49. Internationales Wissenschaftliches Kolloquium
Technische Universität Ilmenau
27.-30. September 2004

M. Fujita / H. Kawai

Control and Estimation of Visual Feedback Systems with a Fixed Camera for Trajectory Tracking Problems

Abstract

This paper deals with the control and the estimation of the visual feedback systems with a fixed camera. Firstly the fundamental representation and the nonlinear observer for the visual feedback system with four coordinate frames is established. Secondly the visual feedback system is composed by the control error system and the estimation error system. Next, we derive the passivity of the visual feedback system. Finally, stability and L_2-gain performance analysis are discussed based on the passivity.

1 Introduction

Robotics and intelligent machines need many information to behave autonomously. Specifically, the combination of mechanical control with visual information, so-called visual feedback control or visual servoing, should become extremely important, when we consider a mechanical system working under dynamical environments [1]. Recently, technological fields which need visual feedback control are undoubtedly increasing, such as the autonomous injection of biological cells [2]. Control will be more important for intelligent machines as future applications.

This paper deals with the control and the estimation of visual feedback systems with a fixed camera. Kelly [3] considered the set-point problems with a static target for the visual feedback system with a planar type manipulator. Cowan *et al.* [4] addressed the problems of the target tracking and the field of view for the 3D visual feedback system by using the navigation functions. More recently, the authors have discussed the rigid body motion (involving both translation and rotation) control problem of visual feedback systems with Eye-in-Hand configuration [5]. This configuration has only three coordinate frames, while visual feedback systems typically use four coordinate frames which consist of a world frame Σ_w, a target object frame Σ_o, a camera frame Σ_c and a hand (end-effector) frame Σ_h as in Fig. 1. Because the camera is attached to the end-effector of robots, the camera frame represents the hand one in Eye-in-Hand configuration.

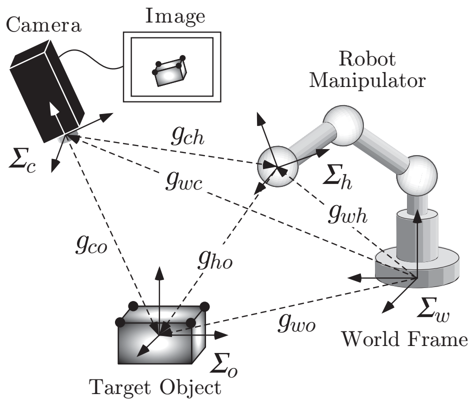

Figure 1: Visual feedback system with a fixed camera

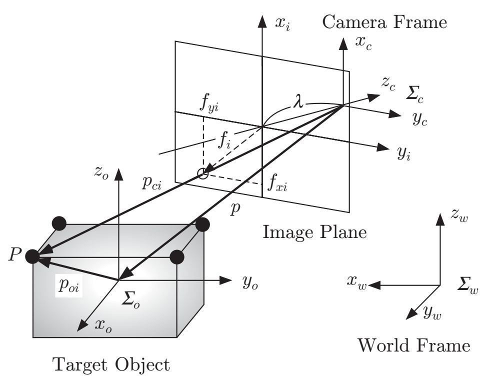

Figure 2: Pinhole camera

In this paper, we deal with the control and the estimation problems for visual feedback systems with the four coordinate frames. Extending the number of the coordinate frames from three to four, this framework can generalize our previous work [5]. In this framework, we can design the control gain and the observer gain separately from each other, while the control problem and the estimation problem of the visual feedback system are considered in the same strategy.

Throughout this paper, we use the notation $e^{\hat{\xi}\theta_{ab}} \in \mathcal{R}^{3\times3}$ to represent the change of the principle axes of a frame Σ_b relative to a frame Σ_a. $\xi_{ab} \in \mathcal{R}^3$ specifies the direction of rotation and $\theta_{ab} \in \mathcal{R}$ is the angle of rotation. Here $\hat{\xi}\theta_{ab}$ denotes $\hat{\xi}_{ab}\theta_{ab}$ for the simplicity of notation. The notation '$\wedge$' (wedge) is the skew-symmetric operator such that $\hat{\xi}\theta = \xi \times \theta$ for the vector cross-product $\times$ and any vector $\theta \in \mathcal{R}^3$. The notation '$\vee$' (vee) denotes the inverse operator to '$\wedge$': i.e., $so(3) \to \mathcal{R}^3$. $g_{ab} = (p_{ab}, e^{\hat{\xi}\theta_{ab}}) \in SE(3)$ is the description of the configuration of a frame Σ_b relative to a frame Σ_a [6]. All proofs in this paper are omitted due to space limitations.

2 Relative Rigid Body Motion in Visual Feedback System

2.1 Fundamental Representation for Visual Feedback System

Visual feedback systems typically use four coordinate frames which consist of a world frame Σ_w, a target object frame Σ_o, a camera frame Σ_c and a hand (end-effector) frame Σ_h as in Fig. 1. Then, $g_{wh} = (p_{wh}, e^{\hat{\xi}\theta_{wh}})$, $g_{wc} = (p_{wc}, e^{\hat{\xi}\theta_{wc}})$ and $g_{wo} = (p_{wo}, e^{\hat{\xi}\theta_{wo}})$ denote the rigid body motion from Σ_w to Σ_h, from Σ_w to Σ_c and from Σ_w to Σ_o, respectively. Using these coordinate frames, we will derive a fundamental representation for the three coordinate frames of the visual feedback system. The rigid body motion g_{co} of the target object, relative to the camera frame Σ_o in Fig. 1, is given by $g_{co} = g_{wc}^{-1}g_{wo}$ which is obtained from the composition rule for rigid body transformations ([6], Chap. 2, pp. 37, eq. (2.24)). Then, a fundamental representation for the three coordinate frames of the visual feedback system are described as follows [5].

$$V_{co}^b = -\mathrm{Ad}_{(g_{co}^{-1})}V_{wc}^b + V_{wo}^b \tag{1}$$

where $V_{co}^b := [v_{co}^T \ \omega_{co}^T]^T$ represents the body velocity of g_{co}. v_{co} and ω_{co} are the velocity of the origin and the angular velocity from Σ_c to Σ_o, respectively. Similarly, $V_{wc}^b := [v_{wc}^T \ \omega_{wc}^T]^T$ and $V_{wo}^b := [v_{wo}^T \ \omega_{wo}^T]^T$ represent the body velocities of g_{wc} and g_{wo}. The notation $\mathrm{Ad}_{(g)}$ denotes the adjoint transformation associated with g_{ab} [6].

2.2 Camera Model

Next, we derive the model of a pinhole camera with a perspective projection as shown in Fig. 2. Let λ be a focal length, $p_{oi} \in \mathcal{R}^3$ and $p_{ci} \in \mathcal{R}^3$ be coordinates of the target object's i-th feature point relative to Σ_o and Σ_c, respectively. Using a transformation of the coordinates, we have $p_{ci} = g_{co}p_{oi}$ where p_{ci} and p_{oi} should be regarded as $[p_{ci}^T \ 1]^T$ and $[p_{oi}^T \ 1]^T$ via the well-known representation in robotics, respectively (see, e.g., [6]). The perspective projection of the i-th feature point onto the image plane gives us the image plane coordinate $f_i := [f_{xi} \ f_{yi}]^T \in \mathcal{R}^2$ as follows

$$f_i = \frac{\lambda}{z_{ci}} \begin{bmatrix} x_{ci} \\ y_{ci} \end{bmatrix} \tag{2}$$

where $p_{ci} := [x_{ci} \ y_{ci} \ z_{ci}]^T$. It is straightforward to extend this model to the m image points case by simply stacking the vectors of the image plane coordinate, i.e. $f := [f_1^T \ \cdots \ f_m^T]^T \in \mathcal{R}^{2m}$. We assume that multiple point features on a known object can be used.

2.3 Nonlinear Observer for Visual Feedback System

The visual information f which includes the relative rigid body motion can be exploited, while the relative rigid body motion g_{co} can not be obtained directly in the visual feedback system. Hence, we consider a nonlinear observer in order to estimate the relative rigid body motion from the image information. In the case of the fixed camera configuration, i.e. $V_{wc}^b = 0$, (1) is rewritten as $V_{co}^b = V_{wo}^b$. Hence, we shall consider the following model which is reproduced from the above relation.

$$\bar{V}_{co}^b = u_e \tag{3}$$

where $\bar{V}_{co}^b := [\bar{v}_{co}^T \ \bar{\omega}_{co}^T]^T$ means the estimated body velocity. Here, $\bar{g}_{co} = (\bar{p}_{co}, e^{\hat{\bar{\xi}}\bar{\theta}_{co}})$ denotes the estimated relative rigid body motion. The new input u_e is to be determined in order to converge the estimated value to the actual relative rigid body motion. Because the design of u_e needs a property of the whole visual feedback system, we will propose u_e in Section 3.

Similarly to (2), the estimated image feature point $\bar{f}_i$ $(i = 1, \cdots, m)$ should be described as

$$\bar{f}_i = \frac{\lambda}{\bar{z}_{ci}} \begin{bmatrix} \bar{x}_{ci} \\ \bar{y}_{ci} \end{bmatrix} \tag{4}$$

where $\bar{p}_{ci} = \bar{g}_{co} p_{oi}$ and $\bar{p}_{ci} := [\bar{x}_{ci} \ \bar{y}_{ci} \ \bar{z}_{ci}]^T$. $\bar{f} := [\bar{f}_1^T \ \cdots \ \bar{f}_m^T]^T \in \mathcal{R}^{2m}$ means the m image points case.

In order to establish the estimation error system, we define the estimation error between the estimated value $\bar{g}_{co}$ and the actual relative rigid body motion g_{co} as $g_{ee} = \bar{g}_{co}^{-1} g_{co}$. Using the notation $e_R(e^{\hat{\xi}\theta})$, the vector of the estimation error is given by $e_e := [p_{ee}^T \ e_R^T(e^{\hat{\xi}\theta_{ee}})]^T$.

As mentioned in [5], the relation between the actual image information and the estimated one can be given by

$$f - \bar{f} = J(\bar{g}_{co}) e_e \tag{5}$$

where $J(\bar{g}_{co})$ is defined in [5]. We assume that the matrix $J(\bar{g}_{co})$ is full column rank for all $\bar{g}_{co} \in SE(3)$. Then, the relative rigid body motion can be uniquely defined by the image feature vector.

The above discussion shows that we can derive the vector of the estimation error e_e from image information f and the estimated value of the relative rigid body motion $\bar{g}_{co}$,

$$e_e = J^{\dagger}(\bar{g}_{co})(f - \bar{f}) \tag{6}$$

where $\dagger$ denotes the pseudo-inverse. Therefore the estimation error e_e can be exploited in the 3D visual feedback control law using image information f obtained from the camera. Hence, the nonlinear observer is constructed by (3)–(4) and the estimation input u_e which can be determined from e_e in (6) with an estimation gain in Section 3.4.

2.4 Estimation Error System

Differentiating the estimation error g_{ee}, we can obtain the estimation error system as follows

$$V_{ee}^b = -\mathrm{Ad}_{(g_{ee}^{-1})} u_e + V_{wo}^b. \tag{7}$$

It should be noted that if the vector of the estimation error is equal to zero, then the estimated relative rigid body motion $\bar{g}_{co}$ equals the actual one g_{co}.

3 Passivity-based Visual Feedback Control

3.1 Control Error System

From the composition rule for rigid body transformations, the relative rigid body motion from Σ_h to Σ_o is described as $g_{ho} = g_{ch}^{-1} g_{co}$. Because g_{co} can not be obtained directly, we represent the relative rigid body motion from Σ_h to Σ_o with the estimated one $\bar{g}_{co}$ as $\bar{g}_{ho} = g_{ch}^{-1}\bar{g}_{co}$. It is supposed that the relative rigid body motion from Σ_c to Σ_h, i.e. g_{ch}, can be measured exactly. Since the problem of the camera calibration is one of important research topics and good solutions to it are reported in some papers (see, e.g., [7]), we will not consider the error of the camera calibration in this paper.

Then, the fundamental representation of the relative rigid body motion $\bar{g}_{ho}$ will be obtained in the same way as (1).

$$\bar{V}_{ho}^b = -\mathrm{Ad}_{(\bar{g}_{ho}^{-1})} V_{wh}^b + u_e \tag{8}$$

where we exploit (3) and the relation $V_{ch}^b = V_{wh}^b$ which is derived from the composition rule. Here we define the control error between the estimated value $\bar{g}_{ho}$ and the reference of the relative rigid body motion g_d as $g_{ec} = g_d^{-1}\bar{g}_{ho}$. It should be remarked that the estimated relative rigid body motion equals the reference one if and only if the control error is equal to the identity matrix in matrix form, i.e. $p_d = \bar{p}_{ho}$ and $e^{\hat{\xi}\theta_d} = e^{\hat{\xi}\bar{\theta}_{ho}}$ iff $g_{ec} = I_4$. Using the notation $e_R(e^{\hat{\xi}\theta})$, the vector of the control error is defined as $e_c := [p_{ec}^T \; e_R^T(e^{\hat{\xi}\theta_{ec}})]^T$. Note that $e_c = 0$ iff $p_{ec} = 0$ and $e^{\hat{\xi}\theta_{ec}} = I_3$.

Similarly to (7), the model of the control error can be obtained as

$$V_{ec}^b = -\mathrm{Ad}_{(\bar{g}_{ho}^{-1})} V_{wh}^b + u_e - \mathrm{Ad}_{(g_{ec}^{-1})} V_d^b \tag{9}$$

where $V_d^b := [v_d^T \; \omega_d^T]^T$ and $\hat{V}_d^b := g_d^{-1}\dot{g}_d$.

3.2 Visual Feedback System

Combining (7) and (9), we construct the visual feedback system as follows

$$\begin{bmatrix} V_{ec}^b \\ V_{ee}^b \end{bmatrix} = \begin{bmatrix} -\mathrm{Ad}_{(\bar{g}_{ho}^{-1})} & I \\ 0 & -\mathrm{Ad}_{(g_{ee}^{-1})} \end{bmatrix} \begin{bmatrix} V_{wh}^b \\ u_e \end{bmatrix} + \begin{bmatrix} -\mathrm{Ad}_{(g_{ec}^{-1})} \\ 0 \end{bmatrix} V_d^b + \begin{bmatrix} 0 \\ I \end{bmatrix} V_{wo}^b. \tag{10}$$

Using the relation of the adjoint transformation, i.e. $\mathrm{Ad}_{(g_{ec}^{-1})} = \mathrm{Ad}_{(\bar{g}_{ho}^{-1})}\mathrm{Ad}_{(\bar{g}_d)}$, the above equation (10) can be rewritten as

$$\begin{bmatrix} V_{ec}^b \\ V_{ee}^b \end{bmatrix} = \begin{bmatrix} -\mathrm{Ad}_{(\bar{g}_{ho}^{-1})} & I \\ 0 & -\mathrm{Ad}_{(g_{ee}^{-1})} \end{bmatrix} u_{ce} + \begin{bmatrix} 0 \\ I \end{bmatrix} V_{wo}^b, \quad u_{ce} := \begin{bmatrix} V_{wh}^b + \mathrm{Ad}_{(g_d)} V_d^b \\ u_e \end{bmatrix} \tag{11}$$

denotes the control input for the visual feedback system. Let us define the error vector of the visual feedback system as $e := \left[e_c^T \; e_e^T\right]^T$. It should be noted that if the vectors of the control error and the estimation error are equal to zero, then the estimated relative rigid body motion $\bar{g}_{ho}$ equals the reference one g_d and the estimated one $\bar{g}_{co}$ equals the actual one g_{co}, respectively. Moreover, the error and the error vector between $\bar{h}_{co}$ and g_{ho} can be also represented as g_{ee} and e_e, while g_{ee} and e_e are defined as the error and the error vector between $\bar{g}_{co}$ and g_{co}. Therefore, the actual relative rigid body motion g_{ho} tends to the reference one g_d when $e \to 0$.

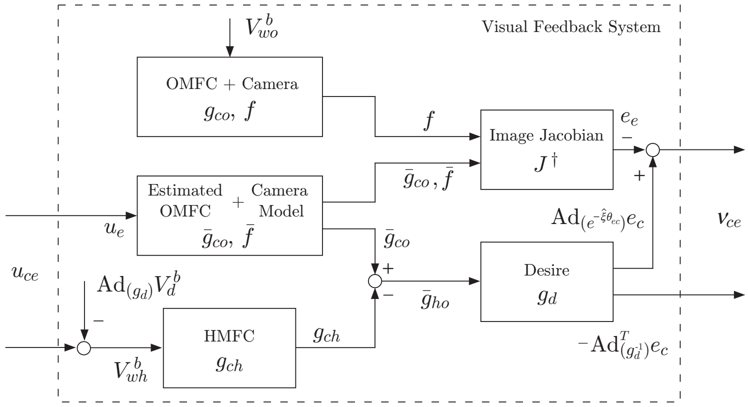

Figure 3: Block diagram of the visual feedback system

3.3 Property of Visual Feedback System

Now, we show an important lemma concerning a relation between the input and the output of the visual feedback system.

Lemma 1 If $V^b_{wo} = 0$, then the visual feedback system (11) satisfies

$$\int_0^T u^T_{ce}\nu_{ce} d\tau \geq -\beta_{ce}, \quad \forall T > 0 \tag{12}$$

where ν_{ce} is defined as

$$\nu_{ce} := N_{ce}e, \quad N_{ce} := \begin{bmatrix} -\mathrm{Ad}^T_{(g_d^{-1})} & 0 \\ \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{ec}})} & -I \end{bmatrix} \tag{13}$$

and β_{ce} is a positive scalar.

The block diagram of the visual feedback system is shown in Fig. 3. OMFC and HMFC represent the object motion relative to the camera frame Σ_o and the hand motion relative to the camera frame Σ_o, respectively.

Remark 1 Let us take u_{ce} as the input and ν_{ce} as its output in Fig. 3. Then, Lemma 1 would suggest that the visual feedback system (11) is *passive* from the input u_{ce} to the output ν_{ce} just formally.

3.4 Stability Analysis for Visual Feedback System

It is well known that there is a direct link between passivity and Lyapunov stability. Thus, we propose the following control input.

$$u_{ce} = -K_{ce}\nu_{ce} = -K_{ce}N_{ce}e, \quad K_{ce} := \begin{bmatrix} K_c & 0 \\ 0 & K_e \end{bmatrix} \tag{14}$$

where $K_c := \mathrm{diag}\{k_{c1}, \cdots, k_{c6}\}$ and $K_e := \mathrm{diag}\{k_{e1}, \cdots, k_{e6}\}$ are the positive gain matrices of x, y and z axes of the translation and the rotation for the control error and the estimation error, respectively. The result with respect to asymptotic stability of the proposed control input (14) can be established as follows.

Theorem 1 If $V^b_{wo} = 0$, then the equilibrium point $e = 0$ for the closed-loop system (11) and (14) is asymptotic stable.

Theorem 1 shows the stability via Lyapunov method for the full 3D visual feedback system. It is interesting to note that stability analysis is based on the passivity as described in (12).

3.5 L_2-gain Performance Analysis for Visual Feedback System

Based on the dissipative systems theory, we consider L_2-gain performance analysis for the visual feedback system (11) in one of the typical problems, i.e. the disturbance attenuation problem. Now, let us define

$$P := N_{ce}^T K_{ce} N_{ce} - \frac{1}{2\gamma^2} W - \frac{1}{2} I \tag{15}$$

where $\gamma \in \mathcal{R}$ is positive and $W := \mathrm{diag}\{0, I\}$. Then we have the following theorem.

Theorem 2 Given a positive scalar γ and consider the gains K_c and K_e such that the matrix P is positive semi-definite, then the closed-loop system (11) and (14) has L_2-gain $\leq \gamma$.

The L_2-gain performance analysis of the visual feedback system is discussed via the dissipative systems theory. In H_∞-type control, we can consider some problems by establishing the adequate generalized plant. This paper has discussed L_2-gain performance analysis for the disturbance attenuation problem. The proposed strategy can be extended for the other-type of generalized plants of the visual feedback systems.

4 Conclusions

This paper dealt with the control and the estimation of the visual feedback systems with a fixed camera. The main contribution of this work is that the visual feedback system with four coordinate frames is constructed in order to generalize our previous works. In this framework, we can design the control gain and the observer gain separately from each other, while the control problem and the estimation problem of the visual feedback system are considered in the same strategy. Stability and L_2-gain performance analysis for the visual feedback system have been discussed based on passivity.

References

[1] S. Hutchinson, G. D. Hager and P. I. Corke, "A Tutorial on Visual Servo Control," *IEEE Trans. Robotics and Automation*, Vol. 12, No. 5, pp. 651–670, 1996.

[2] S. Yu and B. J. Nelson, "Autonomous Injection of Biological Cells Using Visual Servoing," In:D. Rus and S. Singh (Eds.), *Experimental Robotics VII*, Springer–Verlag, pp. 169–178, 2001.

[3] R. Kelly, "Robust Asymptotically Stable Visual Servoing of Planar Robots," *IEEE Trans. Robotics and Automation*, Vol. 12, No. 5, pp. 759–766, 1996.

[4] N. J. Cowan, J. D. Weingarten and D. E. Koditschek, "Visual Servoing via Navigation Functions," *IEEE Trans. Robotics and Automation*, Vol. 18, No. 4, pp. 521–533, 2002.

[5] H. Kawai and M. Fujita An Interpretation of Vision-based Control for Rigid Body Motion: A Geometric Framework, Proc. of the 48th Internationales Wissenschaftliches Kolloquium Paper ID 11-02-02, 2003.

[6] R. Murray, Z. Li and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, 1994.

[7] B. E. Bishop and M. W. Spong, "Adaptive Calibration and Control of 2D Monocular Visual Servo Systems," *Control Engineering Practice*, Vol. 7, No. 3, pp. 423–430, 1999.

Author Information:

Prof. Dr. Masayuki Fujita
Department of Electrical and Electronic Engineering
Kanazawa University
Kodatsuno 2–40-20, Kanazawa 9208667, Japan
Tel: +81-76-234-4848
Fax: +81-76-234-4870
E-mail: fujita@t.kanazawa-u.ac.jp

Dr. Hiroyuki Kawai
Information Technology Research Center
Hosei University
Fujimi 2-17-1, Chiyoda-ku, Tokyo 1028160, Japan
Tel: +81-3-3264-9364
Fax: +81-3-3264-9287
E-mail: hiroyuki@i.hosei.ac.jp