Image-based Dynamic Visual Feedback Control via Passivity Approach

Hiroyuki Kawai†, Toshiyuki Murao‡ and Masayuki Fujita‡

***Abstract*— This paper deals with the image-based dynamic visual feedback control via the passivity approach which is investigated in the position-based one. We construct an energy function using the error in the image plane. The passivity of the visual feedback system is derived from the energy function. We show passivity of the image-based dynamic visual feedback system by combining the passivity of both the visual feedback system and the manipulator dynamics which allows us to prove stability in the sense of Lyapunov. The L_2-gain performance analysis, which deals with the disturbance attenuation problem, is then considered via dissipative systems theory. This paper suggests that the two types of classical visual servoing, i.e. the image-based visual servoing and the position-based visual servoing, can be discussed with the same strategy.**

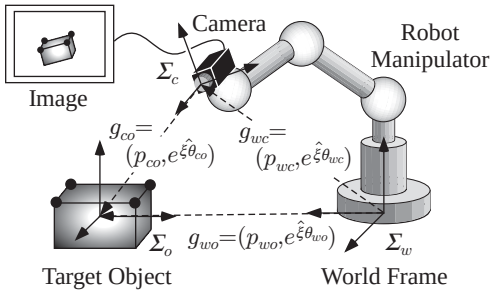

Fig. 1. Eye-in-hand visual feedback system.

I. INTRODUCTION

Visual feedback control of robotic systems involves the fusion of robot kinematics, dynamics, and computer vision to control the motion of the robot in an efficient manner. Visual feedback control is classified into two groups, position-based control and image-based control [1]. In position-based control, the references are given in the three-dimensional Cartesian space. The control objective is to bring a relative pose, which is the pose from a camera to a target *or* from a hand to a target, to a desired pose by using image information. In image-based control, the references are given in an image plane.

The 2 1/2-D visual servoing which incorporates the advantages of both position-based and image-based visual servoing is proposed in order to guarantee robustness with respect to calibration errors [2]. Partitioned visual servoing is considered in [3] in order to guarantee that all features remain in the image. More recently, an approach based on switching between position-based visual servoing and backward motion is investigated for dealing with the field of view problem [4]. However, classical visual servoing algorithms assume that the manipulator dynamics is negligible and does not interact with the visual feedback loop. This assumption, while it holds for kinematic control problems, is invalid for high speed tasks.

Maruyama *et al.* [5] discussed image-based robust control for the set-point problems of the dynamic visual feedback systems that includes the manipulator dynamics. Although the control law guarantees Lyapunov stability and is effective for the dynamic visual feedback system, it is restricted to planar manipulators. For the problem of 3D visual feedback control, Kelly *et al.* [6] considered an image-based controller under the assumption that the object's depths are known. Although good solutions to the set-point problem are reported in those papers, few results have been obtained for the tracking problem of moving target objects in the full 3D dynamic visual feedback system. For the target tracking problem, the authors have proposed position-based control via passivity approach [7], [8], [9]. However, image-based control is suitable than position-based control for some applications. For example, in recent applications of visual feedback system, laparoscopic surgery [10] is investigated by image-based control, because surgeons need a field of view of the surrounding target in the image plane. The objective of image-based visual feedback control is to bring the image feature points f to desired image feature points f_d.

This paper deals with image-based dynamic visual feedback control for a moving target object in 3D workspace with the eye-in-hand configuration as depicted in Fig. 1. Our proposed image-based control method does not need the assumption that the object's depths are known, because they are estimated by the same strategy in the position-based dynamic visual feedback control. In Section II, we provide an image dynamics with a brief summary of our previous work [7]. The visual feedback system is constructed in Section III and in Section IV, we show passivity of the image-based dynamic visual feedback system combined with the manipulator dynamics and present stability and L_2-gain performance analysis by using an energy function. In Section V we present simulation results for image-based dynamic visual feedback control on a 2DOF manipulator. Finally, we offer some conclusions in Section VI.

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. 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., $so(3) \to \mathcal{R}^3$. $\xi_{ab} \in \mathcal{R}^3$ specifies the direction of rotation and $\theta_{ab} \in \mathcal{R}$ is the

†Department of Robotics, Kanazawa Institute of Technology, Ishikawa 921-8501, Japan hiroyuki@neptune.kanazawa-it.ac.jp

‡Department of Mechanical and Control Engineering, Tokyo Institute of Technology, Tokyo 152-8550, Japan

angle of rotation. Here $\hat{\xi}\theta_{ab}$ denotes $\hat{\xi}_{ab}\theta_{ab}$ for the simplicity of notation. We use the 4×4 matrix

$$g_{ab} = \begin{bmatrix} e^{\hat{\xi}\theta_{ab}} & p_{ab} \\ 0 & 1 \end{bmatrix} \tag{1}$$

as the homogeneous representation of $g_{ab} = (p_{ab}, e^{\hat{\xi}\theta_{ab}}) \in SE(3)$ which is the description of the configuration of a frame Σ_b relative to a frame Σ_a. The adjoint transformation associated with g_{ab} is denoted by $\mathrm{Ad}_{(g_{ab})}$ [11].

II. Background

A. Basic Representation for Visual Feedback System

The visual feedback system considered in this paper has the camera mounted on the robot's end-effector as depicted in Fig. 1. Thus, the whole system has three coordinate frames which consist of a world (base of the manipulator) frame Σ_w, a camera (end-effector of the manipulator) frame Σ_c and a target object frame Σ_o. Then, the relative rigid body motion from Σ_c to Σ_o can be represented by g_{co}. Similarly, g_{wc} and g_{wo} denote the rigid body motions from the world frame Σ_w to the camera frame Σ_c and from the world frame Σ_w to the object frame Σ_o, respectively, as shown in Fig. 1.

The relative rigid body motion from Σ_c to Σ_o can be led by using the composition rule for rigid body transformations ([11], Chap. 2, pp. 37, eq. (2.24)) as follows

$$g_{co} = g_{wc}^{-1} g_{wo}. \tag{2}$$

The basic representation of the relative rigid body motion involves the velocity of each rigid body. To this aid, let us consider the velocity of a rigid body as described in [11]. Now, we define the body velocity of the camera relative to the world frame Σ_w as $V_{wc}^b = [v_{wc}^T \ \omega_{wc}^T]^T$, where v_{wc} and ω_{wc} represent the velocity of the origin and the angular velocity from Σ_w to Σ_c, respectively ([11] Chap. 2, eq. (2.55)).

Differentiating (2) with respect to time, the basic representation of the relative rigid body motion g_{co} is described as follows [7].

$$V_{co}^b = -\mathrm{Ad}_{(g_{co}^{-1})} V_{wc}^b + V_{wo}^b \tag{3}$$

where V_{wo}^b is the body velocity of the target object relative to Σ_w. Equation (3) is a standard formula for the relation between the body velocities of three coordinate frames ([11] Chap. 2, pp. 59, Proposition 2.15). Roughly speaking, the relative rigid body motion g_{co} will depend on the difference between the camera velocity V_{wc}^b and the target object velocity V_{wo}^b.

B. Image Dynamics

The pinhole camera model with a perspective projection is shown in Fig. 2. Let λ be a focal length, $p_{oi} \in \mathcal{R}^3$ and $p_{ci} \in \mathcal{R}^3$ be the position vectors 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

$$p_{ci} = g_{co} p_{oi}, \tag{4}$$

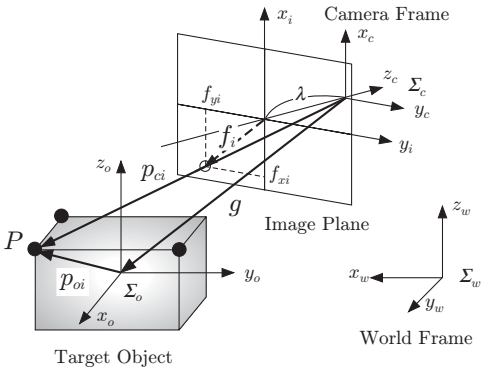

Fig. 2. Perspective projection in a simple camera model.

where p_{ci} and p_{oi} should be regarded, with a slight abuse of notation, as $[p_{ci}^T \ 1]^T$ and $[p_{oi}^T \ 1]^T$ via the well-known homogeneous coordinate representation in robotics, respectively (see, e.g., [11]).

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

$$f_i = \frac{\lambda}{z_{ci}} \begin{bmatrix} x_{ci} \\ y_{ci} \end{bmatrix} \tag{5}$$

where $p_{ci} = [x_{ci} \ y_{ci} \ z_{ci}]^T$. It is straightforward to extend this model to m image points by simply stacking the vectors of the image plane coordinate, i.e.,

$$f(g_{co}) := [f_1^T \ \cdots \ f_m^T]^T \in \mathcal{R}^{2m} \tag{6}$$

and $p_c := [p_{c1}^T \ \cdots \ p_{cm}^T]^T \in \mathcal{R}^{3m}$. We assume that multiple point features on a known object are given.

Differentiating (4) and (5), yields

$$\dot{f}_i = J_i(g_{co}) V_{co}^b \tag{7}$$

where $J_i(g_{co})$ is the image Jacobian defined by

$$J_i(g_{co}) = \begin{bmatrix} \frac{\lambda}{z_{ci}} & 0 & -\frac{\lambda x_{ci}}{z_{ci}^2} \\ 0 & \frac{\lambda}{z_{ci}} & -\frac{\lambda y_{ci}}{z_{ci}^2} \end{bmatrix} e^{\hat{\xi}\theta_{co}} \begin{bmatrix} I_3 & -\hat{p}_{oi} \end{bmatrix}. \tag{8}$$

Using (6) and (7), the time derivative of the image feature vector can be denoted by

$$\dot{f} = J(g_{co}) V_{co}^b \tag{9}$$

where

$$J(g_{co}) = \left[J_1^T(g_{co}) \ \cdots \ J_m^T(g_{co}) \right]^T. \tag{10}$$

Hereafter we assume that the matrix $J(g_{co})$ is full column rank for all $g_{co} \in SE(3)$ and the target object has three feature points, i.e. $\mathrm{rank}(J(g_{co})) = 6$ and $m = 3$. This assumption is required for technical reasons in stability and L_2-gain performance analyses.

The image dynamics is obtained by substituting (3) into (9)

$$\dot{f} = -J(g_{co}) \mathrm{Ad}_{(g_{co}^{-1})} V_{wc}^b + J(g_{co}) V_{wo}^b. \tag{11}$$

Lemma 1: If the target object is static, i.e. $V_{wo}^b = 0$, then the following inequality holds for the image dynamics (11).

$$\int_0^T (V_{wc}^b)^T(\nu_f)dt \geq -\beta_f \tag{12}$$

where $\nu_f := -\mathrm{Ad}_{(g_{co}^{-1})}^T J^T(g_{co})f$ and β_f is a positive scalar.

Proof: Consider the positive definite function

$$V_f = \frac{1}{2}\|f\|^2. \tag{13}$$

Differentiating (13) with respect to time yields

$$\dot{V}_f = f^T\dot{f} = -f^T J(g_{co})\mathrm{Ad}_{(g_{co}^{-1})}V_{wc}^b = (V_{wc}^b)^T\nu_f. \tag{14}$$

Integrating (14) from 0 to T, we obtain

$$\int_0^T (V_{wc}^b)^T\nu_f dt = V_f(T) - V_f(0) \geq -V_f(0) := -\beta_f \tag{15}$$

where β_f is a positive scalar that only depends on the initial states of f. ■

C. Nonlinear Observer

The visual information $f(g_{co})$ which includes the relative rigid body motion can be exploited, while the relative rigid body motion g_{co} can not be obtained directly. Firstly, we shall consider the following model which just comes from the basic representation (3).

$$\bar{V}_{co}^b = -\mathrm{Ad}_{(\bar{g}_{co}^{-1})}V_{wc}^b + u_e. \tag{16}$$

where $\bar{g}_{co}$ and $\bar{V}_{co}^b$ are the estimated value of the relative rigid body motion and the estimated body velocity from Σ_c to Σ_o, respectively. u_e is the input in order to converge the estimated value to the actual relative rigid body motion. Next, the estimation error of the relative rigid body motion from Σ_c to Σ_o, i.e. the error between $\bar{g}_{co}$ and g_{co}, is defined as

$$g_{ee} = \bar{g}_{co}^{-1}g_{co} \tag{17}$$

which is called the estimation error. Using the notation $e_R(e^{\hat{\xi}\theta})$, the vector of the estimation error is given by $e_e := [p_{ee}^T \ e_R^T(e^{\hat{\xi}\theta_{ee}})]^T$. Note that $e_e = 0$ iff $p_{ee} = 0$ and $e^{\hat{\xi}\theta_{ee}} = I_3$. The estimation error system is represented by

$$V_{ee}^b = -\mathrm{Ad}_{(g_{ee}^{-1})}u_e + V_{wo}^b. \tag{18}$$

Similarly to (4) and (5), the estimated image feature point $\bar{f}_i$ $(i = 1, 2, 3)$ is defined as

$$\bar{p}_{ci} = \bar{g}_{co}p_{oi} \tag{19}$$

$$\bar{f}_i = \frac{\lambda}{\bar{z}_{ci}}\begin{bmatrix} \bar{x}_{ci} \\ \bar{y}_{ci} \end{bmatrix} \tag{20}$$

where $\bar{p}_{ci} := [\bar{x}_{ci} \ \bar{y}_{ci} \ \bar{z}_{ci}]^T$. $\bar{f}(\bar{g}_{co}) := [\bar{f}_1^T \ \bar{f}_2^T \ \bar{f}_3^T]^T \in \mathcal{R}^6$ means 3 image points case. The relation between the actual image information and the estimated one can be given by

$$f - \bar{f} = J(\bar{g}_{co})e_e, \tag{21}$$

where $J(\bar{g}_{co})$ has the same form as the image Jacobian defined in (10). Similarly to $J(g_{co})$, we assume that the matrix $J(\bar{g}_{co})$ is full column rank for all $\bar{g}_{co} \in SE(3)$.

Remark 1: Generally, the image Jacobian expresses the relation between the motion of the feature point in the image plane and one in the workspace. As mentioned in [7], the relation between the estimation error in the image plane and one in the workspace can be connected by the same image Jacobian.

III. Image-based Visual Feedback System

Let us consider the image error dynamics in order to establish the image-based visual feedback system. First, we define the image error as follows:

$$f_e = \bar{f} - f_d \tag{22}$$

where $f_d \in \mathcal{R}^6$ denotes the desired image feature vector. Then, the image error dynamics can be represented as

$$\begin{aligned}\dot{f}_e &= J(\bar{g}_{co})\bar{V}_{co}^b \\ &= -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})}V_{wc}^b + J(\bar{g}_{co})u_e.\end{aligned} \tag{23}$$

Combining (18) and (23), we construct the image-based visual feedback system as follows:

$$\begin{bmatrix} \dot{f}_e \\ V_{ee}^b \end{bmatrix} = \begin{bmatrix} -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} & J(\bar{g}_{co}) \\ 0 & -\mathrm{Ad}_{(g_{ee}^{-1})} \end{bmatrix} u_{fe} + \begin{bmatrix} 0 \\ I \end{bmatrix} V_{wo}^b \tag{24}$$

where $u_{fe} := [(V_{wc}^b)^T \ u_e^T]^T$. We define the state of the image-based visual feedback system as $x_{fe} := [f_e^T \ e_e^T]^T$. It should be noted that if the vectors of the image error and the estimation error are equal to zero, then the image feature vector f, the estimated image feature vector $\bar{f}$, and the desired image feature vector f_d coincide. Therefore, the image feature vector f tends to the desired f_d when $x_{fe} \to 0$.

Lemma 2: If the target object is static, i.e. $V_{wo}^b = 0$, then the following inequality holds for the image-based visual feedback dynamics (24).

$$\int_0^T u_{fe}^T\nu_{fe} \geq -\beta_{fe} \tag{25}$$

where

$$\nu_{fe} := N_{fe}x_{fe}, \quad N_{fe} := \begin{bmatrix} -\mathrm{Ad}_{(\bar{g}_{co}^{-1})}^T J(\bar{g}_{co})^T & 0 \\ J(\bar{g}_{co})^T & -I \end{bmatrix} \tag{26}$$

and β_{fe} is a positive scalar.

Proof: Consider the positive definite function

$$V_{fe} = \frac{1}{2}\|f_e\|^2 + \frac{1}{2}\|p_{ee}\|^2 + \phi(e^{\hat{\xi}\theta_{ee}}) \tag{27}$$

where $\phi(e^{\hat{\xi}\theta}) := \frac{1}{2}\mathrm{tr}(I - e^{\hat{\xi}\theta})$ is the error function of the rotation matrix (see e.g. [12]). Differentiating (27) with respect to time yields

$$\begin{aligned}\dot{V}_{fe} &= \begin{bmatrix} f_e^T & e_e^T \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ee}})} \end{bmatrix}\begin{bmatrix} \dot{f}_e \\ V_{ee}^b \end{bmatrix} \\ &= \begin{bmatrix} f_e^T & e_e^T \end{bmatrix}\begin{bmatrix} -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} & J(\bar{g}_{co}) \\ 0 & -I \end{bmatrix} u_{fe} \\ &= u_{fe}^T\nu_{fe}\end{aligned} \tag{28}$$

Integrating (28) from 0 to T, we obtain

$$\int_0^T u_{fe}^T \nu_{fe} dt = V_{fe}(T) - V_{fe}(0) \geq -V_{fe}(0) := -\beta_{fe} \quad (29)$$

where β_{fe} is a positive scalar that only depends on the initial states of f_e and g_{ee}. ■

Let us consider u_{fe} as the input and ν_{fe} as its output. Then, Lemma 2 says that the image-based visual feedback system (24) is passive from the input u_{fe} to the output ν_{fe}.

IV. Image-based Dynamic Visual Feedback Control

A. Image-based Dynamic Visual Feedback System

The dynamics of n-link rigid robot manipulators can be written as

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + g(q) = \tau + \tau_d \quad (30)$$

where q, $\dot{q}$ and $\ddot{q}$ are the joint angle, velocity and acceleration, respectively, τ is the vector of the input torque, and τ_d represents a disturbance input [13]. Equation (30) possesses several important properties which will be used in the sequel. The manipulator dynamics (30) is passive from τ to $\dot{q}$, that is $\int_0^T \tau^T \dot{q} dt \geq -\beta_m$ where β_m is a positive scalar. Moreover, $\dot{M}(q) - 2C(q,\dot{q})$ is skew-symmetric by defining $C(q,\dot{q})$ using the Christoffel symbols. The visual feedback system (24) and the manipulator dynamics (30) will be connected by the passivity.

Now, we will construct an image-based dynamic visual feedback system by connecting the image-based visual feedback system (24) and the manipulator dynamics (30). Since the camera is mounted on the end-effector of the manipulator in the eye-in-hand configuration, the body velocity of the camera V_{wc}^b is given by

$$V_{wc}^b = J_b(q)\dot{q} \quad (31)$$

where $J_b(q)$ is the body manipulator Jacobian [11].

Next, we propose the control law for the manipulator as

$$\begin{aligned}\tau = {} & M(q)\ddot{q}_d + C(q,\dot{q})\dot{q}_d + g(q) \\ & + J_b(q)^T \mathrm{Ad}_{(\bar{g}_{co}^{-1})}^T J(\bar{g}_{co})^T f_e + u_\xi. \end{aligned} \quad (32)$$

where $\dot{q}_d$ and $\ddot{q}_d$ represent the desired joint velocity and acceleration, respectively. The first term in (32) is the compensation of the nonlinear effects.

Here, let us define the error vector with respect to the joint velocity of the manipulator dynamics as

$$\xi := \dot{q} - \dot{q}_d. \quad (33)$$

From the above discussion, we can obtain the following error dynamics

$$M(q)\dot{\xi} + C(q,\dot{q})\xi - J_b(q)^T \mathrm{Ad}_{(\bar{g}_{co}^{-1})}^T J(\bar{g}_{co})^T f_e = u_\xi + \tau_d. \quad (34)$$

Moreover, we design the reference of the joint velocity based on the relation between the camera velocity and the joint one (31) as follows:

$$\dot{q}_d := J_b^\dagger(q) u_d \quad (35)$$

where u_d is the desired body velocity of the camera which will be obtained from the image-based visual feedback system.

Using (24) and (34), the visual feedback system with manipulator dynamics (we call the image-based dynamic visual feedback system) can be derived as follows:

$$\begin{aligned}\begin{bmatrix} \dot{\xi} \\ \dot{f}_e \\ V_{ee}^b \end{bmatrix} = {} & \begin{bmatrix} -M(q)^{-1}\left(C(q,\dot{q})\xi + J_b(q)^T \mathrm{Ad}_{(\bar{g}_{co}^{-1})}^T J(\bar{g}_{co})^T f_e \right) \\ -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} J_b(q)\xi \\ 0 \end{bmatrix} \\ & + \begin{bmatrix} M(q)^{-1} & 0 & 0 \\ 0 & -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} & J(\bar{g}_{co}) \\ 0 & 0 & -\mathrm{Ad}_{(g_{ee}^{-1})} \end{bmatrix} u + \begin{bmatrix} M(q)^{-1} 0 \\ 0 \quad 0 \\ 0 \quad I \end{bmatrix} w \end{aligned} \quad (36)$$

where $u := [u_\xi^T \ u_d^T \ u_e^T]^T$. We define the state and the disturbance of image-based dynamic visual feedback system as $x := [\xi^T \ f_e^T \ e_e^T]^T$ and $w := [\tau_d^T \ (V_{wo}^b)^T]^T$, respectively.

Before constructing the image-based dynamic visual feedback control law, we derive an important lemma.

Lemma 3: If $w = 0$, then the image-based dynamic visual feedback system (36) satisfies

$$\int_0^T u^T \nu dt \geq -\beta, \quad \forall T > 0 \quad (37)$$

where

$$\nu := Nx, \quad N := \begin{bmatrix} I & 0 & 0 \\ 0 & -\mathrm{Ad}_{(\bar{g}_{co}^{-1})}^T J(\bar{g}_{co})^T & 0 \\ 0 & J(\bar{g}_{co})^T & -I \end{bmatrix} \quad (38)$$

and β is a positive scalar.

Proof: Consider the positive definite function

$$V = \frac{1}{2}\xi^T M(q)\xi + \frac{1}{2}\|f_e\|^2 + \frac{1}{2}\|p_{ee}\|^2 + \phi(e^{\hat{\xi}\theta_{ee}}) \quad (39)$$

Differentiating (39) with respect to time yields

$$\begin{aligned}\dot{V} &= \xi^T M(q)\dot{\xi} + \frac{1}{2}\xi^T \dot{M}(q)\xi + f_e^T \dot{f}_e + e_e^T \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ee}})} V_{ee}^b \\ &= \frac{1}{2}\xi^T \left(\dot{M}(q) - 2C(q,\dot{q}) \right)\xi + \xi^T u_\xi \\ &\quad + \begin{bmatrix} f_e^T & e_e^T \end{bmatrix} \begin{bmatrix} -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} & J(\bar{g}_{co}) \\ 0 & -I \end{bmatrix} \begin{bmatrix} u_d \\ u_e \end{bmatrix} \\ &= u^T \nu. \end{aligned} \quad (40)$$

Integrating (40) from 0 to T, we obtain

$$\int_0^T u^T \nu dt = V(T) - V(0) \geq -V(0) := -\beta \quad (41)$$

where β is a positive scalar that only depends on the initial states of ξ, f_e and g_{ee}. ■

Lemma 3 shows that the image-based dynamic visual feedback system preserves the passivity of both the image-based visual feedback system and the manipulator dynamics.

B. Dynamic Visual Feedback Control and Stability Analysis

We now propose the following control input for the interconnected system:

$$u = -K\nu = -KNx, \ K := \begin{bmatrix} K_\xi & 0 & 0 \\ 0 & K_f & 0 \\ 0 & 0 & K_e \end{bmatrix} \quad (42)$$

where $K_\xi := \mathrm{diag}\{k_{\xi 1}, \cdots, k_{\xi n}\}$ the positive gain matrix for each joint axis. $K_f := \mathrm{diag}\{k_{f1}, \cdots, k_{f6}\}$ and $K_e := \mathrm{diag}\{k_{e1}, \cdots, k_{e6}\}$ are the positive gain matrix for image errors and the positive gain matrix of x, y and z axes of the translation and the rotation for the estimation error, respectively.

Theorem 1: If $w = 0$, then the equilibrium point $x = 0$ for the closed-loop system (36) and (42) is asymptotic stable.

Proof: In the proof of Lemma 3, we have already derived that the time derivative of V along the trajectory of the system (36) is formulated as (40). Using the control input (42), (40) can be transformed into

$$\dot{V} = -x^T N^T K N x. \tag{43}$$

This completes the proof. ■

Theorem 1 shows the stability via Lyapunov method. It is interesting to note that stability analysis is based on the passivity as described in (37).

C. L_2-Gain Performance Analysis

The motion of the target object will be regarded as an external disturbance in our approach. Hence, we consider L_2-gain performance analysis for the image-based dynamic visual feedback system (36) in one of the typical problems based on the dissipative system theory, i.e. the disturbance attenuation problem. Here, we define the controlled controlled output as $z = [(\varepsilon x)^T \ (\rho u)^T]^T$ where ε and ρ are positive definite matrices to weight the state x and the auxiliary input u directly. Now, let us define

$$P := N^T K N - \frac{1}{2\gamma^2} W - \frac{1}{2}\|\varepsilon\|^2 - \frac{1}{2}\|\rho K N\|^2 \tag{44}$$

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

Theorem 2: Given a positive scalar γ and consider the control input (42) with the gains K_ξ, K_f and K_e such that the matrix P is positive semi-definite, then the closed-loop system (36) and (42) has L_2-gain $\leq \gamma$.

Proof: Differentiating the positive definite function V defined in (39) along the trajectory of the closed-loop system yields

$$\begin{aligned}\dot{V} = {} & \frac{\gamma^2}{2}\|w\|^2 - \frac{1}{2}\|z\|^2 - \frac{\gamma^2}{2}\|w\|^2 + \frac{1}{2}\|\varepsilon x\|^2 + \frac{1}{2}\|\rho u\|^2 \\ & + x^T \begin{bmatrix} I & 0 & 0 \\ 0 & -J(\bar{g}_{co})\mathrm{Ad}_{(\bar{g}_{co}^{-1})} & J(\bar{g}_{co}) \\ 0 & 0 & -I \end{bmatrix} u + x^T \begin{bmatrix} I & 0 \\ 0 & 0 \\ 0 & \mathrm{Ad}_{(e^{\hat{\xi}\theta_{ee}})} \end{bmatrix} w.\end{aligned}$$

By completing the squares, we have

$$\begin{aligned}\dot{V} + {} & \frac{1}{2}\|z\|^2 - \frac{\gamma^2}{2}\|w\|^2 \\ = {} & -\frac{\gamma^2}{2}\left\| w - \frac{1}{\gamma^2}\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{ee}})} \end{bmatrix} x \right\|^2 + \frac{1}{2}\|\varepsilon x\|^2 \\ & + \frac{1}{2\gamma^2}\left\| \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & \mathrm{Ad}_{(e^{-\hat{\xi}\theta_{ee}})} \end{bmatrix} x \right\|^2 + x^T N^T u + \frac{1}{2}\|\rho u\|^2 \\ \leq {} & \frac{1}{2\gamma^2} W\|x\|^2 + x^T N^T u + \frac{1}{2}\|\varepsilon x\|^2 + \frac{1}{2}\|\rho u\|^2. \end{aligned} \tag{45}$$

Substituting the control input (42) into (45), we obtain

$$\begin{aligned}\dot{V} + {} & \frac{1}{2}\|z\|^2 - \frac{\gamma^2}{2}\|w\|^2 \\ & \leq -x^T N^T K N x + \frac{1}{2\gamma^2} W\|x\|^2 + \frac{1}{2}\|\varepsilon x\|^2 + \frac{1}{2}\|\rho K N x\|^2. \end{aligned} \tag{46}$$

It can be verified that the inequality

$$\dot{V} + \frac{1}{2}\|z\|^2 - \frac{\gamma^2}{2}\|w\|^2 \leq -x^T P x \leq 0 \tag{47}$$

holds if P is positive semi-definite. Integrating (47) from 0 to T and noticing $V(T) \geq 0$, we have

$$\int_0^T \|z\|^2 dt \leq \gamma^2 \int_0^T \|w\|^2 dt + 2V(0), \ \forall T > 0. \tag{48}$$

This completes the proof. ■

We have discussed L_2-gain performance analysis for a disturbance attenuation problem. In other words, L_2-gain performance analysis is exploited to evaluate the tracking performance of the control scheme in the presence of a moving target.

V. Simulation

The simulation results on the two degree-of-freedom manipulator as depicted in Fig. 3 are shown in order to understand our proposed method simply, though it is valid for 3D visual feedback systems. The target object has three feature points and moves for $t = 4.8$ [s] moves along a straight line ($0 \leq t < 2$) and a "Figure 8" motion ($2 \leq t < 4.8$) as depicted in Fig. 4 and Fig. 5, respectively. The desired image feature points are set to $f_d = [-61 \ -61 \ -61 \ 61 \ 61 \ -61]^T$ pixels.

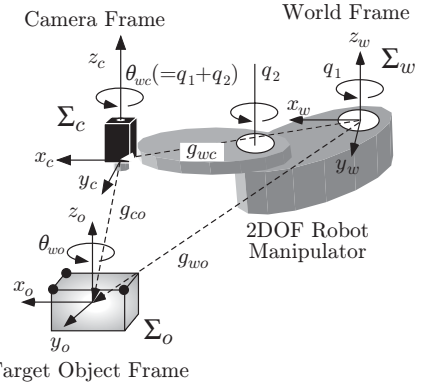

Fig. 3. Coordinate frames for image-based dynamic visual feedback system with two degree of freedom manipulator.

Firstly, we design the weight matrices concerning controlled output as $\varepsilon = \mathrm{diag}\{0.05, 0.05, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1.5, 1.5, 0.25, 0.25, 0.25, 1.5\}$, $\rho = \mathrm{diag}\{0.02, 0.02, 2, 2, 2, 2, 2, 2, 10, 10, 0.1, 0.1, 0.1, 10\} \times 10^{-3}$. Gains K_ξ, K_f and K_e are chosen as follows

$$\begin{aligned}&\text{Gain A} : K_\xi = \mathrm{diag}\{5, 5\}, \ K_f = 10I, \ K_e = 10I \\ &\text{Gain B} : K_\xi = \mathrm{diag}\{20, 10\}, \ K_f = 30I, \ K_e = 30I\end{aligned}$$

Then, the closed-loop system (36) and (42) with gain A has $\gamma = 0.912$ and with gain B has $\gamma = 0.381$.

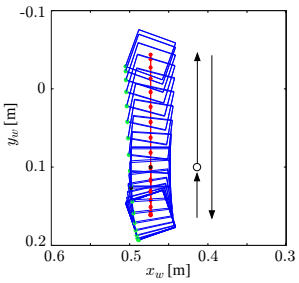

Fig. 4. Trajectory of target object along the straight line in $0 \leq t < 2$

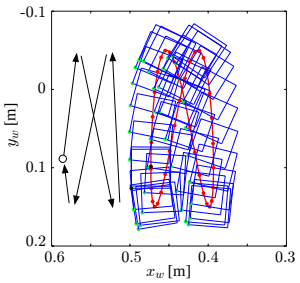

Fig. 5. Trajectory of target object along the "Figure 8" in $2 \leq t < 4.8$

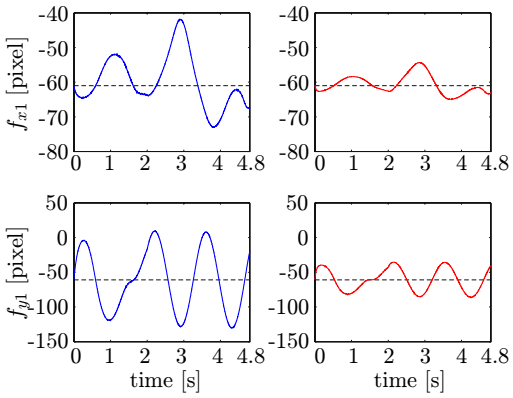

Fig. 6. Image feature point in the image plane with Gain A(left side) and Gain B(right side).

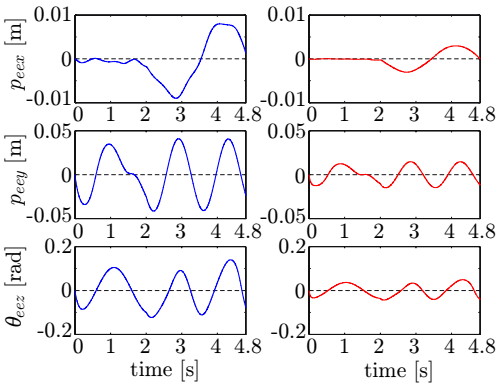

Fig. 7. Estimation error with Gain A(left side) and Gain B(right side).

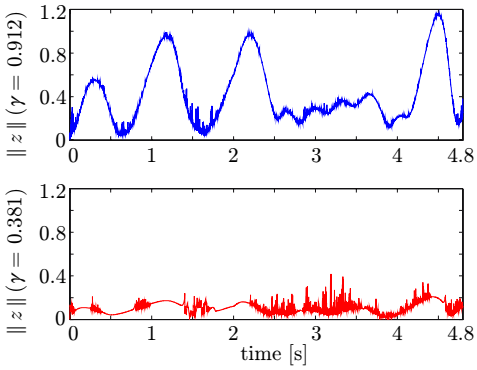

Fig. 8. Euclid norms of z(Gain A: top, Gain B: bottom)

Fig. 6 shows one of the image feature. The dashed lines are the desired position. The estimation error is depicted in Fig. 7. In both figures, the errors in the case of Gain A and Gain B are shown in the left side and the right one, respectively. Fig. 8 shows the norm of z in the case of $\gamma = 0.912$ and $\gamma = 0.381$. In the case of $\gamma = 0.381$, the performance is improved as compared to the case of $\gamma = 0.912$. After all, the simulation results show that L_2-gain is adequate for the performance measure of the image-based dynamic visual feedback control.

VI. Conclusions

This paper dealt with the image-based dynamic visual feedback control for three dimensional target tracking with the eye-in-hand configuration. Our proposed image-based control method does not need object's depths, because they are estimated by the same strategy in the position-based dynamic visual feedback control. The main contribution of this paper is to show that image-based control and position-based control can be discussed with the same passivity approach. In future work we aim to combine image-based control and position-based control in order to incorporate the advantages of both methods.

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] E. Malis, F. Chaumette and S. Boudet, "2-1/2-D Visual Servoing," *IEEE Trans. on Robotics and Automation*, Vol. 15, No. 2, pp. 238–250, 1999.

[3] P. I. Corke and S. A. Hutchinson, "A New Partitioned Approach to Image-Based Visual Servo Control," *IEEE Trans. on Robotics and Automation*, Vol. 17, No. 4, pp. 507–515, 2001.

[4] G. Chesi, K. Hashimoto, D. Prattichizzo and A. Vicino, "Keeping Features in the Field of View in Eye-In-Hand Visual Servoing: A Switching Approach," *IEEE Trans. on Robotics*, Vol. 20, No. 5, pp. 908–913, 2004.

[5] A. Maruyama and M. Fujita, "Robust Control for Planar Manipulators with Image Feature Parameter Potential," *Advanced Robotics*, Vol. 12, No. 1, pp. 67–80, 1998.

[6] R. Kelly, R. Carelli, O. Nasisi, B. Kuchen and F. Reyes, "Stable Visual Servoing of Camera-in-Hand Robotic Systems," *IEEE Trans. Mechatronics*, Vol. 5, No. 1, pp. 39–48, 2000.

[7] H. Kawai and M. Fujita, "Passivity-based Dynamic Visual Feedback Control for Three Dimensional Target Tracking:Stability and L_2-gain Performance Analysis," *Proc. of the 2004 American Control Conference*, pp. 1522–1527, 2004.

[8] M. Fujita, H. Kawai and M. W. Spong, "Passivity-based Dynamic Visual Feedback Control for Three Dimensional Target Tracking:Stability and L_2-gain Performance Analysis," IEEE Transactions on Control Systems Technology(submitted).

[9] T. Murao, H. Kawai and M. Fujita "Passivity-based Control of Visual Feedback Systems with Dynamic Movable Camera Configuration," *Proc. of the 44th IEEE Conference on Decision and Control and European Control Conference*, pp. 5360-5365, Seville, Spain, Dec. 2005.

[10] K. Omote *et al.*, "Self-Guided Robotic Camera Control for Laparoscopic Surgery Compared with Human Camera Control," *The American Journal of Surgery*, Vol. 177, No. 4, pp. 321–324, 1999.

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

[12] F. Bullo and R. Murray, "Tracking for Fully Actuated Mechanical Systems: a Geometric Framework," *Automatica*, Vol. 35, No. 1, pp. 17–34, 1999.

[13] M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*, John Wiley & Sons, 1989.