

# Robust 2.5D Visual Servoing for Robot Manipulators\*

Y. Fang,<sup>1</sup> W. E. Dixon,<sup>2</sup> D. M. Dawson,<sup>1</sup> and J. Chen<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634-0915

<sup>2</sup>Eng. Science and Tech. Division-Robotics, Oak Ridge National Lab., P.O. Box 2008, Oak Ridge, TN 37831-6305

“The submitted manuscript has been authored by a contractor of the U.S. Government under contract DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.”

To appear at the IEEE American Control Conference, June 4-6, 2003, Denver, CO

Keywords: 2.5D Visual Servoing, Robust Control, and Nonlinear Lyapunov-Based Control

E-mail: dixonwe@ornl.gov, Telephone: (865)574-9025

---

\* This research was supported in part U.S. DOE Office of Biological and Environmental Research (OBER), Environmental Management Sciences Program (EMSP) project ID No. 82797 at ORNL, a subcontract to ORNL by the Florida Department of Citrus through the University of Florida, and by U.S. NSF Grant DMI-9457967, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

# Robust 2.5D Visual Servoing for Robot Manipulators\*

Y. Fang<sup>†</sup>, W. E. Dixon<sup>‡</sup>, D. M. Dawson<sup>†</sup> and J. Chen<sup>†</sup>

<sup>†</sup>Department of Electrical & Computer Engineering, Clemson University, Clemson, SC 29634-0915

<sup>‡</sup>Eng. Science and Tech. Div. - Robotics, Oak Ridge National Lab., P.O. Box 2008, Oak Ridge, TN 37831-6305

email: yfang, ddawson, jianc@ces.clemson.edu; dixonwe@ornl.gov

## Abstract

In this paper, the 3-dimensional (3D) position and orientation of a camera held by the end-effector of a robot manipulator is regulated to a constant desired position and orientation despite (i) the lack of depth information of the actual or desired camera position from a target, (ii) the lack of a geometric model of the target object, and (iii) uncertainty regarding both the angle and axis of rotation of the camera with respect to the robot end-effector (i.e., the orientation extrinsic camera parameters). By fusing 2D image-space and projected 3D task-space information (i.e., 2.5D visual servoing), a robust controller is developed that ensures exponential regulation of the position and orientation of the camera. The stability of the controller is proven through a Lyapunov-based analysis.

## 1 Introduction

Motivated by the significant impact that may be realized by enabling robotic systems with the ability to perform tasks based on a sense of perception, a myriad of research has been directed at vision related issues. Some key issues that have limited the robustness of vision-based robotic control are camera calibration and the fact that the image space is a 2-dimensional (2D) projection of the 3D task-space. Specifically, based on the fact that the camera output is in the image-space and the robot control is computed in terms of the task-space (joint space), an optic model is often employed to relate image-space data to the task-space. To relate the image-space to the task-space, both intrinsic and extrinsic parameters<sup>1</sup> of the optic model are required. Motivated by the desire to incorporate robustness to these parameters, several adaptive and robust controllers have been designed (e.g., see [7, 11, 18]). Unfortunately, much of the previous work either constrains the visual servoing problem to a planar case so that the effects of the unknown depth information can be held constant or assumes that the depth can be measured via an additional sensor (e.g., ultrasonic sensors, laser-based sensors, additional cameras) or estimated from multiple cameras.

\*This research was supported in part U.S. DOE Office of Biological and Environmental Research (OBER), Environmental Management Sciences Program (EMSP) project ID No. 82797 at ORNL, a subcontract to ORNL by the Florida Department of Citrus through the University of Florida, and by U.S. NSF Grant DMI-9457967, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

<sup>1</sup>The camera calibration parameters are composed of the so-called intrinsic parameters (i.e., image center, camera scale factors, and camera magnification factor) and extrinsic parameters (i.e., camera position and orientation).

Researchers have also been motivated by the desire to compensate for the lack of depth information from the 2D image data (without requiring additional cameras or additional sensors). To this end, several researchers have recently developed partitioned approaches that exploit a combination of 3D task-space information and 2D image-space information. For example, in the series of papers by Malis and Chaumette (e.g., [1, 2, 14, 15]) various kinematic control strategies (coined 2.5D visual servo controllers) exploit the fact that the interaction between translation and rotation components can be decoupled through a homography. Specifically, information from the task-space (obtained through a projective Euclidean reconstruction from the image data) is utilized to regulate the rotation error system, while information from the 2D image-space is utilized to control the translation error system. In [6], Deguchi utilizes a homography relationship and an epipolar condition to decouple the rotation and translation components and then illustrates how two types of visual controllers can be developed from the decoupled information. More recently, Corke and Hutchinson [5] also developed a hybrid image-based visual servoing scheme that decouples rotation and translation components about the  $z$ -axis from the remaining degrees of freedom. One drawback of some of the previous controllers is the claim that a constant estimate of the aforementioned depth information can be utilized in lieu of the exact value (although, no stability analysis is provided to support this claim). That is, as stated in [15], an off-line learning stage is required to estimate the distance of the desired camera position to the reference plane. Motivated by the desire to actively compensate for the aforementioned depth information, [3] developed an adaptive kinematic controller to ensure uniformly ultimately bounded (UUB) set-point regulation of the image space errors while compensating for the unknown depth information, provided conditions on the translational velocity and the bounds on uncertain depth parameters are satisfied. In [4], Conticelli et al. proposed a 3D depth estimation procedure that exploits a prediction error provided a positive definite condition on the interaction matrix is satisfied. In [17], Taylor et al. developed a kinematic controller that utilizes a constant, best-guess estimate of the calibration parameters to achieve local set-point regulation; although, several conditions on the rotation and calibration matrix are required. In [8], Fang et al. recently developed a 2.5D visual servo controller to asymptotically regulate a manipulator end-effector by exploiting Lyapunov-based techniques to develop an adaptive update law that compensated for an unknown depth parameter. Built on the results of [8], Fang et al. in [9] designed a homography-based visual servo controller that asymptotically regulates the position of a wheeled mobile robot despite non-

holonomic constraints and parametric uncertainty in the depth parameter. Although the results in [8] and [9] were achieved despite unknown depth information, the intrinsic and extrinsic camera parameters were required to be known; hence, motivation exists to develop controllers that can compensate for the intrinsic and extrinsic camera parameters. Recently in [13], Malis and Chaumette proposed a 2.5D visual servo controller to address the unknown calibration parameters. Specifically, the approach in [13] is to use a constant, “best-guess” estimate of the intrinsic and extrinsic calibration parameters and then discuss the resulting stability of the controller.

In this paper, feature points extracted from images taken from the desired and current camera position and orientation are related through a homography, so that the 3D position and orientation of a single camera held by the end-effector of a robot manipulator are exponentially regulated to the constant, desired position and orientation. The result is facilitated by decomposing the homography into translation and rotation components and then combining 2D image-space and projected 3D task-space information (i.e., 2.5D visual servoing). As described in [15], the advantages of the 2.5D visual servoing strategy are that a 3D model of the target is not required and the target remains in the camera’s field-of-view. The exponential regulation result is achieved despite (i) the lack of depth information of the actual or desired camera position from a target, (ii) the lack of a geometric model of the target object, and (iii) parametric uncertainty regarding both the angle and axis of rotation of the camera with respect to the robot end-effector. That is, in contrast to the asymptotic results given in [8] and [9] that require calibration of the intrinsic and extrinsic camera parameters, the stability result in this paper is robust to the extrinsic calibration parameters related to the orientation of the camera with respect to the robot end-effector. In contrast to the strategy provided in [13], the controller in this paper compensates for the unknown extrinsic parameters through nonlinear feedback rather than by using a “best-guess” estimate; however, this paper assumes that the intrinsic camera parameters are exactly known.

## 2 Model Development

### 2.1 Camera Model

Consider two orthogonal coordinate systems, denoted by  $\mathcal{F}$  and  $\mathcal{F}^*$ , where  $\mathcal{F}$  is attached to a camera that is held by the robot end-effector, and  $\mathcal{F}^*$  is a fixed coordinate system that represents the constant, desired position and orientation of  $\mathcal{F}$ . Also consider a reference plane  $\pi$  that is defined by four<sup>2</sup> target points  $O_i \forall i = 1, 2, 3, 4$ . The actual and desired 3D coordinates of  $O_i$  expressed in terms of  $\mathcal{F}$  and  $\mathcal{F}^*$  are denoted by  $x_i(t), y_i(t), z_i(t) \in \mathbb{R}$  and  $x_i^*, y_i^*, z_i^* \in \mathbb{R}$ , respectively, and are defined as elements of  $\bar{m}_i(t), \bar{m}_i^* \in \mathbb{R}^3$  as follows

$$\bar{m}_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^T \quad (1)$$

$$\bar{m}_i^* = \begin{bmatrix} x_i^* & y_i^* & z_i^* \end{bmatrix}^T. \quad (2)$$

<sup>2</sup>In general, only 3 points are required to define a plane, however, in the subsequent analysis, 4 target points located on the plane  $\pi$  are required. If the points are not coplanar [15], then at least 8 pairs of points are required to estimate the homography (e.g., using the algorithm presented in [14]).

The task-space coordinates given in (1) and (2) can be expressed as normalized coordinates, denoted by  $m_i(t), m_i^* \in \mathbb{R}^3$ , as follows

$$m_i = \frac{\bar{m}_i}{z_i} = \begin{bmatrix} \frac{x_i}{z_i} & \frac{y_i}{z_i} & 1 \end{bmatrix}^T \quad (3)$$

$$m_i^* = \frac{\bar{m}_i^*}{z_i^*} = \begin{bmatrix} \frac{x_i^*}{z_i^*} & \frac{y_i^*}{z_i^*} & 1 \end{bmatrix}^T \quad (4)$$

where the standard assumption is made that  $z_i(t), z_i^*$  are positive. The normalized task-space coordinates of each target point can also be expressed in terms of pixel coordinates, denoted by  $u_i(t), v_i(t) \in \mathbb{R}$ , through the following global invertible transformation

$$p_i = Am_i \quad (5)$$

where  $p_i(t) \in \mathbb{R}^3$  is defined as follows

$$p_i = \begin{bmatrix} u_i & v_i & 1 \end{bmatrix}^T \quad (6)$$

and  $A \in \mathbb{R}^{3 \times 3}$  is a known, constant, and invertible intrinsic camera calibration matrix. The desired normalized task-space coordinates given in (4) can also be expressed in terms of pixel coordinates, denoted by  $u_i^*, v_i^* \in \mathbb{R}$ , through the following relationship

$$p_i^* = Am_i^*. \quad (7)$$

where  $p_i^* \in \mathbb{R}^3$  is defined as follows

$$p_i^* = \begin{bmatrix} u_i^* & v_i^* & 1 \end{bmatrix}^T. \quad (8)$$

Based on the fact that the matrix  $A$  is known and invertible, the relationships given in (5) and (7) can be used to compute  $m_i(t)$  and  $m_i^*$  of (3) and (4), respectively.

Motivated by the desire to further relate the 2D image-space information to the 3D task-space information, a projective homography, denoted by  $G(t) \in \mathbb{R}^{3 \times 3}$ , can be utilized to relate the image points  $p_i(t)$  of (5) to the image points  $p_i^*$  of (7) in the following manner [10]

$$p_i = \alpha_i G p_i^*. \quad (9)$$

In (9),  $\alpha_i(t) \in \mathbb{R}$  denotes an unknown scaling factor defined as follows

$$\alpha_i = \frac{z_i^*}{z_i} \quad (10)$$

where  $z_i(t)$  and  $z_i^*$  were defined in (1) and (2), respectively. After utilizing (5), (7), and (9), the relationship given in (9) can be rewritten in terms of the normalized task-space coordinates as follows

$$m_i = \alpha_i H m_i^* \quad (11)$$

where  $H(t) \in \mathbb{R}^{3 \times 3}$  denotes an Euclidean homography that is defined as follows

$$H = A^{-1}GA. \quad (12)$$

By utilizing various techniques (e.g., see [10, 19]),  $H(t)$  can now be decomposed into rotation and translation components as follows

$$H = R + x_h n^{*T}. \quad (13)$$

In (13),  $R(t) \in \mathbb{R}^{3 \times 3}$  denotes the measurable rotation from the actual task-space coordinates to the desired task-space coordinates,  $n^* \in \mathbb{R}^3$  denotes the constant unit normal from  $\mathcal{F}^*$  to  $\pi$ , and  $x_h(t) \in \mathbb{R}^3$  denotes the scaled translation vector from  $\mathcal{F}$  to  $\mathcal{F}^*$ . The actual translation from  $\mathcal{F}$  to  $\mathcal{F}^*$  denoted

by  $x_f(t) \in \mathbb{R}^3$  is unmeasurable; however, it can be expressed in terms of the known scaled translation vector  $x_h(t)$  as follows (see Figure 1)

$$x_f = x_h d^* \quad (14)$$

where  $d^* \in \mathbb{R}$  denotes the constant, unknown distance from  $\mathcal{F}^*$  to  $\pi$  along  $n^*$ .

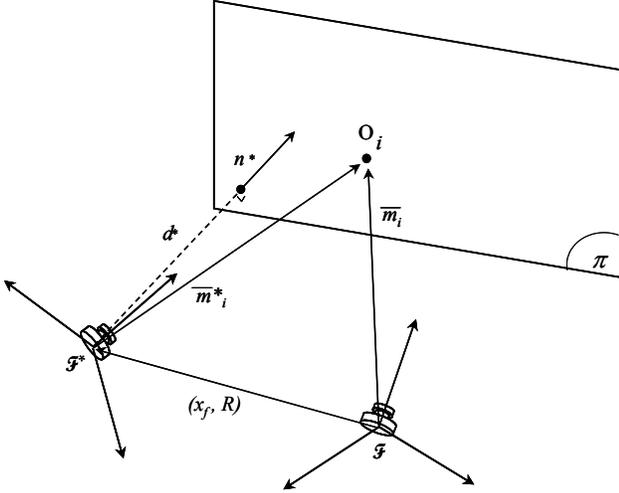


Figure 1: Coordinate frame relationships.

**Remark 1** To compute the homography introduced in (12),  $G(t)$  must first be computed. After normalizing  $G(t)$  with respect to the  $g_{33}(t)$  component, the expression given in (9) can be written as follows

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \frac{\alpha_i}{g_{33}} \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & 1 \end{bmatrix} \begin{bmatrix} u_i^* \\ v_i^* \\ 1 \end{bmatrix} \quad (15)$$

where  $g_{ij}(t)$  denote the components of  $G(t) \forall i, j = 1, 2, 3$ , (since (15) is defined up to a scalar, at least one component of  $G(t)$  is nonzero and without loss of generality we assume that  $g_{33}(t) \neq 0$ ). Since each pair of target points provide the 3 equations given in (15), 4 coplanar points must be selected such that 12 linearly independent equations can be generated from (15) to solve for the 12 unknowns (i.e., 4 unknowns from  $\frac{\alpha_i}{g_{33}} \forall i = 1, 2, 3, 4$  and the 8 unknowns from the normalized  $G(t)$ ). From the system of 12 equations and 12 unknowns, the 4 unknowns  $\frac{\alpha_i}{g_{33}} \forall i = 1, 2, 3, 4$  can be factored out leaving 8 equations with 8 unknowns. The set of set of linear equations used to determine the unknown components of the normalized  $G(t)$  matrix can be written in the following form

$$\Lambda \xi_1 = \xi_2 \quad (16)$$

where  $\Lambda(u_i, v_i, u_i^*, v_i^*) \in \mathbb{R}^{8 \times 8}$  is a known matrix,  $\xi_1 \in \mathbb{R}^8$  denotes the unknown vector composed of the normalized components of  $G(t)$ , and  $\xi_2(u_i, v_i) \in \mathbb{R}^8$  denotes a measurable vector. To determine the unknown parameters from (16) there are two issues that must be addressed. The first issue is ambiguity in the image. Specifically, if the four points are selected so that the rotation of the image may be ambiguous (e.g., a square or circle) then some additional image feature must be used to

remove the ambiguity (e.g., color of the image points) or the solution to (16) will be incorrect. Another issue that must be taken into consideration is that the points must be selected to ensure that  $\Lambda(u_i, v_i, u_i^*, v_i^*)$  is invertible. From the first relationship in (9), a set of 12 linear equations given by the 4 target point pairs  $(p_i^*, p_i(t))$  with 3 equations per target pair can be used to determine the projective homography up to a scalar multiple (i.e., the product  $\alpha_i(t)G(t)$  can be determined). Various techniques can then be used (e.g., see [10, 19]) to decompose the Euclidean homography, to obtain  $\alpha_i(t)$ ,  $G(t)$ ,  $H(t)$ , and the rotation and translation signals  $R(t)$ ,  $x_h(t)$ , and  $n^*$ . Hence,  $R(t)$ ,  $x_h(t)$ ,  $n^*$ , and the depth ratio  $\alpha_i(t)$  are all known signals that can be used for control synthesis.

## 2.2 Control Objective

The control objective of this paper is to regulate the mismatch between the actual and desired 3D task-space camera position to zero. Specifically, the control objective is to regulate the rotation mismatch between  $\mathcal{F}$  and  $\mathcal{F}^*$  (i.e.,  $R(t)$  given in (13)) to the identity matrix, while regulating the translation mismatch between  $\mathcal{F}$  and  $\mathcal{F}^*$  to zero. To this end, we define a rotation error-like signal  $e_\omega(t) \in \mathbb{R}^3$  as follows [15]

$$e_\omega = u\theta \quad (17)$$

where  $u(t) \in \mathbb{R}^3$  represents a unit rotation axis and  $\theta(t) \in \mathbb{R}$  denotes the rotation about  $u(t)$  that is assumed to be confined to the following region

$$-\pi < \theta(t) < \pi. \quad (18)$$

The parameterization  $u(t)\theta(t)$  is related to the rotation matrix  $R(t)$  by the following expression

$$R = I_3 + \sin \theta [u]_\times + 2 \sin^2 \frac{\theta}{2} [u]_\times^2 \quad (19)$$

where the notation  $I_i$  denotes an  $i \times i$  identity matrix, and the notation  $[u]_\times$  denotes the following skew-symmetric matrix

$$[u]_\times \triangleq \begin{bmatrix} 0 & -\zeta_3 & \zeta_2 \\ \zeta_3 & 0 & -\zeta_1 \\ -\zeta_2 & \zeta_1 & 0 \end{bmatrix} \quad \forall \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix}. \quad (20)$$

To quantify the translation mismatch between  $\mathcal{F}$  and  $\mathcal{F}^*$ , we define a translation error signal, denoted by  $e_v(t) \in \mathbb{R}^3$ , as follows

$$e_v = m_e - m_e^* \quad (21)$$

where  $m_e(t) \in \mathbb{R}^3$  denotes the extended coordinates [15] of an image point on  $\pi$  in terms of  $\mathcal{F}$  and is defined as follows<sup>3</sup>

$$m_e = [m_{e1} \quad m_{e2} \quad m_{e3}]^T = \begin{bmatrix} \frac{x_1}{z_1} & \frac{y_1}{z_1} & \ln(z_1) \end{bmatrix}^T \quad (22)$$

and  $m_e^* \in \mathbb{R}^3$  denotes the extended coordinates of the corresponding desired image point on  $\pi$  in terms of  $\mathcal{F}^*$  and is defined as follows

$$m_e^* = [m_{e1}^* \quad m_{e2}^* \quad m_{e3}^*]^T = \begin{bmatrix} \frac{x_1^*}{z_1^*} & \frac{y_1^*}{z_1^*} & \ln(z_1^*) \end{bmatrix}^T \quad (23)$$

<sup>3</sup>Any point  $O_i$  on  $\pi$  can be utilized in the subsequent development; however, to reduce the notational complexity, we have elected to select the image point  $O_1$ , and hence, the subscript 1 is utilized in lieu of  $i$  in the subsequent development.

where  $\ln(\cdot)$  denotes the natural logarithm. The first two elements of  $e_v(t)$  can be measured based on the fact that  $m_i(t)$  and  $m_i^*$  can be computed from the image-space by using (5) and (7). Given (21)-(23), the fact that

$$\ln(z_1) - \ln(z_1^*) = -\ln(\alpha_1), \quad (24)$$

and the fact that  $\alpha_1(t)$  can be determined through the decomposition of the Euclidean homography, it is clear that the third element of  $e_v(t)$  is also measurable.

**Remark 2** To obtain  $u(t)$  and  $\theta(t)$  from a given rotation matrix  $R(t)$ , the following expressions can be utilized [16]

$$\cos \theta = \frac{1}{2} (\text{tr}(R) - 1) \quad (25)$$

$$[u]_{\times} = \frac{R - R^T}{2 \sin(\theta)} \quad (26)$$

where the notation  $\text{tr}(\cdot)$  denotes the trace of a matrix.

**Remark 3** Given the structure of the skew symmetric matrix given in (20), the following property can be easily proven

$$\zeta^T [\zeta]_{\times} = [0 \ 0 \ 0] \quad \forall \zeta \in \mathbb{R}^3. \quad (27)$$

## 3 Error System Development

### 3.1 Open-Loop Dynamics

After taking the time derivative of (17) and (21) and performing some mathematical manipulation, the open-loop dynamics for  $e_v(t)$  and  $e_\omega(t)$  can be expressed in terms of the camera velocities as follows [8, 15]

$$\dot{e}_\omega = -L_\omega \omega_c \quad (28)$$

$$z_1^* \dot{e}_v = -\alpha_1 L_v v_c + z_1^* L_{(v,\omega)} \omega_c \quad (29)$$

where  $v_c(t), \omega_c(t) \in \mathbb{R}^3$  represent the linear and angular velocities of the camera expressed in  $\mathcal{F}$ , respectively. In (29) and (28), the measurable matrices  $L_v(m_{e1}, m_{e2}), L_{(v,\omega)}(m_{e1}, m_{e2}), L_\omega(u, \theta) \in \mathbb{R}^{3 \times 3}$  are defined as follows

$$L_v = \begin{bmatrix} 1 & 0 & -m_{e1} \\ 0 & 1 & -m_{e2} \\ 0 & 0 & 1 \end{bmatrix} \quad (30)$$

$$L_{(v,\omega)} = \begin{bmatrix} m_{e1}m_{e2} & -1 - m_{e1}^2 & m_{e2} \\ 1 + m_{e2}^2 & -m_{e1}m_{e2} & -m_{e1} \\ -m_{e2} & m_{e1} & 0 \end{bmatrix} \quad (31)$$

and

$$L_\omega = I_3 - \frac{\theta}{2} [u]_{\times} + \left( 1 - \frac{\text{sinc}(\theta)}{\text{sinc}^2\left(\frac{\theta}{2}\right)} \right) [u]_{\times}^2 \quad (32)$$

where

$$\text{sinc}(\theta(t)) = \frac{\sin \theta(t)}{\theta(t)}. \quad (33)$$

The open-loop dynamics given in (29) and (28) are expressed in terms of the linear and angular camera velocities. Since the camera is mounted on the end-effector of the robot manipulator, a relationship can be developed to relate the linear and angular camera velocities to the actual kinematic control input signals given by the linear and angular velocities of the robot

end-effector. This relationship is dependent on the extrinsic calibration parameters related to the orientation of the camera with respect to the end-effector. Typically, visual servo control designs assume that these extrinsic parameters are known. To enhance the robustness and accuracy of the subsequent control development, we relax the assumption that these extrinsic parameters are known. Specifically, the relationship between the linear and angular velocity of the camera with respect to the end-effector is given as follows [8]

$$\begin{bmatrix} v_c \\ \omega_c \end{bmatrix} = \begin{bmatrix} R_{off}^T R_0^T & 0 \\ 0 & R_{off}^T R_0^T \end{bmatrix} \begin{bmatrix} v_e \\ \omega_e \end{bmatrix} \quad (34)$$

where  $R_{off} \in \mathbb{R}^{3 \times 3}$  denotes an unknown, constant rotation offset between the camera frame  $\mathcal{F}$  and the end-effector frame  $\mathcal{F}_e$ ,  $R_0(t) \in \mathbb{R}^{3 \times 3}$  represents the measurable, time-varying rotation between the robot frame  $\mathcal{F}_o$  and the end-effector frame  $\mathcal{F}_e$ , and  $v_e(t), \omega_e(t) \in \mathbb{R}^3$  denote the linear velocity and angular velocity of the end-effector expressed in  $\mathcal{F}_e$ , respectively. After substituting the relationship given in (34) into (29) and (28) for  $v_c(t)$  and  $\omega_c(t)$ , the open-loop kinematics can be written in terms of the end-effector velocity inputs as follows

$$z_1^* \dot{e}_v = -\alpha_1 L_v R_{off}^T R_0^T v_e + z_1^* L_{(v,\omega)} R_{off}^T R_0^T \omega_e \quad (35)$$

$$\dot{e}_\omega = -L_\omega R_{off}^T R_0^T \omega_e. \quad (36)$$

**Remark 4** The matrix  $\frac{1}{2}(R_{off} + R_{off}^T)$  is positive definite, and there exist some positive constants  $\lambda_1$  and  $\lambda_2$  that satisfy the following inequalities (see the Appendix)

$$\lambda_1 \|\zeta\|^2 \leq \zeta^T \left( \frac{R_{off} + R_{off}^T}{2} \right) \zeta \leq \lambda_2 \|\zeta\|^2 \quad \forall \zeta \in \mathbb{R}^3. \quad (37)$$

In the subsequent stability analysis, we will also exploit the following fact

$$x^T R_{off} f x = x^T R_{off}^T f x = x^T \left( \frac{R_{off} + R_{off}^T}{2} \right) x \quad \forall x \in \mathbb{R}^3. \quad (38)$$

### 3.2 Control Development

Given the open-loop error dynamics in (35) and (36), and the subsequent stability analysis, the end-effector velocity control signals are designed as follows

$$v_e = \frac{1}{\alpha_1} \left( \frac{k_v}{f(m_{e1}, m_{e2})} R_0 L_v^T e_v + R_0 v_R \right) \quad (39)$$

$$\omega_e = k_\omega R_0 e_\omega \quad (40)$$

where  $k_v, k_\omega \in \mathbb{R}$  denotes positive control gains. In (39),  $f(m_{e1}, m_{e2}) \in \mathbb{R}$  denotes the following nonnegative function

$$f(m_{e1}, m_{e2}) = \frac{1}{6} m_{e1}^2 + \frac{1}{6} m_{e2}^2 + \frac{1}{3} - \frac{1}{3} \sqrt{\left( \frac{1}{2} m_{e1}^2 + \frac{1}{2} m_{e2}^2 + 1 \right)^2 - 1}, \quad (41)$$

and the auxiliary variable  $v_R(t) \in \mathbb{R}$  is defined as follows

$$v_R = k_{n1} k_w^2 \rho^2 L_v^T e_v. \quad (42)$$

In (42),  $k_{n1} \in \mathbb{R}$  denotes a positive control gain, and  $\rho(|m_{e1}|, |m_{e2}|) \in \mathbb{R}$  denotes a positive bounding function that must be selected to satisfy the following inequality

$$\rho(|m_{e1}|, |m_{e2}|) \geq \|z_1^* L_v^{-1} L_{(v,\omega)}\|. \quad (43)$$

After substituting (39), (40), and (42) into (35) and (36), the following closed-loop error dynamics are obtained

$$\begin{aligned} z_1^* \dot{e}_v &= -\frac{k_v}{f(m_{e1}, m_{e2})} L_v R_{off}^T L_v^T e_v \\ &\quad -k_{n1} k_\omega \rho^2 L_v R_{off}^T L_v^T e_v \\ &\quad +k_\omega L_v \left( z_1^* L_v^{-1} L_{(v,\omega)} R_{off}^T e_\omega \right) \end{aligned} \quad (44)$$

$$\dot{e}_\omega = -k_\omega L_\omega R_{off}^T e_\omega. \quad (45)$$

**Remark 5** *Motivation for the structure of  $f(m_{e1}, m_{e2})$  of (41) is due to the fact that the the following relationship can be determined (proof available upon request)*

$$x^T \left( L_v L_v^T \right) x \geq f(m_{e1}, m_{e2}) \|x\|^2 \quad \forall x \in \mathbb{R}^3. \quad (46)$$

The structure of (41) is also motivated by the fact that if  $m_{e1}(t), m_{e2}(t) \in L_\infty$ , then  $f(m_{e1}, m_{e2}) \in L_\infty$  and  $f(m_{e1}, m_{e2})$  can be lower bounded by a positive constant  $c_1 \in \mathbb{R}$  as follows (proof available upon request)

$$f(m_{e1}, m_{e2}) > c_1. \quad (47)$$

## 4 Stability Analysis

**Theorem 1** *The control input given in (39), (42), and (40) ensures that all signals are bounded during closed-loop operation and that the translation and rotation error signals,  $e_v(t)$  and  $e_\omega(t)$ , defined in (21) and (17), respectively, are exponentially regulated in the sense that*

$$\|e_v(t)\|, \|e_\omega(t)\| \leq \sqrt{\frac{\lambda_4}{\lambda_3}} \|z(0)\| \exp(-\lambda_0 t) \quad (48)$$

provided the control gain  $k_{n1}$  introduced in (42) is selected to satisfy the following constraint

$$k_{n1} > \frac{1}{\lambda_1 k_\omega} \quad (49)$$

where the vector  $z(t) \in \mathbb{R}^6$  is defined as follows

$$z = \begin{bmatrix} e_\omega^T & e_v^T \end{bmatrix}^T \quad (50)$$

and the positive constants  $\lambda_3$ ,  $\lambda_4$  and  $\lambda_0 \in \mathbb{R}$  are defined as follows

$$\lambda_3 = \min \left\{ \frac{1}{2}, \frac{z_1^*}{2} \right\} \quad \lambda_4 = \max \left\{ \frac{1}{2}, \frac{z_1^*}{2} \right\} \quad (51)$$

$$\lambda_0 = \frac{1}{\lambda_4} \min \left\{ \left( \lambda_1 k_\omega - \frac{1}{\lambda_1 k_{n1}} \right), \lambda_1 k_v \right\}. \quad (52)$$

**Proof:** To prove Theorem 1, we define a non-negative function denoted by  $V(t) \in \mathbb{R}$  as follows

$$V = \frac{1}{2} e_\omega^T e_\omega + \frac{1}{2} z_1^* e_v^T e_v \quad (53)$$

where  $V(t)$  can be lower and upper bounded as follows

$$\lambda_3 \|z\|^2 \leq V \leq \lambda_4 \|z\|^2 \quad (54)$$

where  $z(t)$  was defined in (50), and the positive constants  $\lambda_3$  and  $\lambda_4$  were defined in (51). After taking the time derivative of (53) and then utilizing (44) and (45), the following expression can be obtained

$$\begin{aligned} \dot{V} &= -\frac{k_v}{f(m_{e1}, m_{e2})} \left( L_v^T e_v \right)^T R_{off}^T L_v^T e_v \\ &\quad -k_{n1} k_\omega \rho^2 \left( L_v^T e_v \right)^T R_{off}^T L_v^T e_v \\ &\quad +k_\omega e_v^T L_v \left( z_1^* L_v^{-1} L_{(v,\omega)} \right) R_{off}^T e_\omega - k_\omega e_\omega^T L_\omega R_{off}^T e_\omega. \end{aligned} \quad (55)$$

Given (27) and the definition of  $e_\omega(t)$  in (17), it is straightforward to prove that

$$e_\omega^T L_\omega = e_\omega^T. \quad (56)$$

By utilizing (37), (38), (43), and (56), we can rewrite (55) as follows

$$\begin{aligned} \dot{V} &\leq -\lambda_1 k_\omega \|e_\omega\|^2 - \lambda_1 \frac{k_v}{f(m_{e1}, m_{e2})} \left\| L_v^T e_v \right\|^2 \\ &\quad + \left[ k_\omega \rho \left\| L_v^T e_v \right\| \|e_\omega\| - \lambda_1 k_{n1} k_\omega \rho^2 \left\| L_v^T e_v \right\|^2 \right] \end{aligned} \quad (57)$$

where the fact that

$$\left\| R_{off}^T e_\omega \right\| = \|e_\omega\| \quad (58)$$

has been utilized. By applying the nonlinear damping argument [12] to the bracketed term in (57) and utilizing (46), the following expression can be obtained

$$\dot{V} \leq -\left( \lambda_1 k_\omega - \frac{1}{\lambda_1 k_{n1}} \right) \|e_\omega\|^2 - \lambda_1 k_v \|e_v\|^2. \quad (59)$$

After utilizing (49), (50), (52), and (54), the expression in (59) can be upper bounded as follows

$$\dot{V} \leq -\lambda_0 V. \quad (60)$$

The differential inequality of (60) can now be solved to yield the following expression

$$V(t) \leq V(0) \exp(-\lambda_0 t). \quad (61)$$

By utilizing (54), the following upper bound for  $z(t)$  can be developed from (61)

$$\|z(t)\| \leq \sqrt{\frac{\lambda_4}{\lambda_3}} \|z(0)\| \exp(-\lambda_0 t). \quad (62)$$

Based on (50), it is straightforward that  $e_v(t)$ ,  $e_\omega(t)$  are bounded by the exponential envelope given in (48). Standard signal chasing arguments can now be applied to conclude that all of the signals in the closed-loop system remain bounded. ■

## 5 Conclusions

In this paper, Lyapunov-based design and analysis techniques are utilized to develop a robust visual servoing controller that ensures exponential regulation of the camera translation and rotation errors. The result is obtained using a single camera, despite uncertainty associated with the both the angle and

the axis of rotation of the camera with respect to the robot end-effector and unmeasurable depth information. To facilitate the result, a Euclidean homography is decomposed into separate translation and rotation components. By decomposing the homography in this manner, both 2D image-space and projected 3D task-space (i.e., 2.5D visual servoing) information are exploited to construct the controller. Based on the results given in our previous work (see [8]), the result in this paper can easily be extended to incorporate the robot dynamics via integrator backstepping. Future efforts will also target experimental verification of the control strategy.

## References

[1] F. Chaumette and E. Malis, “2 1/2 D Visual Servoing: A Possible Solution to Improve Image-Based and Position-Based Visual Servoings,” *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 630-635, 2000.

[2] F. Chaumette, E. Malis, and S. Boudet, “2D 1/2 Visual Servoing With Respect to a Planar Object,” *Proc. of the Workshop on New Trends in Image-Based Robot Servoing*, pp. 45-52, 1997.

[3] F. Conticelli and B. Allotta, “Nonlinear Controllability and Stability Analysis of Adaptive Image-Based Systems,” *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 2, 2001.

[4] F. Conticelli and B. Allotta, “Discrete-Time Robot Visual Feedback in 3-D Positioning Tasks with Depth Adaptation,” *IEEE/ASME Transactions on Mechatronics*, Vol. 6, No. 3, pp. 356-363, 2001.

[5] P. I. Corke and S. A. Hutchinson, “A New Hybrid Image-Based Visual Servo Control Scheme,” *Proc. of the IEEE Conference on Decision and Control*, pp. 2521-2527, 2000.

[6] K. Deguchi, “Optimal Motion Control for Image-Based Visual Servoing by Decoupling Translation and Rotation,” *Proceedings of the Intl. Conf. on Intelligent Robots and Systems*, B.C. Canada, pp. 705-711, Oct. 1998.

[7] W. E. Dixon, D. M. Dawson, E. Zergeroglu, and A. Behal, “Adaptive Tracking Control of a Wheeled Mobile Robot via an Uncalibrated Camera System,” *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, Vol. 31, No. 3, pp. 341-352, 2001.

[8] Y. Fang, A. Behal, W. E. Dixon and D. M. Dawson, “Adaptive 2.5D Visual Servoing of Kinematically Redundant Robot Manipulators,” *Proc. of the IEEE Conference on Decision and Control*, Las Vegas, NV, December 2002, pp. 2860-2865.

[9] Y. Fang, D. M. Dawson, W. E. Dixon, and M. S. de Queiroz, “2.5D Visual Servoing of Wheeled Mobile Robots,” *Proc. of the IEEE Conference on Decision and Control*, Las Vegas, NV, December 2002, pp. 2866-2871.

[10] O. Faugeras and F. Lustman, “Motion and Structure from Motion in a Piecewise Planar Environment,” *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 2, No. 3, pp. 485-508, 1988.

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

[12] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*, John Wiley & Sons, New York, NY, 1995.

[13] E. Malis and F. Chaumette, “Theoretical Improvements in the Stability Analysis of a New Class of Model-Free Visual Servoing Methods,” *IEEE Transactions on Robotics and Automation*, Vol. 18, No. 2, pp. 176-186, 2002.

[14] E. Malis and F. Chaumette, “2 1/2 D Visual Servoing With Respect to Unknown Objects Through A New Estimation Scheme of Camera Displacement,” *International Journal of Computer Vision*, Vol. 37, No. 1, pp. 79-97, 2000.

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

[16] M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*, John Wiley and Sons, Inc: New York, NY, 1989.

[17] C. J. Taylor and J. P. Ostrowski, “Robust Vision-Based Pose Control,” *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 2734-2740, 2000.

[18] E. Zergeroglu, D. M. Dawson, M. de Queiroz, and A. Behal, “Vision-Based Nonlinear Tracking Controllers in the Presence of Parametric Uncertainty,” *IEEE/ASME Transactions on Mechatronics*, Vol. 6, No. 3, pp. 322-337, Sept. 2001.

[19] Z. Zhang and A. R. Hanson, “Scaled Euclidean 3D Reconstruction Based on Externally Uncalibrated Cameras,” *IEEE Symp. on Computer Vision*, pp. 37-42, 1995.

## Appendix

The constant rotation matrix  $R_{off}$  can be written as a function of a rotation angle  $\theta_0$  and rotation axis  $u_0$  as follows

$$R_{off} = I_3 + \sin(\theta_0) [u_0]_{\times} + 2 \sin^2\left(\frac{\theta_0}{2}\right) [u_0]_{\times}^2 \quad (63)$$

where we assume that

$$|\theta_0| < \frac{\pi}{2}. \quad (64)$$

After utilizing the fact that

$$R_{off}^T = I_3 - \sin(\theta_0) [u_0]_{\times} + 2 \sin^2\left(\frac{\theta_0}{2}\right) [u_0]_{\times}^2 \quad (65)$$

and the following property [8]

$$[u_0]_{\times}^2 = u_0 u_0^T - I_3, \quad (66)$$

the following expression can be obtained

$$\frac{R_{off} + R_{off}^T}{2} = \left[1 - 2 \sin^2\left(\frac{\theta_0}{2}\right)\right] I_3 + 2 \sin^2\left(\frac{\theta_0}{2}\right) u_0 u_0^T. \quad (67)$$

Given the following property

$$x^T \left(u_0 u_0^T\right) x = \left(u_0^T x\right)^2 \geq 0 \quad \forall x \in \mathbb{R}^3, \quad (68)$$

the matrix  $u_0 u_0^T$  is positive semi-definite. Based on this fact, we can utilize (64) to prove that the expression given in (67) is positive definite. Hence, the inequality given in (37) follows directly.