行政院國家科學委員會專題研究計畫成果報告

機器臂視覺伺服控制與視覺系統參數校準
Robot Vision Servo Control and Calibration of the Camera System

計畫編號: NSC89-2213-E-009-216
執行期間: 89年8月1日至90年7月31日
主持人: 林錫寬教授 交通大學電機與控制系
計畫參與人員: 方政加 交通大學電機與控制系

1 Abstract

This project presents a camera calibration method and a novel visual servo control scheme. The camera calibration method takes into account the lens distortion and can reach the reconstruction errors of about 0.2 mm. We propose a criterion to measure the performance of the depth estimation. which turns out to be a control objective function. The proposed visual servo control scheme has good performance in both the depth estimation and the visual control.

Keywords: Robot Control, Visual Servo Control, Camera Calibration

中文摘要

此計畫提出攝影機參數校準的新方法和機器臂視覺伺服控制的新法則。攝影機參數校準方法考慮到鏡頭曲度，也達到重建精度為0.2mm。機器臂視覺伺服控制法，我們先提出深度估測的性能指標，再依此導出新的控制法則。這個控制法不僅有良好的控制性能，同時能提供良好的深度估測。因為篇幅關係，攝影機參數校準不討論。

關鍵詞: 機器臂控制、視覺伺服控制、鏡頭參數校準

2 Introduction

In the literature, there are two types of the visual servo controller: one is *feature-based method* [3, 4, 5, 9, 6, 8, 11] the other is *position-based method* [10]. The main distinction is the input. The unknown depths are in the feature Jacobian matrix of feature-based control, the depth estimation is needed in visual servo control. Although there have been several methods proposed for the depth estimation, the popular one is the extended Kalman filter (EKF) [10, 7, 12].

However, little attention has been paid to the effect of the velocity of the camera on the depth observability. Dayawansa *et al.* [1] proposed a necessary and sufficient condition for the perspective observability problem.

In this project, we propose a criterion to measure the performance of the depth estimation. which turns out to be a control objective function. We suggest to use the variance as an index for the performance of the depth estimation. This index is also verified by simulations and experiments as a rule of thumb for the performance evaluation of the EKF, especially for slow camera velocity. We then try to develop a new visual servo control scheme by making the index as small as possible while having little effect on the control performance, so that the depth estimation is improved. Finally, a simulation example shows that the resulting control scheme reaches this goal.

3 Performance Index for Depth Estimation

Consider a point P with coordinates (X, Y, Z) with respect to the camera frame E_{XYZ}. The image of point P projected onto the image plane is denoted by p with coordinates $(x, y, 0)$ with respect to E_{xyz}. Let $\gamma_x = f_e/S_x$ and $\gamma_y = f_e/S_y$, where S_x, S_y are, respectively, the horizontal and vertical lengths per pixel on the camera sensing array.

Assume that the linear velocity and the angular velocity of the camera are $\mathbf{v}$ and $\boldsymbol{\omega}$, respectively, with respect to E_{XYZ}. Let $\boldsymbol{\xi} \equiv [x,\ y,\ Z]^T$ and $\mathbf{u} \equiv [\mathbf{v}^T,\ \boldsymbol{\omega}^T]^T$. The present system is the nonlinear system of equations

$$\dot{\boldsymbol{\xi}} = \mathbf{f}(\boldsymbol{\xi}, \mathbf{u}) \equiv \mathbf{G}(\boldsymbol{\xi})\mathbf{u} \tag{1}$$

$$\boldsymbol{\psi} = \mathbf{h}(\boldsymbol{\xi}) \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \boldsymbol{\xi} \tag{2}$$

where

$$\mathbf{G}(\boldsymbol{\xi}) = \begin{bmatrix} & & & \bar{\mathbf{J}}(\boldsymbol{\xi}) & & \\ 0 & 0 & -1 & -\frac{yZ}{\gamma_y} & \frac{xZ}{\gamma_x} & 0 \end{bmatrix} \tag{3}$$

$$\bar{\mathbf{J}}(\boldsymbol{\xi}) = \begin{bmatrix} -\frac{\gamma_x}{Z} & 0 & \frac{x}{Z} & \frac{xy}{\gamma_y} & -\frac{\gamma_x^2 + x^2}{\gamma_x} & \frac{\gamma_x}{\gamma_y} y \\ 0 & -\frac{\gamma_y}{Z} & \frac{y}{Z} & \frac{\gamma_y^2 + y^2}{\gamma_y} & -\frac{xy}{\gamma_x} & -\frac{\gamma_y}{\gamma_x} x \end{bmatrix} \tag{4}$$

In the state vector $\boldsymbol{\xi}$, x and y are the visual measurements, so the only unknown is the depth Z. Therefore, the state estimation of the system of (1) and (2) is equivalent to the *depth estimation* for the given camera velocity.

Let the linear velocity of the camera be decomposed into

$$\mathbf{v}(t) = \alpha_1(t)\mathbf{v}_1(\boldsymbol{\xi}) + \alpha_2(t)\mathbf{v}_2(\boldsymbol{\xi}) + \alpha_3(t)\mathbf{v}_3(\boldsymbol{\xi}) \tag{5}$$

where $\alpha_1(t)$, $\alpha_2(t)$, and $\alpha_3(t)$ are bounded functions of t, and

$$\mathbf{v}_1(\boldsymbol{\xi}) = \begin{bmatrix} -\gamma_x \\ 0 \\ x \end{bmatrix},\ \mathbf{v}_2(\boldsymbol{\xi}) = \begin{bmatrix} 0 \\ -\gamma_y \\ y \end{bmatrix},\ \mathbf{v}_3(\boldsymbol{\xi}) = \begin{bmatrix} \frac{x}{\gamma_x} \\ \frac{y}{\gamma_y} \\ 1 \end{bmatrix} \tag{6}$$

Proposition 1: *Consider the system of* (1) *and* (2) *with input* $\mathbf{u} = [\mathbf{v}^T,\ \boldsymbol{\omega}^T]^T$. *Suppose that the variation rate of the depth is small enough that the nonlinear depth estimator (such as the extended Kalman filter) can be approximated as a linear least squares estimator. When the norm of the linear velocity of the camera* $\|\mathbf{v}(t)\|$ *is fixed, a larger* $\mathcal{J}_o(\boldsymbol{\xi}, \mathbf{v})$ *guarantees a faster convergent rate of the depth estimation, where*

$$\mathcal{J}_o(\boldsymbol{\xi}, \mathbf{v}) = \mathbf{v}^T \mathbf{A}(\boldsymbol{\xi})\mathbf{v} \tag{7}$$

in which

$$\mathbf{A}(\boldsymbol{\xi}) = \frac{\mathbf{v}_1(\boldsymbol{\xi})\mathbf{v}_1^T(\boldsymbol{\xi})}{q_{11}} + \frac{\mathbf{v}_2(\boldsymbol{\xi})\mathbf{v}_2^T(\boldsymbol{\xi})}{q_{22}} \tag{8}$$

Since $\mathbf{v}_3$ *is orthogonal to* $\mathbf{v}_1$ *and* $\mathbf{v}_2$, $\mathbf{v}_3$ *is then in the null space of* $\mathbf{A}(\boldsymbol{\xi})$. ■

4 Visual Servo Control

The results of the last section will be used to design a visual servo control scheme. The concept is to correct the linear velocity $\mathbf{v}$ of the camera by increasing $\mathcal{J}_o(\boldsymbol{\xi}, \mathbf{v})$ as possible.

4.1 Control Scheme

First, we introduce the visual servo control scheme with the damped least-squares method (DLSM). Suppose that there are n feature points (x_i, y_i). Applying (**??**) to n feature points, we obtain

$$\dot{\mathbf{f}} = \mathbf{J}\mathbf{u} \tag{9}$$

where the feature vector $\mathbf{f}$ and the visual Jacobian matrix $\mathbf{J}$ are

$$\mathbf{f} = \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{bmatrix},\ \mathbf{J} = \begin{bmatrix} \bar{\mathbf{J}}(\boldsymbol{\xi}_1) \\ \vdots \\ \bar{\mathbf{J}}(\boldsymbol{\xi}_n) \end{bmatrix} \tag{10}$$

Note that $\mathbf{J} \in \mathcal{R}^{2n\times 6}$. The control purpose is to design the velocity of the camera $\mathbf{u}$, so that the rate of change of the feature points in the image of the camera follows the desired one.

The damped least-squares control scheme is to minimize $||\mathbf{J}\mathbf{u} - \dot{\mathbf{f}}^*||^2 + \rho_s^2||\mathbf{u}||^2$, where $\dot{\mathbf{f}}^*$ is the feature velocity command and $\rho_s \in \mathcal{R}$ is the damping factor which represents the weighting of $\|\mathbf{u}\|^2$ with respect to the feature velocity error $\|\dot{\mathbf{f}} - \dot{\mathbf{f}}^*\|$. The control command $\mathbf{u}^*$ is the optimal solution

$$\mathbf{u}^* = (\mathbf{J}^T\mathbf{J} + \rho_s^2\mathbf{I})^{-1}\mathbf{J}^T\dot{\mathbf{f}}^* \tag{11}$$

where $\mathbf{I}$ is the identity matrix. A nonzero ρ_s makes $(\mathbf{J}^T\mathbf{J}+\rho_s^2\mathbf{I})$ positive definite, even if $\mathbf{J}^T\mathbf{J}$ is singular. Although this control scheme can alleviate the singularity problem, the input velocity $\mathbf{u}^*$ may not help the performance of the depth estimation Z_i, which is required by (4). To compensate for this drawback, Proposition **??** motivates us to minimize the following objective function $\mathcal{J}(\mathbf{u})$:

$$\mathcal{J}(\mathbf{u}) = \|\mathbf{J}\mathbf{u}-\dot{\mathbf{f}}^*\|^2+\rho_s^2\|\mathbf{u}\|^2-\frac{\rho_s^2}{n}\sum_{i=1}^{n}\rho_{oi}^2\mathcal{J}_o(x_i, y_i, \mathbf{u}) \tag{12}$$

where

$$\mathcal{J}_o(x_i, y_i, \mathbf{u}) = \mathbf{u}^T \bar{\mathbf{A}}_i \mathbf{u} \quad (13)$$

$$\bar{\mathbf{A}}_i = \begin{bmatrix} \mathbf{A}(\boldsymbol{\xi}_i) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{R}^{6\times 6} \quad (14)$$

in which $\mathbf{A}(\boldsymbol{\xi}_i)$ is a matrix function defined in (8) for the ith point. The first two terms on the right-hand side of (12) are those in the damped least-squares method. The term of $\mathcal{J}_o(\boldsymbol{\xi}, \mathbf{v})$ in (12) is for improving the depth estimation. The factor ρ_{oi} is a weighting factor to compromise the control error and the performance of the depth estimation.

Since matrix $\mathbf{A}(\boldsymbol{\xi}_i)$ is symmetric, $\bar{\mathbf{A}}_i$ is orthogonally diagonalizable: $\bar{\mathbf{A}}_i = \mathbf{U}_i^T \boldsymbol{\Lambda}_i \mathbf{U}_i$, where $\boldsymbol{\Lambda}_i = \text{diag}(\sigma_{i1}, \sigma_{i2}, 0, 0, 0, 0) \in \mathcal{R}^{6\times 6}$, σ_{i1} and σ_{i2} are the positive eigenvalues of $\mathbf{A}(\boldsymbol{\xi}_i)$, and $\mathbf{U}_i$ is an orthogonal matrix. Thus, (13) is rewritten as

$$\mathcal{J}_o(x_i, y_i, \mathbf{u}) = \mathbf{u}^T (\mathbf{U}_i^T \boldsymbol{\Lambda}_i \mathbf{U}_i) \mathbf{u} \quad (15)$$

By calculus, we set $\partial \mathcal{J}(\mathbf{u})/\partial \mathbf{u} = \mathbf{0}$ to obtain the optimal solution $\mathbf{u}^*$ to (12) as

$$\mathbf{u}^* = \mathbf{W}^{-1} \mathbf{J}^T \dot{\mathbf{f}}^* \quad (16)$$

where

$$\mathbf{W} = \mathbf{J}^T \mathbf{J} + \rho_s^2 \left[\mathbf{I} - \frac{1}{n} \sum_{i=1}^{n} \rho_{oi}^2 (\mathbf{U}_i^T \boldsymbol{\Lambda}_i \mathbf{U}_i) \right] \quad (17)$$

which is a positive definite symmetrical matrix if $\rho_{oi}^2 \max\{\sigma_{i1}, \sigma_{i2}\} \leq 1$, $\forall i = 1, \cdots, n$. We shall call the control law (16) the *observabilized damped least-squares method* (abbreviated as ODLSM). When $\rho_{oi}^2 < 1/\max\{\sigma_{i1}, \sigma_{i2}\}$, $\forall i = 1, \cdots, n$, it follows from (12) that $\mathcal{J}(\mathbf{u})$ is always positive.

Suppose the true values of depths can be obtained by a depth estimator like the extended Kalman filter in a few seconds. Thereafter, $\mathbf{J}$ in (10) is approximately calculated by the estimated values of depths. We can then expect that the overall system is asymptotically convergent. It is verified by the simulation described in the following.

4.2 Simulation Example

The simulation example considers three image points of the corners of a triangle pattern. The initial image coordinates of three corners are located at $(10, 63)$, $(90, 51)$, and $(29, 153)$ in pixels and their real initial depths are all 550 mm but unknown in this simulation. The desired image has three corner images at $(-50, -68)$, $(50, -68)$, and $(-50, 43)$ in pixels. Note that the final depths are all 450 mm corresponding to the desired image feature. Suppose that the estimates of the initial depths are $\hat{Z}_1 = 595$ mm, $\hat{Z}_2 = 599$ mm, and $\hat{Z}_3 = 596$ mm, i.e., the initial estimation errors are about 50 mm. The intrinsic parameters of the camera are as follows: the effective focal length $f_e = 16.53$ mm, the horizontal length per pixel $S_x = 0.0161$ mm/pixel, and the vertical length per pixel $S_y = 0.0189$ mm/pixel. The noise covariance matrices in EKF are $\mathbf{R} = \text{diag}(0.25 \text{ pixel}^2, \ 0.25 \text{ pixel}^2)$, and $\mathbf{Q} = \text{diag}(0.25 \text{ pixel}^2, \ 0.25 \text{ pixel}^2, \ 25^2 \text{ mm}^2)$. Both DLSM and ODLSM controllers (see (11) and (16)) have the same following data: the sampling period $T_s = 200$ ms, the proportional gain $K_p = 0.65$, the damping factor $\rho_s = 0.003$, and the weighting factor $\rho_{oi} = 0.9/\max\{\sigma_{i1}, \sigma_{i2}\}$, $i = 1, 2, 3$, for each corner.

The history of the error norm of the estimated depths, i.e., $\sqrt{\sum_{i=1}^{3} (Z_i - \hat{Z}_i)^2/3}$, in Fig. 1(a) reveals that the depth estimation in ODLSM controller is superior to that in DLSM controller. The steady-state error of the depth estimate in DLSM is about 7 mm, while that in ODLSM nearly vanishes for the same EKF estimator.

The feature errors in both the ODLSM and DLSM are almost the same and converge to zero as is shown in Fig. 1(b). That means using a moderate small ρ_{oi} in ODLSM affects little the convergence performance of the visual servo control.

Consequently, the advantage of the ODLSM is that a good depth estimate can be achieved while the convergence performance is retained.

5 Conclusion

This project presents a depth estimation crieterion and a novel visual servo control scheme. A simulation shows that using the proposed control scheme the improvement of the depth estimation is achieved without any sacrifice of

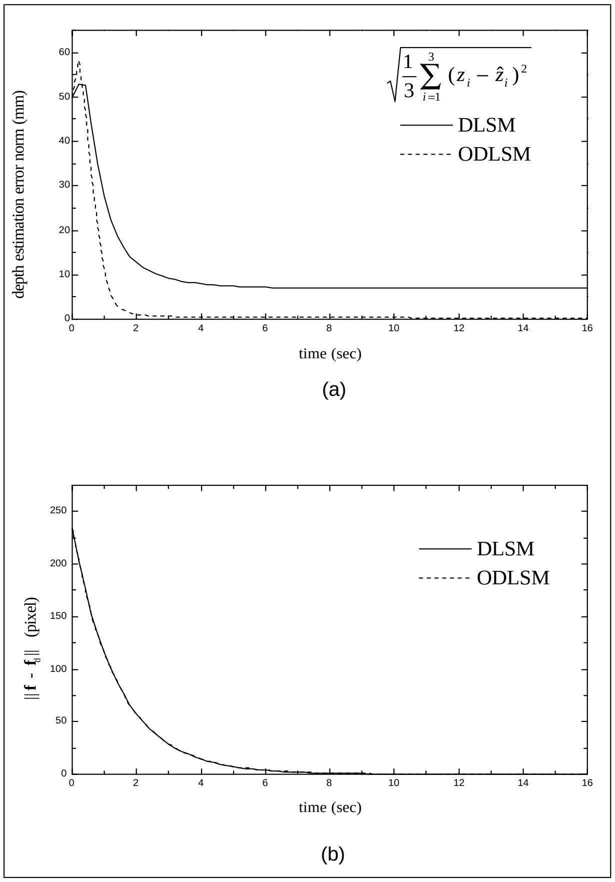

Fig. 1: (a) The norm of the depth estimation errors, (b) The norm of the feature feedback errors by DLSM and ODLSM.

the convergence performance. Due to limit of space, the result of the camera calibration method is omitted.

References

[1] W. P. Dayawansa, B. K. Ghosh, C. Martin, and X. Wang, "A necessary and sufficient condition for the perspective observability problem," *System & Control Letters*, vol. 25, pp. 159–166, 1995.

[2] C. E. Smith, S. A. Brandt, and N. P. Papanikolopoulos, "Eye-in-hand robotic tasks in uncalibrated environments," *IEEE Trans. Robotics Automat.*, vol. 13, no. 6, pp. 903–914, 1997.

[3] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Active self-calibration of robotic eyes and hand-eye relationships with model identification," *IEEE Trans. Robotics Automat.*, vol. 9, no. 1, pp. 14–34, 1993.

[4] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Adaptive robotic visual tracking: theory and experiments," *IEEE Trans. Robotics Automat.*, vol. 38, no. 3, pp. 429–445, 1993.

[5] J. T. Feddema and C. S. G. Lee, "Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera," *IEEE Trans. Syst. Man Cybern.*, vol. 20, no. 5, pp. 1172–1183, 1990.

[6] A. J. Koivo and N. Houshangi, "Real-time vision feedback for servoing robotic manipulator with self-tuning controller," *IEEE Trans. Syst. Man Cybern.*, vol. 21, no. 1, pp. 134–141, 1991.

[7] L. Matthies and T. Kanade, "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. Computer Vision*, vol. 3, pp. 209–236, 1989.

[8] W. Jang and Z. Bein, "Feature-based visual servoing of an eye-in-hand robot with improved tracking performance," in *proc. IEEE int. Conf. Robotics Automat.*, 1991, pp. 2254–2260.

[9] J. T. Feddema, C. S. G. Lee, and O.R. Mitchell, "Weighted selection of image features for resolved rate visual feedback control," *IEEE Trans. Robotics Automat.*, vol. 7, no. 1, pp. 31–47, 1991.

[10] W. J. Wilson , C. C. W Hulls, and G. S. Bell, "Relative end-effector control using cartesian position based visual servoing," *IEEE Trans. Robotics Automat.*, vol. 12, no. 5, pp. 684–696, 1996.

[11] K. Hashimoto, T. Ebine, and H. Kimura, "Visual servoing with hand-eye manipulator – optimal control approach," *IEEE Trans. Robotics Automat.*, vol. 12, no. 5, pp. 766–773, 1996.

[12] C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, New York, Spring-Verlag, 1991.