

# Visual Motion Observer-based Pose Control with Panoramic Camera via Passivity Approach

Hiroyuki Kawai, Toshiyuki Murao and Masayuki Fujita

**Abstract**—This paper considers the vision-based estimation and control with a panoramic camera via passivity approach. First, a hyperbolic projection of a panoramic camera is presented. Next, using standard body-attached coordinate frames (the world frame, mirror frame, camera frame and object frame), we represent the body velocity of the relative rigid body motion (position and orientation). After that, we propose a visual motion observer to estimate the relative rigid body motion from the measured camera data. We show that the estimation error system with a panoramic camera has the passivity which allows us to prove stability in the sense of Lyapunov. After that, stability and  $L_2$ -gain performance analysis for the closed-loop system are discussed. Finally, simulation and experimental results are shown in order to confirm the proposed method.

## I. INTRODUCTION

Vision based control uses the computer vision data to control the motion of a robot in an efficient manner. The combination of mechanical control with visual information, so-called visual feedback control or visual servoing, is important when we consider a mechanical system working under dynamical environments [1].

Recently, Lippiello *et al.* [2] presented a position-based visual servoing for a hybrid eye-in-hand/eye-to-hand multi-camera configuration by using the extended Kalman filter and a multiarm robotic cell. Gans and Hutchinson [3] proposed a hybrid switched-system control which utilizes image-based and position-based visual feedback control. Hu *et al.* [4] considered a homography-based robust visual servo control for the uncertainty of the camera calibration. In our previous works, we discussed the dynamic visual feedback control for 3D target tracking based on passivity [5][6]. Although these previous works give us the new vision-based robot control theory systematically, most of the works use a simple perspective projection by a pinhole camera.

On the other hand, omnidirectional cameras are useful for recognizing unknown surroundings widely. Geyer and Daniilidis [7] presented a unifying theory for all central panoramic systems, i.e., an equivalence of catadioptric and spherical projections. Mariottini *et al.* [8] reviewed the several epipolar geometry estimation algorithms by using a omnidirectional camera and give us Epipolar Geometry Toolbox which is a simulation environment with MATLAB. Fomena and Chaumette [9] considered improvements on modeling features for visual servoing using a spherical projection. Although these novel methods need a desired image a priori, these works focus on the theoretical contributions. From the more practical point of view, Vidal *et al.* [10]

H. Kawai is with Department of Robotics, Kanazawa Institute of Technology, Ishikawa 921-8501, Japan [hiroyuki@neptune.kanazawa-it.ac.jp](mailto:hiroyuki@neptune.kanazawa-it.ac.jp)  
T. Murao is with Master Program of Innovation for Design and Engineering Advanced Institute of Industrial Technology, Tokyo 140-0011, Japan  
M. Fujita is with Department of Mechanical and Control Engineering, Tokyo Institute of Technology, Tokyo 152-8550, Japan

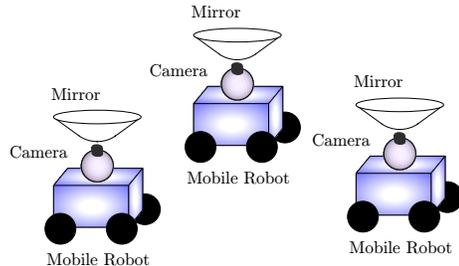


Fig. 1. Omnidirectional vision-based formation of mobile robots.

used motion segmentation techniques to estimate the position of each leader for the vision-based formation control of mobile robots with central panoramic cameras. Although the vision-based pose synchronization has been proposed in our previous work [11], a perspective projection model of a pinhole camera is used in order to estimate relative poses.

This paper deals with vision-based control by using a panoramic camera for mobile robot systems as depicted in Fig. 1. We propose the visual motion observer with a panoramic camera in order to estimate the relative rigid body motion (position and orientation). The main contribution of this paper is to show that the estimation error system with a panoramic camera has the passivity which allows us to prove stability in the sense of Lyapunov. After that, stability and  $L_2$ -gain performance analysis for the closed-loop system are discussed. Finally, simulation and experimental results are shown in order to confirm the proposed method.

## II. PANORAMIC CAMERA PROJECTION

### A. Hyperbolic Projection of Panoramic Camera

In this paper, we consider a panoramic camera which consists of a pinhole camera and a hyperbolic mirror as shown in Fig. 2. So, the pinhole camera catches reflected images through the hyperbolic mirror. We first review the panoramic camera model [8] in order to represent a image feature in our framework. Visual feedback systems by using a panoramic camera use four coordinate frames which consist of a world frame  $\Sigma_w$ , a mirror frame  $\Sigma_m$ , a camera frame  $\Sigma_c$ , and a object frame  $\Sigma_o$  as in Fig. 2. Let  $p_{mo} \in \mathcal{R}^3$  and  $e^{\hat{\xi}\theta_{mo}} \in SO(3)$  be the position vector and the rotation matrix from the mirror frame  $\Sigma_m$  to the object frame  $\Sigma_o$ . Then, the relative rigid body motion from  $\Sigma_m$  to  $\Sigma_o$  can be represented by  $g_{mo} = (p_{mo}, e^{\hat{\xi}\theta_{mo}}) \in SE(3)$ <sup>1</sup>. Similarly,  $g_{wm} = (p_{wm}, e^{\hat{\xi}\theta_{wm}})$ ,  $g_{wc} = (p_{wc}, e^{\hat{\xi}\theta_{wc}})$  and  $g_{wo} = (p_{wo}, e^{\hat{\xi}\theta_{wo}})$  denote the rigid body motions from the world frame  $\Sigma_w$  to the mirror frame  $\Sigma_m$ , from the world frame  $\Sigma_w$  to the camera frame  $\Sigma_c$  and from the world frame  $\Sigma_w$  to the object frame  $\Sigma_o$ , respectively.

<sup>1</sup>The notation of the homogeneous transform is referred to [5].

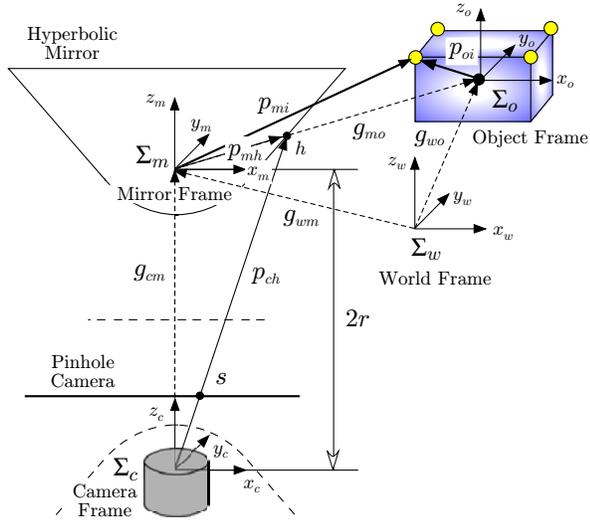


Fig. 2. Panoramic camera model (pinhole camera and hyperbolic mirror).

Here,  $p_{wo}$  is projected at  $s$  through the origin of the camera frame  $\Sigma_c$ , after being projected at  $p_{mh}$  through the origin of the mirror frame  $\Sigma_m$  as shown in Fig. 2. Let  $a$  and  $b$  be the hyperbolic mirror parameters which satisfy

$$\frac{(z_{mh} + r)^2}{a^2} - \frac{x_{mh}^2 + y_{mh}^2}{b^2} = 1 \quad (1)$$

with eccentricity  $r = \sqrt{a^2 + b^2}$ . Eq. (1) means a constraint with respect to  $p_{mh}$ . The transformation to obtain the projection  $s$  in the camera frame (see, Fig. 2) is given by

$$s = \frac{1}{z_{ch}} \Lambda \left( e^{\hat{\xi}\theta_{cm}} (\alpha e^{-\hat{\xi}\theta_{wm}} (p_{wo} - p_{wm})) + p_{cm} \right) \quad (2)$$

where  $\alpha$  is a scalar and defined as  $s := [f_x \ f_y \ 1]^T \in \mathcal{R}^3$  where  $f_x$  and  $f_y$  are the coordinates of  $x$ -axis and  $y$ -axis onto the image plane, respectively [8]. In this paper,  $\Lambda$  is assumed as the ideal internal calibration matrix of the pinhole camera and defined as  $\Lambda := \text{diag}\{\lambda, \lambda, 1\}$  where  $\lambda$  is a focal length.

From Fig. 2, the relation between the camera frame  $\Sigma_c$  and the mirror frame  $\Sigma_m$  can be represented as

$$p_{cm} = [0 \ 0 \ 2r]^T, \quad e^{\hat{\xi}\theta_{cm}} = I_3 \quad (3)$$

and it is assumed that these parameters are known. Moreover,  $p_{mh}$  can be represented as follows:

$$p_{mh} = \alpha p_{mo} \quad (4)$$

Because  $\alpha p_{mo}$  has to satisfy the constraint (1), the following relation holds

$$\frac{(\alpha z_{mo} + r)^2}{a^2} - \frac{\alpha^2 x_{mo}^2 + \alpha^2 y_{mo}^2}{b^2} = 1. \quad (5)$$

Solving (5) for  $\alpha$ , we obtain

$$\alpha(p_{mo}) = \frac{b^2(-r z_{mo} \pm a \|p_{mo}\|)}{b^2 z_{mo}^2 - a^2 x_{mo}^2 - a^2 y_{mo}^2} \quad (6)$$

where  $\alpha(p_{mo})$  represents that  $\alpha$  depends on  $p_{mo}$  explicitly<sup>2</sup>.

<sup>2</sup>Although  $\alpha(p_{mo})$  has two solutions, the suitable solution will be selected by  $p_{mo}$  [8].

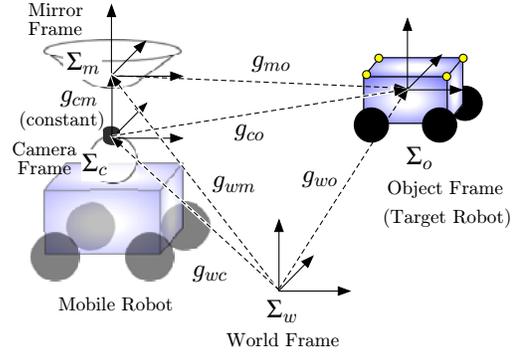


Fig. 3. Coordinate frames for mobile robots.

From Eqs. (2) and (3) and  $z_{ch} = 2r + z_{mh}$ , the hyperbolic projection of the panoramic camera can be represented as

$$s = \Lambda \frac{1}{2r + \alpha(p_{mo})z_{mo}} \left( \alpha(p_{mo})p_{mo} + p_{cm} \right) \quad (7)$$

where we exploit  $z_{mh} = \alpha(p_{mo})z_{mo}$  and the composition rule, i.e.,  $p_{mo} = e^{-\hat{\xi}\theta_{wm}} (p_{wo} - p_{wm})$ .

### B. Image Feature for Panoramic Camera

Let  $p_{oi} \in \mathcal{R}^3$  and  $p_{mi} \in \mathcal{R}^3$  be the position vectors of the target object's  $i$ -th feature points ( $i = 1, \dots, n$ ,  $(n \geq 4)$ ) relative to  $\Sigma_o$  and  $\Sigma_m$ , respectively (see Fig. 2). Using a transformation of the coordinates, we have

$$p_{mi} = g_{mo} p_{oi} \quad (8)$$

where  $p_{mi}$  and  $p_{oi}$  should be regarded, with a slight abuse of notation, as  $[p_{mi}^T \ 1]^T$  and  $[p_{oi}^T \ 1]^T$  via the well-known homogeneous coordinate representation in robotics, respectively (see, e.g., [12]).

The hyperbolic 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

$$f_i = \frac{\lambda \alpha(p_{mi})}{2r + \alpha(p_{mi})z_{mi}} \begin{bmatrix} x_{mi} \\ y_{mi} \end{bmatrix} \quad (9)$$

where  $\alpha(p_{mi})$  means that  $p_{mo}$  in Eq. (6) is replaced with  $p_{mi} = [x_{mi} \ y_{mi} \ z_{mi}]^T$ . It is straightforward to extend this model to  $n$  image points by simply stacking the vectors of the image plane coordinate, i.e.,

$$f(g_{mo}) := [f_1^T \ \dots \ f_n^T]^T \in \mathcal{R}^{2n}. \quad (10)$$

and  $p_m := [p_{m1}^T \ \dots \ p_{mn}^T]^T \in \mathcal{R}^{3n}$ . We assume that multiple feature points on a known object are given. Although the problem of extracting the feature points from the target object is interesting in its own right, we will not focus on this problem and merely assume that the image feature are obtained by well-known techniques [13]. From Eq. (8), the image feature  $f$  only depends on the relative rigid body motion  $g_{mo}$ .

### III. BODY VELOCITY FOR PANORAMIC CAMERA

The relative rigid body motion  $g_{mo}$  is discussed in this section, because the image feature  $f$  only depends on  $g_{mo}$ . We recall that visual feedback systems by using a panoramic camera use four coordinate frames and the relative rigid

body motion from  $\Sigma_m$  to  $\Sigma_o$  can be represented by  $g_{mo} = (p_{mo}, e^{\hat{\xi}\theta_{mo}})$  as shown in Fig. 3.

The relative rigid body motion from  $\Sigma_m$  to  $\Sigma_o$  can be led by using the composition rule for rigid body transformations ([12], Chap. 2, pp. 37, eq. (2.24)) as follows:

$$g_{mo} = g_{wm}^{-1} g_{wo}. \quad (11)$$

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 [12]. We define the body velocity of the mirror relative to the world frame  $\Sigma_w$  as  $V_{wm}^b = [v_{wm}^T \ \omega_{wm}^T]^T$ , where  $v_{wm}$  and  $\omega_{wm}$  represent the velocity of the origin and the angular velocity from  $\Sigma_w$  to  $\Sigma_m$ , respectively ([12] Chap. 2, eq. (2.55)).

Differentiating (11) with respect to time, the body velocity of the relative rigid body motion  $g_{mo}$  can be written as follows (See [5]):

$$V_{mo}^b = -\text{Ad}_{(g_{mo}^{-1})} V_{wm}^b + V_{wo}^b. \quad (12)$$

where  $V_{wo}^b$  is the body velocity of the target object relative to  $\Sigma_w$ . Eq. (12) is a standard formula for the relation between the body velocities of three coordinate frames ([12] Chap. 2, pp. 59, Proposition 2.15). Because the camera velocity  $V_{wc}^b$  is adequate as an input rather than the mirror velocity  $V_{wo}^b$  in this framework, we lead the body velocity of the relative rigid body motion with the camera velocity.

The body velocity of the mirror frame relative to  $\Sigma_w$  will be denoted as

$$\begin{aligned} \hat{V}_{wm}^b &= g_{wm}^{-1} \dot{g}_{wm} = g_{wm}^{-1} \dot{g}_{wc} g_{cm} \\ &= g_{wm}^{-1} g_{wc} g_{wc}^{-1} \dot{g}_{wc} g_{cm} = g_{cm}^{-1} \hat{V}_{wc}^b g_{cm}. \end{aligned} \quad (13)$$

From the property concerning the adjoint transformation,  $V_{wm}^b$  can be transformed into

$$V_{wm}^b = \text{Ad}_{(g_{cm}^{-1})} V_{wc}^b. \quad (14)$$

Thus, Eq. (12) can be transformed into

$$V_{mo}^b = -\text{Ad}_{(g_{mo}^{-1})} \text{Ad}_{(g_{cm}^{-1})} V_{wc}^b + V_{wo}^b. \quad (15)$$

This is the body velocity of the relative rigid body motion for the panoramic camera. While  $g_{cm}$  is known information from Eq. (3),  $g_{mo}$  and  $g_{wo}$ , i.e.  $V_{mo}^b$  and  $V_{wo}^b$ , are unknown information in the visual feedback system. Then, the control objective is described as follows.

**Control Objective:** The controlled mobile robot follows the target robot, i.e., the relative rigid body motion  $g_{co}$  is coincided with the desired one  $g_d$ .

Because  $g_{cm}$  is known a priori,  $g_{co}$  can be obtained from  $g_{mo}$  by the composition rule  $g_{co} = g_{cm} g_{mo}$ . Thus, we consider the estimate of  $g_{mo}$  for the above control objective.

#### IV. VISION-BASED ESTIMATION

##### A. Image Jacobian for Panoramic Camera

Since the measurable information is only the image feature  $f$  from the panoramic camera, we consider a visual motion observer in order to estimate the relative rigid body motion  $g_{mo}$  from the image feature  $f$ . Using the body velocity of the relative rigid body motion (15), we choose estimates  $\bar{g}_{mo} = (\bar{p}_{mo}, e^{\hat{\xi}\bar{\theta}_{mo}})$  and  $\bar{V}_{mo}^b$  of the relative rigid body motion and velocity, respectively as

$$\bar{V}_{mo}^b = -\text{Ad}_{(\bar{g}_{mo}^{-1})} \text{Ad}_{(g_{cm}^{-1})} V_{wc}^b + u_e. \quad (16)$$

The new input  $u_e = [v_{ue}^T \ \omega_{ue}^T]^T$  is to be determined in order to drive the estimated values  $\bar{g}_{mo}$  and  $\bar{V}_{mo}^b$  to their actual values.

Similarly to (8) and (9), the estimated image feature  $\bar{f}_i$  ( $i = 1, \dots, n$ ) is defined as

$$\bar{p}_{mi} = \bar{g}_{mo} p_{oi} \quad (17)$$

$$\bar{f}_i = \frac{\lambda \alpha (\bar{p}_{mi})}{2r + \alpha (\bar{p}_{mi}) \bar{z}_{mi}} \begin{bmatrix} \bar{x}_{mi} \\ \bar{y}_{mi} \end{bmatrix} \quad (18)$$

where  $\bar{p}_{mi} := [\bar{x}_{mi} \ \bar{y}_{mi} \ \bar{z}_{mi}]^T$ .  $\bar{f}(\bar{g}_{mo}) := [\bar{f}_1^T \ \dots \ \bar{f}_n^T]^T \in \mathcal{R}^{2n}$  means the  $n$  image points case.

In order to establish the estimation error system, we define the estimation error between the estimated value  $\bar{g}_{mo}$  and the actual relative rigid body motion  $g_{mo}$  as

$$g_{ee} := \bar{g}_{mo}^{-1} g_{mo} \quad (19)$$

in other words,  $p_{ee} = e^{-\hat{\xi}\bar{\theta}_{mo}}(p_{mo} - \bar{p}_{mo})$  and  $e^{\hat{\xi}\theta_{ee}} = e^{-\hat{\xi}\bar{\theta}_{mo}} e^{\hat{\xi}\theta_{mo}}$ . We next define the error vector of the rotation matrix  $e^{\hat{\xi}\theta_{ab}}$  as  $e_R(e^{\hat{\xi}\theta_{ab}}) := \text{sk}(e^{\hat{\xi}\theta_{ab}})^\vee$  where  $\text{sk}(e^{\hat{\xi}\theta_{ab}})$  denotes  $\frac{1}{2}(e^{\hat{\xi}\theta_{ab}} - e^{-\hat{\xi}\theta_{ab}})$ . Using this notation, the vector of the estimation error is given by  $e_e := [p_{ee}^T \ e_R^T(e^{\hat{\xi}\theta_{ee}})]^T$ . From the above, we derive a relation between the actual and the estimated image feature. Suppose the attitude estimation error  $\theta_{ee}$  is small enough that we can let  $e^{\hat{\xi}\theta_{ee}} \simeq I + \text{sk}(e^{\hat{\xi}\theta_{ee}})$ . Then we have the following relation between the actual feature point  $p_{mi}$  and the estimated one  $\bar{p}_{mi}$

$$p_{mi} - \bar{p}_{mi} = e^{\hat{\xi}\bar{\theta}_{mo}} \begin{bmatrix} I & -\hat{p}_{oi} \end{bmatrix} \begin{bmatrix} p_{ee} \\ e_R(e^{\hat{\xi}\theta_{ee}}) \end{bmatrix}. \quad (20)$$

Using a first-order Taylor expansion approximation, the relation between the actual image feature and the estimated one can be expressed as

$$f_i - \bar{f}_i = \begin{bmatrix} \frac{\partial f_i}{\partial x_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} & \frac{\partial f_i}{\partial y_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} \\ \frac{\partial f_i}{\partial z_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} \end{bmatrix} (p_{mi} - \bar{p}_{mi}) \quad (21)$$

where the partial differentiations  $\partial f_i / \partial x_{mi}$ ,  $\partial f_i / \partial y_{mi}$  and  $\partial f_i / \partial z_{mi}$  are represented as Eqs. (22)–(24) at the bottom of the next page, respectively.

Let us define the image feature error as  $f_e := f(g_{mo}) - \bar{f}(\bar{g}_{mo})$ . Hence, the relation between the actual image feature and the estimated one can be given by

$$f_e = J(\bar{g}_{mo}) e_e, \quad (25)$$

where  $J(\bar{g}_{mo}) : SE(3) \rightarrow \mathcal{R}^{2n \times 6}$  is defined as

$$J(\bar{g}_{mo}) := [J_1^T(\bar{g}_{mo}) \ J_2^T(\bar{g}_{mo}) \ \dots \ J_m^T(\bar{g}_{mo})]^T \quad (26)$$

$$J_i(\bar{g}_{mo}) := \begin{bmatrix} \frac{\partial f_i}{\partial x_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} & \frac{\partial f_i}{\partial y_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} \\ \frac{\partial f_i}{\partial z_{mi}} \Big|_{p_{mi}=\bar{p}_{mi}} \end{bmatrix} \times e^{\hat{\xi}\bar{\theta}_{mo}} \begin{bmatrix} I & -\hat{p}_{oi} \end{bmatrix}. \quad (i = 1, \dots, n) \quad (27)$$

We assume that the matrix  $J(\bar{g}_{mo})$  is full column rank for all  $\bar{g}_{mo} \in SE(3)$ . Then, the relative rigid body motion can be uniquely defined by the image feature vector. Because this may not hold in some cases when  $n = 3$ , it is known that

$n \geq 4$  is desirable for the full column rank of the image Jacobian.

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

$$e_e = J^\dagger(\bar{g}_{mo})f_e \quad (28)$$

where  $\dagger$  denotes the pseudo-inverse. Therefore the estimation error  $e_e$  can be exploited in the 3D visual feedback control law using image feature  $f$  obtained from the panoramic camera.

*Remark 1:* If we select one and the imaginary unit as  $a$  and  $b$  numerically, i.e.,  $a = 1, b = i$ , then the image Jacobian for the panoramic camera (27) equals to the pinhole's one [5] as follows:  $\frac{\partial f_i}{\partial x_{mi}} = \frac{\lambda}{z_{mi}} [1 \ 0]^T$ ,  $\frac{\partial f_i}{\partial y_{mi}} = \frac{\lambda}{z_{mi}} [0 \ 1]^T$ ,  $\frac{\partial f_i}{\partial z_{mi}} = -\frac{\lambda}{z_{mi}^2} [x_{mi} \ y_{mi}]^T$ . Thus our previous work [5] can be regarded as a special case of this study, although the pinhole camera has different applications from the panoramic one in the practical view.

### B. Passivity of Estimation Error System

In the same way as Eq. (12), the estimation error system can be represented by

$$V_{ee}^b = -\text{Ad}_{(g_{ee}^{-1})} u_e + V_{wo}^b. \quad (29)$$

Then, we have the following lemma relating the input  $u_e$  to the vector form of the estimation error  $e_e$ .

*Lemma 1:* If  $V_{wo}^b = 0$ , then the following inequality holds for the estimation error system (29).

$$\int_0^T u_e^T(-e_e)dt \geq -\beta_e \quad (30)$$

where  $\beta_e$  is a positive scalar.

*Proof:* Consider the positive definite function

$$V_e = \frac{1}{2} \|p_{ee}\|^2 + \phi(e^{\hat{\xi}\theta_{ee}}) \quad (31)$$

where  $\phi(e^{\hat{\xi}\theta_{ab}}) := \frac{1}{2} \text{tr}(I - e^{\hat{\xi}\theta_{ab}})$  is the error function of the rotation matrix [14]. Evaluating the time derivative of  $V_e$

along the trajectories of (29) gives us

$$\begin{aligned} \dot{V}_e &= p_{ee}^T e^{\hat{\xi}\theta_{ee}} e^{-\hat{\xi}\theta_{ee}} \dot{p}_{ee} + e_R^T (e^{\hat{\xi}\theta_{ee}}) e^{\hat{\xi}\theta_{ee}} \omega_{ee} \\ &= e_e^T \text{Ad}_{(e^{\hat{\xi}\theta_{ee}})} V_{ee}^b = u_e^T(-e_e). \end{aligned} \quad (32)$$

Integrating (32) from 0 to  $T$  yields

$$\int_0^T u_e^T(-e_e)d\tau = \int_0^T \dot{V}_e d\tau \geq -V(0) \geq -\beta_e \quad (33)$$

where  $\beta_e$  is the positive scalar which only depends on the initial state of  $g_{ee}$ .  $\blacksquare$

*Remark 2:* Let us consider the vector form of the estimation error  $-e_e$  as its output. Then, Lemma 1 says that the estimation error system (29) is *passive* from the input  $u_e$  to the output  $-e_e$ . In fact, the body velocity of the relative rigid body motion (15) has passivity, the estimation error system preserves its passivity.

### C. Visual Motion Observer

Based on the above passivity property of the estimation error system, we consider the following control law.

$$u_e = K_e e_e \quad (34)$$

where  $K_e := \text{diag}\{k_{e1}, \dots, k_{e6}\}$  is the positive gain matrix of  $x, y$  and  $z$  axes of the translation and the rotation for the estimation error.

*Theorem 1:* If  $V_{wo}^b = 0$ , then the equilibrium point  $e_e = 0$  for the closed-loop system (29) and (34) is asymptotic stable.

*Proof:* Theorem 1 can be easily proved by Lemma 1, hence the proof is omitted.  $\blacksquare$

Fig. 4 shows a block diagram of the visual motion observer with a panoramic camera. By the proposed visual motion observer, the unmeasurable motion  $g_{co}$  will be exploited as the part of control law. Our proposed visual motion observer is composed just as Luenberger observer for linear systems.

*Remark 3:* The estimation error vector is configured by available information (i.e., the measurement and the estimate) though it is defined by unavailable one. This is one of the main contributions of this paper.

$$\frac{\partial f_i}{\partial x_{mi}} = \frac{2r\lambda\alpha_x(p_{mi})}{(2r + \alpha(p_{mi})z_{mi})^2} \begin{bmatrix} x_{mi} \\ y_{mi} \end{bmatrix} + \frac{\lambda\alpha(p_{mi})}{2r + \alpha(p_{mi})z_{mi}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (22)$$

$$\frac{\partial f_i}{\partial y_{mi}} = \frac{2r\lambda\alpha_y(p_{mi})}{(2r + \alpha(p_{mi})z_{mi})^2} \begin{bmatrix} x_{mi} \\ y_{mi} \end{bmatrix} + \frac{\lambda\alpha(p_{mi})}{2r + \alpha(p_{mi})z_{mi}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (23)$$

$$\frac{\partial f_i}{\partial z_{mi}} = \frac{2r\lambda\alpha_z(p_{mi})}{(2r + \alpha(p_{mi})z_{mi})^2} \begin{bmatrix} x_{mi} \\ y_{mi} \end{bmatrix} - \frac{\lambda\alpha^2(p_{mi})}{(2r + \alpha(p_{mi})z_{mi})^2} \begin{bmatrix} x_{mi} \\ y_{mi} \end{bmatrix} \quad (24)$$

where

$$\alpha_x(p_{mi}) = \frac{\partial\alpha(p_{mi})}{\partial x_{mi}} = \frac{-2a^2b^2rx_{mi}z_{mi}\|p_{mi}\| \pm ab^2x_{mi}(r^2z_{mi}^2 + a^2\|p_{mi}\|^2)}{(b^2z_{mi}^2 - a^2x_{mi}^2 - a^2y_{mi}^2)^2\|p_{mi}\|}$$

$$\alpha_y(p_{mi}) = \frac{\partial\alpha(p_{mi})}{\partial y_{mi}} = \frac{-2a^2b^2ry_{mi}z_{mi}\|p_{mi}\| \pm ab^2y_{mi}(r^2z_{mi}^2 + a^2\|p_{mi}\|^2)}{(b^2z_{mi}^2 - a^2x_{mi}^2 - a^2y_{mi}^2)^2\|p_{mi}\|}$$

$$\alpha_z(p_{mi}) = \frac{\partial\alpha(p_{mi})}{\partial z_{mi}} = \frac{b^2r(b^2z_{mi}^2 + a^2x_{mi}^2 + a^2y_{mi}^2)\|p_{mi}\| \mp ab^2z_{mi}(r^2(x_{mi}^2 + y_{mi}^2) + b^2\|p_{mi}\|^2)}{(b^2z_{mi}^2 - a^2x_{mi}^2 - a^2y_{mi}^2)^2\|p_{mi}\|}$$

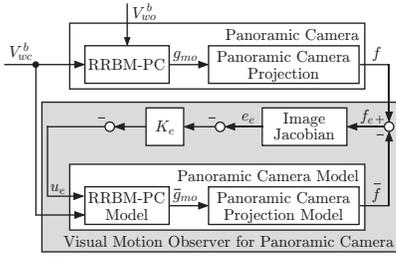


Fig. 4. Block diagram of visual motion observer.

## V. VISUAL MOTION OBSERVER-BASED POSE CONTROL

### A. Pose Control Error System

Let us consider the dual of the estimation error system, which we call the pose control error system, in order to achieve the control objective. First, we define the pose control error as follows:

$$g_{ec} = g_d^{-1} g_{co}, \quad (35)$$

which represents the error between the relative rigid body motion  $g_{co}$  and the reference one  $g_d$ . It should be remarked that  $g_{co}$  can be calculated by using the estimated relative rigid body motion  $\tilde{g}_{mo}$ , the estimation error vector  $e_e = [p_{ee}^T \ e_R^T(e^{\hat{\xi}\theta_{ee}})]$  and the known mirror parameter  $g_{cm}$  equivalently, although  $g_{co}$  can't be measured directly (see [6] for more details). Using the notation  $e_R(e^{\hat{\xi}\theta})$ , the vector of the pose control error is defined as  $e_c := [p_{ec}^T \ e_R^T(e^{\hat{\xi}\theta_{ec}})]^T$ .

The reference of the relative rigid body motion  $g_d$  is assumed to be constant in this paper, i.e.,  $\dot{g}_d = 0$  and hence  $V_{ec}^b = V_{co}^b$ . Thus, the pose control error system can be represented as

$$V_{ec}^b = -\text{Ad}_{(g_{ec}^{-1})} \left( \text{Ad}_{(g_d^{-1})} V_{wc}^b \right) + V_{wo}^b. \quad (36)$$

This is dual to the estimation error system. Similar to the estimation error system, the pose control error system also preserves the passivity property.

### B. Passivity of Visual Motion Error System

Combining the estimation error system (29) and the pose control one (36), we construct the visual motion observer-based pose control error system (we call the visual motion error system) as follows:

$$\begin{bmatrix} V_{ec}^b \\ V_{ee}^b \end{bmatrix} = \begin{bmatrix} -\text{Ad}_{(g_{ec}^{-1})} & 0 \\ 0 & -\text{Ad}_{(g_{ee}^{-1})} \end{bmatrix} u + \begin{bmatrix} I \\ I \end{bmatrix} V_{wo}^b \quad (37)$$

where  $u := [(\text{Ad}_{(g_d^{-1})} V_{wc}^b)^T \ u_e^T]^T$ . Let us define the error vector of the visual motion error system as  $x := [e_c^T \ e_e^T]^T$  which consists of the pose control error vector  $e_c$  and the estimation error vector  $e_e$ . It should be noted that if the vectors of the pose control error and the estimation one are equal to zero, then the actual relative rigid body motion  $g_{co}$  tends to the reference one  $g_d$  when  $x \rightarrow 0$ .

Next, we show an important relation between the input and the output of the visual motion error system.

*Lemma 2:* If  $V_{wo}^b = 0$ , then the visual motion error system (37) satisfies

$$\int_0^T u^T(-x)dt \geq -\beta, \quad \forall T > 0 \quad (38)$$

where  $\beta$  is a positive scalar.

*Proof:* Lemma 2 can be proved by using the positive definite function  $V = \frac{1}{2} \|p_{ec}\|^2 + \phi(e^{\hat{\xi}\theta_{ec}}) + \frac{1}{2} \|p_{ee}\|^2 + \phi(e^{\hat{\xi}\theta_{ee}})$ , hence the proof is omitted. ■

*Remark 4:* Let us take  $u$  as the input and  $-x$  as its output. Thus, Lemma 2 implies that the visual motion error system (37) is *passive* from the input  $u$  to the output  $-x$ .

### C. Visual Motion Pose Control and Stability Analysis

Based on the above passivity property of the visual motion error system, we consider the following control law.

$$u = Kx, \quad K := \begin{bmatrix} K_c & 0 \\ 0 & K_e \end{bmatrix} \quad (39)$$

where  $K_c := \text{diag}\{k_{c1}, \dots, k_{c6}\}$  is the positive gain matrix of  $x$ ,  $y$  and  $z$  axes of the translation and the rotation for the pose control error.

*Theorem 2:* If  $V_{wo}^b = 0$ , then the equilibrium point  $x = 0$  for the closed-loop system (37) and (39) is asymptotic stable.

*Proof:* Theorem 2 can be easily proved by Lemma 2, hence the proof is omitted. ■

Theorem 2 shows Lyapunov stability for the closed-loop system. If the camera velocity is decided directly, the control objective is achieved by using the proposed control law (39). Although the dead angle problem [9] is not considered explicitly in this framework, our proposed method will overcome this problem by dealing with a collision avoidance as in [11].

### D. $L_2$ -Gain Performance Analysis

In this subsection, we utilize  $L_2$ -gain performance analysis to evaluate the tracking performance of the control scheme in the presence of a moving target robot. The motion of the target robot is regarded as an external disturbance.

In order to derive a simple and practical gain condition, we redefine  $K_e = k_e I$  where  $k_e$  is a positive scalar.

*Theorem 3:* Given a positive scalar  $\gamma$ , assume

$$k_{c,\min} > \frac{\gamma^2(2k_e - 1) + 2(k_e - 1)}{2\{\gamma^2(2k_e - 1) - 1\}} \quad (40)$$

$$k_e > \frac{\gamma^2 + 1}{2\gamma^2} \quad (41)$$

where  $k_{c,\min}$  means the minimum value in  $K_c$ . Then the closed-loop system (37) and (39) has  $L_2$ -gain  $\leq \gamma$ .

*Proof:* The proof is omitted due to its similarity to the proof of Theorem 3 in [5]. ■

The estimation gain does not have to be restricted to a scalar, although it cause slightly complicated gain conditions compared with Eqs. (40) and (41). for the reason that Schur complement can not be used. In this framework,  $\gamma$  can be considered as an indicator of the tracking performance.

## VI. SIMULATION AND EXPERIMENTAL RESULTS

Both the proposed estimation and control methods are confirmed by the simulation, while the experiment is carried out by using the proposed visual motion observer only.

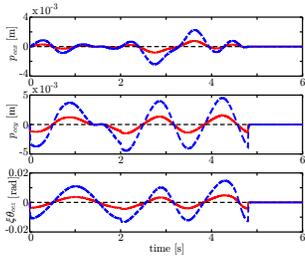


Fig. 5. Estimation error  $e_e$  (Gain A: Dashed, Gain B: Solid).

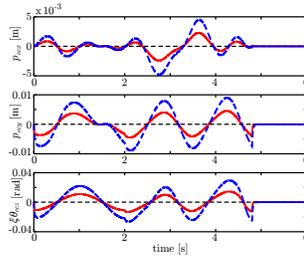


Fig. 6. Pose control error  $e_c$  (Gain A: Dashed, Gain B: Solid).



Fig. 7. Panoramic camera.



Fig. 8. Target mobile robot.

### A. Simulation Results

In this subsection, we present simulation results for stability and  $L_2$ -gain performance analysis in the case of a moving target robot. The target robot has four feature points and moves by  $t = 4.8$  [s]. The gains for the control law  $u$  (39) were empirically selected as follows:

Gain A)  $\gamma = 0.123$ ,  $K_c = 50I$ ,  $k_e = 100$

Gain B)  $\gamma = 0.082$ ,  $K_c = 100I$ ,  $k_e = 300$ .

The simulation results are presented in Figs. 5 and 6 which illustrate the estimation error  $e_e$  and the pose control one  $e_c$ , respectively. In these figures, we focus on the errors of the translations of  $x$  and  $y$  and the rotation of  $z$ . In Figs. 5 and 6, the dashed line and the solid line are the errors in the case of  $\gamma = 0.123$  and  $\gamma = 0.082$ , respectively.

In the case of the static target robot, i.e., after  $t = 4.8$  [s], all errors in Figs. 5 and 6 tend to zero. Therefore, asymptotic stability can be confirmed through the simulation. In the presence of the moving target robot as disturbances by  $t = 4.8$  [s], the tracking performance is improved for the smaller values of  $\gamma$  from Figs. 5 and 6. Thus the simulation results show that  $L_2$ -gain is adequate for the performance measure of the visual motion observer-based pose control.

### B. Experimental Results

In this subsection, we describe experimental results with respect to the proposed visual motion observer with a panoramic camera for a static target object as shown in Figs. 7 and 8. A panoramic camera consists of a MTV-7310 camera and a hyperbolic mirror. The video signals are acquired by a frame grabber board PICOLO DILLIGENT (Euresys) and an image processing software HALCON (MVTech Software GmbH).

The experiment was carried out with an appropriate initial estimation error. The experimental results are presented in Figs. 9 and 10. In Fig. 9, the dashed lines and the solid ones mean the image features and the estimated ones with respect to the  $x$ -axis, i.e.,  $f_{xi}$  and  $\hat{f}_{xi}$  ( $i = 1, \dots, 4$ ), respectively. The estimated image features coincide with the actual ones after  $t = 1$  [s]. From Fig. 10 which illustrates the estimation error of the translations of  $x$  and  $y$  and the rotation of  $z$ ,

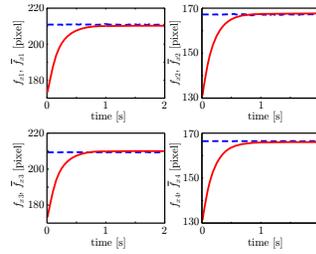


Fig. 9. Image features and estimated ones w.r.t. the  $x$ -axis.

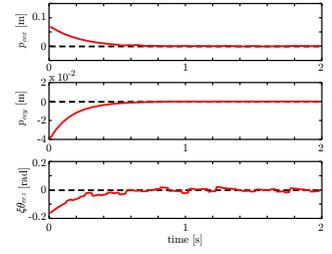


Fig. 10. Estimation error  $e_e$  for the static target object with  $k_e = 5$ .

we can confirm that the estimation error  $e_e$  tends to zero by using the proposed visual motion observer.

## VII. CONCLUSIONS

This paper considers the vision-based estimation and control with a panoramic camera based on the passivity. The main contribution of this paper is to show that the estimation error system with a panoramic camera has the passivity which allows us to prove stability in the sense of Lyapunov. The visual motion error system which consists of the estimation error system and the pose control one preserves the passivity. Our previous work [5] can be regarded as a special case of this study.

## REFERENCES

- [1] F. Chaumette and S. A. Hutchinson, "Visual Servoing and Visual Tracking," In: B. Siciliano and O. Khatib (Eds), *Springer Handbook of Robotics*, Springer-Verlag, pp. 563–583, 2008.
- [2] V. Lippiello, B. Siciliano and L. Villani, "Position-Based Visual Servoing in Industrial Multirobot Cells Using a Hybrid Camera Configuration," *IEEE Trans. on Robotics*, Vol. 23, No. 1, pp. 73–86, 2007.
- [3] N. R. Gans and S. A. Hutchinson, "Stable Visual Servoing Through Hybrid Switched-System Control," *IEEE Trans. on Robotics*, Vol. 23, No. 3, pp. 530–540, 2007.
- [4] G. Hu, W. MacKunis, N. Gans, W. E. Dixon, J. Chen, A. Behal, and D. Dawson, "Homography-Based Visual Servo Control With Imperfect Camera Calibration," *IEEE Trans. on Automatic Control*, Vol. 54, No. 6, pp. 1318–1324, 2009.
- [5] 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 Trans. on Control Systems Technology*, Vol. 15, No. 1, pp. 40–52, 2007.
- [6] T. Murao, H. Kawai and M. Fujita "Predictive Visual Feedback Control with Eye-in-to-Hand Configuration via Stabilizing Receding Horizon Approach," *Proc. of the 17th IFAC World Congress on Automatic Control*, pp. 5341–5346, 2008.
- [7] C. Geyer and K. Daniilidis, "A Unifying Theory for Central Panoramic Systems and Practical Implications," In: D. Vernon (Ed), *Computer Vision - ECCV 2000*, Springer-Verlag, pp. 445–461, 2000.
- [8] G. L. Mariottini and D. Prattichizzo, "EGT for Multiple View Geometry and Visual Servoing," *IEEE Robotics & Automation Magazine*, Vol. 12, No. 4, pp. 26–39, 2005.
- [9] R. T. Fomena and F. Chaumette, "Improvements on Visual Servoing From Spherical Targets Using a Spherical Projection Model," *IEEE Trans. on Robotics*, Vol. 25, No. 4, pp. 874–886, 2009.
- [10] R. Vidal, O. Shakernia and S. Sastry, "Following the flock," *IEEE Robotics & Automation Magazine*, Vol. 11, No. 4, pp. 14–20, 2004.
- [11] M. Fujita, T. Hatanaka, N. Kobayashi, T. Ibuki and M. W. Spong, "Visual Motion Observer-based Pose Synchronization: A Passivity Approach," *Proc. of the 48th IEEE Conf. on Decision and Control*, pp. 2402–2407, 2009.
- [12] R. Murray, Z. Li and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, 1994.
- [13] D. Forsyth and J. Ponce, *Computer Vision - A Modern Approach*, Prentice-Hall, 2003.
- [14] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems*, Springer-Verlag, 2004.