the task space and it is this error that is used by the control system.

IBVS approaches have seen increasing popularity, largely due to the shortcomings of PBVS systems. With PBVS, any errors in calibration of the vision system will lead to errors in the 3-D reconstruction and subsequently to errors during task execution. In addition, since the control law for PBVS is defined in terms of the 3-D workspace, there is no mechanism by which the image is directly regulated. Thus, it is possible that objects of interest (including features that are being used by the visual servo system) can exit the camera's field of view.

However, there are problems associated with IBVS systems. For an IBVS system, the control law involves the mapping between image space velocities and velocities in the robot's workspace. This mapping is encoded in the image Jacobian (which will be briefly reviewed in Section II). Singularities or poor conditioning in this Jacobian (which occur as a function of the relative position and motion of the camera and the object under observation) lead to control problems. Secondly, since control is effected with respect to the image, there is no direct control

Manuscript received November 29, 2000. This paper was recommended for publication by Associate Editor P. Allen and Editor A. De Luca upon evaluation of the reviewers' comments.

P. I. Corke is with CSIRO Manufacturing Science & Technology, Pinjarra Hills, Australia 4069.

S. A. Hutchinson is with the Beckman Institute for Advanced Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

Publisher Item Identifier S 1042-296X(01)08447-6.