Adaptive Homography-Based Visual Servo Tracking*

J. Chen,[1] D. M. Dawson,[1] W. E. Dixon,[2] and A. Behal[1]

[1]Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634-0915
[2]Eng. Science and Tech. Division-Robotics, Oak Ridge National Lab., P.O. Box 2008, Oak Ridge, TN 37831-6305

To appear at the IEEE International Conference on Intelligent Robots and Systems (IROS 2003)
Las Vegas, NV, October 27 - 31, 2003

Keywords: Homograhpy, Adaptive 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.

Adaptive Homography-Based Visual Servo Tracking*

J. Chen†, D. M. Dawson†, W. E. Dixon‡, and A. Behal†

†Department of Electrical & Computer Engineering, Clemson University, Clemson, SC 29634-0915

‡Eng. Science and Tech. Div. - Robotics, Oak Ridge Nat. Lab., P.O. Box 2008, Oak Ridge, TN 37831-6305

E-mail: dixonwe@ornl.gov

Abstract— **In this paper, a homography-based adaptive visual servo controller is described to enable a robot end-effector to track a desired Euclidean space trajectory as determined by a sequence of images for the fixed camera configuration. To achieve the objective, a Lyapunov-based adaptive control strategy is employed to actively compensate for unknown depth measurements and the lack of an object model. The error systems are constructed as a hybrid of pixel information and reconstructed Euclidean variables obtained by comparing the images and decomposing a homographic relationship.**

I. Introduction

A key issue that impacts camera-based visual servo control is the relationship between the Euclidean-space and the image-space. One factor that impacts this relationship is the fact that the image-space is a 2-dimensional (2D) projection of the 3D Euclidean-space. To compensate for the lack of depth information from the 2D image data, some researchers have focused on the use of alternate sensors (e.g., laser and sound ranging technologies). While some applications may be suited to alternative vision sensors that provide depth information, many applications are ill suited for such technologies. Other researchers have explored the use of a camera-based vision system in conjunction with other sensors along with some sensor fusion method or the use of additional cameras in a stereo configuration that triangulate on corresponding images. However, the practical drawbacks of incorporating additional sensors include: increased cost, increased complexity, decreased reliability, and increased processing burden to condition and fuse sensor data. Motivated by these practical insights, recent research has focused on monocular camera-based visual servo strategies that rely on analytic techniques to address the lack of depth information. One strategy that has recently been employed involves the use of partitioning methods that exploit a combination of reconstructed 3D Euclidean information and 2D image-space information. For example, in the series of papers by Malis and Chaumette (e.g., [1], [2], [18], [19]) various kinematic control strategies exploit the fact that the interaction between translation and rotation components can be decoupled through a homography (i.e., a projective coordinate transformation). Specifically, information combined from the task-space (obtained through a Euclidean reconstruction from the image data) and the 2D image-space is utilized to regulate the translation and rotation error systems. In [8], 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. Corke and Hutchinson [5] also developed a hybrid image-based visual servoing scheme that decouples rotation and translation components from the remaining degrees of freedom. One drawback of some of the aforementioned controllers are claims (without a supporting proof) that a constant, best-guess estimate of the depth information can be utilized in lieu of the exact value. Motivated by the desire to actively compensate for unmeasurable depth information, Conticelli developed an adaptive kinematic controller in [3] to ensure uniformly ultimately bounded (UUB) set-point regulation, 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 [10] and [11], Fang et al. recently developed 2.5D visual servo controllers to asymptotically regulate a manipulator end-effector and a mobile robot, respectively, by developing an adaptive update law that actively compensates for an unknown depth parameter. In [12], Fang et al. also developed a camera-in-hand regulation controller that incorporated a robust control structure to compensate for uncertainty in the extrinsic calibration parameters.

After examining the literature, it is clearly evident that much of the previous visual servo controllers have only been designed to address the regulation problem. That is, the objective of most control designs is to force a hand-held camera to a Euclidean position defined by a static reference image. Unfortunately, many practical applications require a robotic system to move along a predefined or dynamically changing trajectory. For example, a human operator may predefine an image trajectory through a high-level interface, and this trajectory may need to be modified on-the-fly to respond to obstacles moving in and out of the environment. Moreover, it is well known that a regulating controller may produce erratic behavior and require excessive initial control torques if the initial error is large.

*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, and by U.S. NSF Grant DMI-9457967, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

Motivated by the need for new advancements to meet visual servo tracking applications, previous research has concentrated on developing different types of path planning techniques in the image-space (e.g., see [6], [21], [22], [23]). More recently, Mezouar and Chaumette developed a path-following image-based visual servo algorithm in [20] where the path to a goal point is generated via a potential function that incorporates motion constraints. In [7], Cowan et al. develop a hybrid position/image-space controller that forces a manipulator to a desired endpoint while avoiding obstacles and ensuring the object remains in the field-of-view by avoiding pitfalls such as self-occlusion.

In contrast to the approaches in [7] and [20] in which a path is planned as a means to reach a desired setpoint, hybrid tracking controllers are proposed in this paper where the robot end-effector is required to track a prerecorded time-varying reference trajectory. To develop the hybrid controllers, a homography-based visual servoing approach is utilized. The motivation for using this approach is that the visual servo control problem can be incorporated with a Lyapunov-based control design strategy to overcome many practical and theoretical obstacles associated with more traditional purely image-based approaches. Specifically, one of the challenges of this problem is that the translation error system is corrupted by an unknown depth related parameter. By formulating a Lyapunov-based argument, an adaptive update law is developed to actively compensate for the unknown depth parameter. In addition, the proposed approach facilitates: i) translation/rotational control in the full six degree-of-freedom task-space without the requirement of an object model, ii) partial servoing on pixel data that yields improved robustness and increases the likelihood that the centroid of the object remains in the camera field-of-view [19], and iii) the use of an image Jacobian that is only singular for multiples of 2π, and hence, eliminates the serious problem of singular image Jacobians inherent in many of the purely image-based controllers. The proposed homography-based controllers target the fixed camera configuration.

II. Geometric Model

To make the subsequent development more tractable, four target points located on an object (i.e., the end-effector of a robot manipulator) denoted by O_i $\forall i = 1, 2, 3, 4$ are considered to be coplanar[1] and not colinear. Based on this assumption, consider a fixed plane, denoted by π^*, that is defined by a reference image of the object. In addition, consider the actual and desired motion of the plane containing the end-effector feature points, denoted by π and π_d, respectively (see Fig. 1). To develop a relationship between the planes, an inertial coordinate system, denoted by $\mathcal{I}$, is defined where the origin coincides with the center of a fixed camera. The 3D coordinates of the target points on π, π_d, and π^* can be respectively expressed in terms of $\mathcal{I}$ as follows

$$\begin{aligned} \bar{m}_i(t) &\triangleq \begin{bmatrix} x_i(t) & y_i(t) & z_i(t) \end{bmatrix}^T \\ \bar{m}_{di}(t) &\triangleq \begin{bmatrix} x_{di}(t) & y_{di}(t) & z_{di}(t) \end{bmatrix}^T \\ \bar{m}_i^* &\triangleq \begin{bmatrix} x_i^* & y_i^* & z_i^* \end{bmatrix}^T \end{aligned} \tag{1}$$

under the standard assumption that the distances from the origin of $\mathcal{I}$ to the target points remains positive (i.e., $z_i(t)$, $z_{di}(t)$, $z_i^* > \varepsilon$ where ε denotes an arbitrarily small positive constant). Orthogonal coordinate systems $\mathcal{F}$, $\mathcal{F}_d$, and $\mathcal{F}^*$ are attached to the planes π, π_d, and π^*, respectively, where the origin of the coordinate systems coincides with the object (see Fig. 1). To relate the coordinate systems, let $R(t), R_d(t), R^* \in SO(3)$ denote the rotation between $\mathcal{F}$ and $\mathcal{I}$, $\mathcal{F}_d$ and $\mathcal{I}$, and $\mathcal{F}^*$ and $\mathcal{I}$, respectively, and let $x_f(t)$, $x_{fd}(t)$, $x_f^* \in \mathbb{R}^3$ denote the respective translation vectors expressed in the coordinates of $\mathcal{I}$. As also illustrated in Fig. 1, $n^* \in \mathbb{R}^3$ denotes the constant unit normal to the plane π^* expressed in the coordinates of $\mathcal{I}$, $s_i \in \mathbb{R}^3$ denotes the constant coordinates of the target points located on the object reference frame, and the constant distance $d^* \in \mathbb{R}$ from the origin of $\mathcal{I}$ to π^* along the unit normal is given by

$$d^* = n^{*T} \bar{m}_i^* \ . \tag{2}$$

From the geometry between the coordinate frames depicted in Fig. 1, the following relationships can be developed

$$\begin{aligned} \bar{m}_i &= x_f + R s_i \\ \bar{m}_{di} &= x_{fd} + R_d s_i \\ \bar{m}_i^* &= x_f^* + R^* s_i \ . \end{aligned} \tag{3}$$

After solving the third equation in (3) for s_i and then substituting the resulting expression into the first and second equations, the following relationships can be obtained

$$\bar{m}_i = \bar{x}_f + \bar{R} \bar{m}_i^* \qquad \bar{m}_{di} = \bar{x}_{fd} + \bar{R}_d \bar{m}_i^* \tag{4}$$

where $\bar{R}(t)$, $\bar{R}_d(t) \in SO(3)$ and $\bar{x}_f(t)$, $\bar{x}_{fd}(t) \in \mathbb{R}^3$ are new rotational and translational variables, respectively, defined as follows

$$\begin{aligned} \bar{R} &= R (R^*)^T \qquad & \bar{R}_d &= R_d (R^*)^T \\ \bar{x}_f &= x_f - \bar{R} x_f^* \qquad & \bar{x}_{fd} &= x_{fd} - \bar{R}_d x_f^* \ . \end{aligned} \tag{5}$$

From (2), it is easy to see how the relationships in (4) can now be expressed as follows

$$\begin{aligned} \bar{m}_i &= \left(\bar{R} + \frac{\bar{x}_f}{d^*} n^{*T} \right) \bar{m}_i^* \\ \bar{m}_{di} &= \left(\bar{R}_d + \frac{\bar{x}_{fd}}{d^*} n^{*T} \right) \bar{m}_i^* . \end{aligned} \tag{6}$$

Remark 1: The subsequent development requires that the constant rotation matrix R^* be known. This is a mild assumption since the constant rotation matrix R^* can be obtained a priori using various methods (e.g., a second camera, Euclidean measurements).

[1] It should be noted that that if that if four coplanar target points are not available then the subsequent development can exploit the classic eight-points algorithm [18] with no four of the eight target points being coplanar.

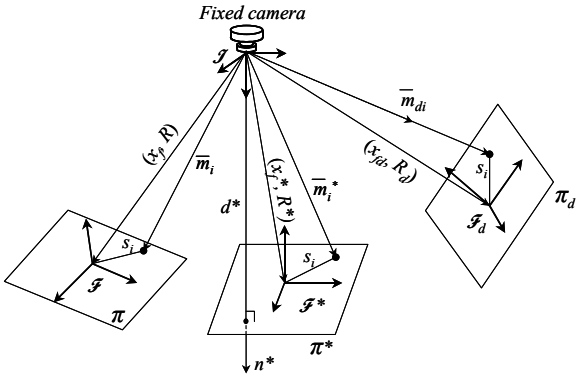

Fig. 1. Coordinate frame relationships.

III. Euclidean Reconstruction

The relationship given by (6) provides a means to quantify a translation and rotation error between $\mathcal{F}$ and $\mathcal{F}^*$ and between $\mathcal{F}_d$ and $\mathcal{F}^*$. Since the Euclidean position of $\mathcal{F}$, $\mathcal{F}_d$, and $\mathcal{F}^*$ cannot be directly measured, a Euclidean reconstruction is developed in this section to obtain the position and rotational error information by comparing multiple images. Specifically, comparisons are made between the current image acquired by the fixed camera, the reference image obtained a priori, and the a priori known sequence of images that define the trajectory of $\mathcal{F}_d$. To facilitate the subsequent development, the normalized Euclidean coordinates of the points on π, π_d, and π^* can be respectively expressed in terms of $\mathcal{I}$ as $m_i(t)$, $m_{di}(t)$, $m_i^* \in \mathbb{R}^3$, as follows

$$m_i \triangleq \frac{\bar{m}_i}{z_i} = \begin{bmatrix} \frac{x_i}{z_i} & \frac{y_i}{z_i} & 1 \end{bmatrix}^T \tag{7}$$

$$m_{di} \triangleq \frac{\bar{m}_{di}}{z_{di}} = \begin{bmatrix} \frac{x_{di}}{z_{di}} & \frac{y_{di}}{z_{di}} & 1 \end{bmatrix}^T \tag{8}$$

$$m_i^* \triangleq \frac{\bar{m}_i^*}{z_i^*} = \begin{bmatrix} \frac{x_i^*}{z_i^*} & \frac{y_i^*}{z_i^*} & 1 \end{bmatrix}^T. \tag{9}$$

The rotation and translation between the coordinate systems can now be related in terms of the normalized coordinates as follows

$$m_i = \underbrace{\frac{z_i^*}{z_i}}_{\alpha_i} \underbrace{\left(\bar{R} + \bar{x}_h n^{*T}\right)}_{H} m_i^* \tag{10}$$

$$m_{di} = \underbrace{\frac{z_i^*}{z_{di}}}_{\alpha_{di}} \underbrace{\left(\bar{R}_d + \bar{x}_{hd} n^{*T}\right)}_{H_d} m_i^* \tag{11}$$

where $\alpha_i(t)$, $\alpha_{di}(t) \in \mathbb{R}$ denote invertible depth ratios, $H(t)$, $H_d(t) \in \mathbb{R}^{3\times 3}$ denote Euclidean homographies, and $\bar{x}_h(t)$, $\bar{x}_{hd}(t) \in \mathbb{R}^3$ denote scaled translation vectors that are defined as follows

$$\bar{x}_h = \frac{\bar{x}_f}{d^*} \qquad \bar{x}_{hd} = \frac{\bar{x}_{fd}}{d^*}. \tag{12}$$

Each target point on π, π_d, and π^* will have a projected pixel coordinate expressed in terms of $\mathcal{I}$, denoted by $u_i(t)$, $v_i(t) \in \mathbb{R}$ for π, $u_{di}(t)$, $v_{di}(t) \in \mathbb{R}$ for π_d, and u_i^*, $v_i^* \in \mathbb{R}$ for π^*, that are defined as elements of $p_i(t)$ (i.e., the actual time-varying target points), $p_{di}(t)$ (i.e., the desired time-varying target point trajectory), and p_i^* (i.e., the constant reference target points), respectively, as follows

$$p_i \triangleq \begin{bmatrix} u_i & v_i & 1 \end{bmatrix}^T \qquad p_{di} \triangleq \begin{bmatrix} u_{di} & v_{di} & 1 \end{bmatrix}^T \tag{13}$$
$$p_i^* \triangleq \begin{bmatrix} u_i^* & v_i^* & 1 \end{bmatrix}^T.$$

To calculate the Euclidean homography given in (10) and (11) from pixel information, the projected 2D pixel coordinates of the target points are related to $m_i(t)$, $m_{di}(t)$, and m_i^* by the following pin-hole lens models [13]

$$p_i = Am_i \qquad p_{di} = Am_{di} \qquad p_i^* = Am_i^* \tag{14}$$

where $A \in \mathbb{R}^{3\times 3}$ is a known, constant, and invertible intrinsic camera calibration matrix. After substituting (14) into (10) and (11), the following relationships can be developed

$$p_i = \alpha_i \underbrace{\left(AHA^{-1}\right)}_{G} p_i^* \qquad p_{di} = \alpha_{di} \underbrace{\left(AH_dA^{-1}\right)}_{G_d} p_i^* \tag{15}$$

where $G(t) = [g_{ij}(t)]$, $G_d(t) = [g_{dij}(t)]$ $\forall i, j = 1, 2, 3 \in \mathbb{R}^{3\times 3}$ denote projective homographies. From the first relationship in (15), a set of 12 linearly independent equations given by the 4 target point pairs $(p_i^*, p_i(t))$ with 3 independent 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). From the definition of $G(t)$ given in (15), various techniques can then be used (e.g., see [14], [26]) to decompose the Euclidean homography, to obtain $\alpha_i(t)$, $G(t)$, $H(t)$, and the rotation and translation signals $\bar{R}(t)$ and $\bar{x}_h(t)$, and n^*. Likewise, by using the target point pairs $(p_i^*, p_{di}(t))$, the desired Euclidean homography can be decomposed to obtain $\alpha_{di}(t)$, $G_d(t)$, $H_d(t)$, and the desired rotation and translation signals $\bar{R}_d(t)$ and $\bar{x}_{hd}(t)$. The rotation matrices $R(t)$ and $R_d(t)$ can be computed from $\bar{R}(t)$ and $\bar{R}_d(t)$ by using (5) and the fact that R^* is assumed to be known. Hence, $R(t)$, $\bar{R}(t)$, $R_d(t)$, $\bar{R}_d(t)$, $\bar{x}_h(t)$, $\bar{x}_{hd}(t)$, and the depth ratios $\alpha_i(t)$ and $\alpha_{di}(t)$ are all known signals that can be used for control synthesis.

IV. Control Objective

The objective is to develop a visual servo controller that ensures that the trajectory of $\mathcal{F}$ tracks $\mathcal{F}_d$ (i.e., $\bar{m}_i(t)$ tracks $\bar{m}_{di}(t)$), where the trajectory of $\mathcal{F}_d$ is constructed relative to the reference camera position/orientation given by $\mathcal{F}^*$. To ensure that $\bar{m}_i(t)$ tracks $\bar{m}_{di}(t)$ from the Euclidean reconstruction given in (10) and (11), the tracking control objective can be stated as follows[2]: $\bar{R}(t) \to \bar{R}_d(t)$, $m_1(t) \to$

[2]Any point O_i 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.

$m_{d1}(t)$, and $z_1(t) \to z_{d1}(t)$ (and hence, $\bar{x}_h(t) \to \bar{x}_{hd}(t)$). The 3D control objective is complicated by the fact that only 2D image information is measureable. That is, while the development of the homography provides a means to reconstruct some Euclidean information, the formulation of a controller is challenging due to the fact that the time varying signals $z_1(t)$ and $z_{d1}(t)$ are not measurable. In addition, it is desirable to servo on actual pixel information (in lieu of reconstructed Euclidean information) to improve robustness to intrinsic camera calibration parameters and to increase the likelihood that the object will stay in the field of view of the camera.

To reformulate the control objective in light of these issues, a hybrid translation tracking error, denoted by $e_v(t) \in \mathbb{R}^3$, is defined as follows

$$e_v = p_e - p_{ed} \tag{16}$$

where $p_e(t)$, $p_{ed}(t) \in \mathbb{R}^3$ are defined as follows

$$\begin{aligned} p_e &= \left[\begin{array}{ccc} u_1 & v_1 & -\ln(\alpha_1) \end{array} \right]^T \\ p_{ed} &= \left[\begin{array}{ccc} u_{d1} & v_{d1} & -\ln(\alpha_{d1}) \end{array} \right]^T \end{aligned} \tag{17}$$

and $\ln(\cdot)$ denotes the natural logarithm. A rotation tracking error, denoted by $e_\omega(t) \in \mathbb{R}^3$, is defined as follows

$$e_\omega \triangleq \Theta - \Theta_d \tag{18}$$

where $\Theta(t)$, $\Theta_d(t) \in \mathbb{R}^3$ denote the axis-angle representation of $\bar{R}(t)$ and $\bar{R}_d(t)$ as follows [25]

$$\Theta = u(t)\theta(t) \quad \Theta_d = u_d(t)\theta_d(t). \tag{19}$$

For the representations in (19), $u(t)$, $u_d(t) \in \mathbb{R}^3$ represent unit rotation axes, and $\theta(t), \theta_d(t) \in \mathbb{R}$ denote the respective rotation angles about $u(t)$ and $u_d(t)$ that are assumed to be confined to the following regions

$$-\pi < \theta(t) < \pi \qquad -\pi < \theta_d(t) < \pi \ . \tag{20}$$

Based on the error system formulations in (16) and (18), the control objective can be stated as the desire to regulate the tracking error signals $e_v(t)$ and $e_\omega(t)$ to zero. If the tracking error signals $e_v(t)$ and $e_\omega(t)$ are regulated to zero then the object can be proven to be tracking the desired trajectory (details of this proof are available upon request).

Remark 2: As stated in [25], the axis-angle representation in (19) is not unique, in the sense that a rotation of $-\theta(t)$ about $-u(t)$ is equal to a rotation of $\theta(t)$ about $u(t)$. A particular solution for $\theta(t)$ and $u(t)$ can be determined as follows [25]

$$\theta_p = \cos^{-1}\left(\frac{1}{2}\left(\mathrm{tr}\left(\bar{R}\right) - 1\right)\right) \qquad [u_p]_\times = \frac{\bar{R} - \bar{R}^T}{2\sin(\theta_p)} \tag{21}$$

where the notation $\mathrm{tr}(\cdot)$ denotes the trace of a matrix and $[u_p]_\times$ denotes the 3×3 skew-symmetric expansion of $u_p(t)$. From (21), it is clear that

$$0 \le \theta_p(t) \le \pi. \tag{22}$$

While (22) is confined to a smaller region than $\theta(t)$ in (20), it is not more restrictive in the sense that

$$u_p\theta_p = u\theta. \tag{23}$$

The constraint in (22) is consistent with the computation of $[u(t)]_\times$ in (21) since a clockwise rotation (*i.e.*, $-\pi \le \theta(t) \le 0$) is equivalent to a counterclockwise rotation (*i.e.*, $0 \le \theta(t) \le \pi$) with the axis of rotation reversed. Hence, based on (23) and the functional structure of the object kinematics, the particular solutions $\theta_p(t)$ and $u_p(t)$ can be used in lieu of $\theta(t)$ and $u(t)$ without loss of generality and without confining $\theta(t)$ to a smaller region. Since, we do not distinguish between rotations that are off by multiples of 2π, all rotational possibilities are considered via the parameterization of (19) along with the computation of (21). Likewise, particular solutions can be found in the same manner for $\theta_d(t)$ and $u_d(t)$.

Remark 3: To develop a tracking control design, it is typical that the desired trajectory is used as a feedforward component in the control design. Hence, for a kinematic controller the desired trajectory is required to be at least first order differentiable and at least second order differentiable for a dynamic level controller. To this end, a sufficiently smooth function (e.g., a spline function) is used to fit the sequence of target points to generate the desired trajectory $p_{di}(t)$; hence, it is assumed that $p_{ed}(t)$ and $\dot{p}_{ed}(t)$ are bounded functions of time. From the projective homography introduced in (15), $p_{di}(t)$ can be expressed in terms of the a priori known, functions $\alpha_{di}(t)$, $H_d(t)$, $\bar{R}_d(t)$, and $\bar{x}_{hd}(t)$. Since these signals can be obtained from the prerecorded sequence of images, sufficiently smooth functions can also be generated for these signals by fitting a sufficiently smooth spline function to the signals. Hence, in practice, the a priori developed smooth functions $\alpha_{di}(t)$, $\bar{R}_d(t)$,and $\bar{x}_{hd}(t)$ can be constructed as bounded functions with bounded time derivatives. Based on the assumption that $\bar{R}_d(t)$ is a bounded first order differentiable function with a bounded derivative, (21) can be used to conclude that $u_d(t)$ and $\theta_d(t)$ are bounded first order differentiable functions with a bounded derivative; hence, $\Theta_d(t)$ and $\dot{\Theta}_d(t)$ can be assumed to be bounded. In the subsequent tracking control development, the desired signals $\dot{p}_{ed}(t)$ and $\dot{\Theta}_d(t)$ will be used as a feedforward control term.

V. Control Formulation

A. Open-Loop Error System

To develop the open-loop error system for $e_\omega(t)$, we take the time derivative of (18) to obtain the following expression

$$\dot{e}_\omega = L_\omega R\omega_e - \dot{\Theta}_d \tag{24}$$

where $L_\omega(t) \in \mathbb{R}^{3\times3}$ denotes a Jacobian-like matrix (see [10], [17]) and $\omega_e(t) \in \mathbb{R}^3$ denotes the angular velocity of the object expressed in $\mathcal{F}$. The determinant of $L_\omega(t)$ is only singular for multiples of 2π (i.e., out of the assumed workspace); therefore, $L_\omega(t)$ is invertible in the assumed workspace. To develop the open-loop error system for $e_v(t)$,

we take the time derivative of (16) to obtain the following expression

$$z_1^* \dot{e}_v = \alpha_1 A_e L_v R \left[v_e + [\omega_e]_\times s_1 \right] - z_1^* \dot{p}_{ed} \tag{25}$$

where $v_e(t) \in \mathbb{R}^3$ denotes the linear velocity of the object expressed in $\mathcal{F}$. In (25), $A_e \in \mathbb{R}^{3\times 3}$ is defined as follows

$$A_e = A - \begin{bmatrix} 0 & 0 & u_0 \\ 0 & 0 & v_0 \\ 0 & 0 & 0 \end{bmatrix} \tag{26}$$

where u_0, $v_0 \in \mathbb{R}$ denote the pixel coordinates of the principal point (i.e., the image center that is defined as the frame buffer coordinates of the intersection of the optical axis with the image plane), and the auxiliary Jacobian-like matrix $L_v(t) \in \mathbb{R}^{3\times 3}$ is defined as

$$L_v = \begin{bmatrix} 1 & 0 & -\frac{x_1}{z_1} \\ 0 & 1 & -\frac{y_1}{z_1} \\ 0 & 0 & 1 \end{bmatrix} . \tag{27}$$

Remark 4: It is easy to show that the product $A_e L_v$ is an invertible upper triangular matrix from (26) and (27).

B. Closed-Loop Error System

Based on the structure of the open-loop error systems and subsequent stability analysis, the angular and linear camera velocity control inputs for the object are defined as follows

$$\omega_e = R^T L_\omega^{-1} (\dot{\Theta}_d - K_\omega e_\omega) \tag{28}$$

$$v_e = -\frac{1}{\alpha_1} R^T \left(A_e L_v \right)^{-1} \left(K_v e_v - \hat{z}_1^* \dot{p}_{ed} \right) - [\omega_e]_\times \hat{s}_1 . \tag{29}$$

In (28) and (29), K_ω, $K_v \in \mathbb{R}^{3\times 3}$ denote diagonal matrices of positive constant control gains, and $\hat{z}_1^*(t) \in \mathbb{R}$, $\hat{s}_1(t) \in \mathbb{R}^3$ denote parameter estimates that are generated according to the following differential equations

$$\dot{\hat{z}}_1^*(t) = -\gamma_1 e_v^T \dot{p}_{ed} \tag{30}$$

$$\dot{\hat{s}}_1 = -\alpha_1 \Gamma_2 [\omega_e]_\times R^T L_v^T A_e^T e_v \tag{31}$$

where $\gamma_1 \in \mathbb{R}$ denotes a positive constant adaptation gain, and $\Gamma_2 \in \mathbb{R}^{3\times 3}$ denotes a positive constant diagonal adaptation gain matrix. After substituting (28) into (24), the following closed-loop error dynamics can be obtained

$$\dot{e}_\omega = -K_\omega e_\omega . \tag{32}$$

After substituting (29) into (25), the closed-loop translation error dynamics can be determined as follows

$$z_1^* \dot{e}_v = -K_v e_v + \alpha_1 A_e L_v R [\omega_e]_\times \tilde{s}_1 - \tilde{z}_1^* \dot{p}_{ed} \tag{33}$$

where the parameter estimation error signals $\tilde{z}_1^*(t) \in \mathbb{R}$ and $\tilde{s}_1(t) \in \mathbb{R}^3$ are defined as follows

$$\tilde{z}_1^* = z_1^* - \hat{z}_1^* \qquad \tilde{s}_1 = s_1 - \hat{s}_1 . \tag{34}$$

VI. Stability Analysis

Theorem 1: The adaptive update laws defined in (30) and (31) along with the control inputs designed in (28) and (29) ensure that $e_\omega(t)$ and $e_v(t)$ are asymptotically driven to zero in the sense that

$$\lim_{t\to\infty} \|e_\omega(t)\| , \ \|e_v(t)\| = 0 . \tag{35}$$

Proof: To prove Theorem 1, a non-negative function $V(t) \in \mathbb{R}$ is defined as follows

$$V \triangleq \frac{1}{2} e_\omega^T e_\omega + \frac{z_1^*}{2} e_v^T e_v + \frac{1}{2\gamma_1} \tilde{z}_1^{*2} + \frac{1}{2} \tilde{s}_1^T \Gamma_2^{-1} \tilde{s}_1 . \tag{36}$$

After taking the time derivative of (36) and then substituting for the closed-loop error systems developed in (32) and (33), utilizing the time derivative of (34), and substituting the adaptive update laws designed in (30) and (31), the following simplified expression can be obtained

$$\dot{V} = -e_\omega^T K_\omega e_\omega - e_v^T K_v e_v \tag{37}$$

where the fact that $[\omega_e]_\times^T = -[\omega_e]_\times$ was utilized. Based on (34), (36), and (37), it can be determined that $e_\omega(t)$, $e_v(t)$, $\tilde{z}_1^*(t)$, $\hat{z}_1^*(t)$, $\tilde{s}_1(t)$, $\hat{s}_1(t) \in \mathcal{L}_\infty$ and that $e_\omega(t)$, $e_v(t) \in \mathcal{L}_\in$ [9]. Based on the assumption that $\dot{\Theta}_d(t)$ is designed as a bounded function, the expressions given in (18) and (28) can be used to conclude that $\omega_e(t) \in \mathcal{L}_\infty$. Since $e_v(t) \in \mathcal{L}_\infty$, (7), (13), (14), (16), (17), and (27) can be used to prove that $m_1(t)$, $L_v(t) \in \mathcal{L}_\infty$. Given that $\dot{p}_{ed}(t)$ is assumed to be bounded function, the expressions in (29) - (33) can be used to conclude that $\dot{\hat{z}}_1^*(t)$, $\dot{\hat{s}}_1(t)$, $v_e(t)$, $\dot{e}_v(t)$, $\dot{e}_\omega(t) \in \mathcal{L}_\infty$. Since $e_\omega(t)$, $e_v(t) \in \mathcal{L}_\in$ and $e_\omega(t)$, $\dot{e}_\omega(t)$, $e_v(t)$, $\dot{e}_v(t) \in \mathcal{L}_\infty$, Barbalat's Lemma [24] can be used to prove the result given in (35). □

VII. Conclusion

In this paper, an adaptive visual servo controller is developed for the fixed camera configuration to enable the end-effector of a robot manipulator to track a desired trajectory determined by an a priori available sequence of images. The controller is formulated using a hybrid composition of image-space pixel information and reconstructed Euclidean information that is obtained via projective homography relationships between the actual image, a reference image, and the desired image. To achieve the objective, a Lyapunov-based adaptive control strategy is employed to actively compensate for the lack of unknown depth measurements and unknown object model parameters.

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*, 2000, pp. 630-635.

[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*, 1997, pp. 45-52.

[3] F. Conticelli and B. Allotta, "Nonlinear Controllability and Stability Analysis of Adaptive Image-Based Systems," *IEEE Transactions on Robotics and Automation*, 17(2), April 2001.

[4] F. Conticelli and B. Allota, "Discrete-Time Robot Visual Feedback in 3-D Positioning Tasks with Depth Adaptation,"

IEEE/ASME Transactions on Mechatronics, 6(3), Sept. 2001, pp. 356-363.

[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*, 2000, pp. 2521-2527.

[6] N. J. Cowan and D. Koditschek, "Planar Image-based Visual Servoing as a Navigation Problem," *Proc. of the IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 1720-1725.

[7] N. J. Cowan, J. D. Weingarten, and D. E. Koditschek, "Visual servoing via Navigation Functions," *IEEE Trans. Robot. Automat.,* Vol. 18, Aug. 2002, pp. 521-533.

[8] K. Deguchi, "Optimal Motion Control for Image-Based Visual Servoing by Decoupling Translation and Rotation," *Proceedings of the Int. Conf. on Intelligent Robots and Systems, Oct. 1998,* pp. 705-711.

[9] M. de Queiroz, D. Dawson, S. Nagarkatti, and F. Zhang, *Lyapunov-based Control of Mechanical Systems*, Birkhauser, New York, 2000.

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

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

[12] Y. Fang, W. E. Dixon, D. M. Dawson, and J. Chen, "Robust 2.5D Visual Servoing for Robot Manipulators," *Proceedings of the 2003 IEEE American Control Conference*, June 2003, pp. 3311-3316.

[13] O. Faugeras, *Three-Dimensional Computer Vision*, The MIT Press, Cambridge Massachusetts, 2001.

[14] O. Faugeras and F. Lustman, "Motion and Structure From Motion in a Piecewise Planar Environment," *International Journal of Pattern Recognition and Artificial Intelligence,* 2(3), 1988, pp. 485-508.

[15] C. A. Felippa, *A Systematic Approach to the Element-Independent Corotational Dynamics of Finite Elements*, Center for Aerospace Structures Document Number CU-CAS-00-03, College of Engineering, University of Colorado, January 2000.

[16] E. Malis, "Contributions à la modélisation et à la commande en asservissement visuel," Ph.D. Dissertation, University of Rennes I, IRISA, France, Nov. 1998.

[17] 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*, 18(2), April 2002, pp. 176-186.

[18] 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*, 37(1), June 2000, pp. 79-97.

[19] E. Malis, F. Chaumette, and S. Bodet, "2 1/2 D Visual Servoing," *IEEE Transactions on Robotics and Automation*, 15(2), April 1999, pp. 238-250.

[20] Y. Mezouar and F. Chaumette, "Path Planning for Robust Image-based Control," *IEEE Transactions on Robotics and Automation*, 18(4), August 2002, pp. 534-549.

[21] E. Rimon and D. E. Koditschek, "Exact Robot Navigation Using Artificial Potential Functions," *IEEE Trans. Robot. Automat.,* Vol. 8, Oct. 1992, pp. 501-518.

[22] A. Ruf and R. Horaud, "Visual Trajectories From Uncalibrated Stereo," *Proc. IEEE Int. Conf. Intelligent Robots Systems*, Sept. 1997, pp. 83-91.

[23] R. Singh, R. M. Voyle, D. Littau, and N. P. Papanikolopoulos, "Alignment of an Eye-In-Hand System to Real Objects using Virtual Images," *Proc. Workshop Robust Vision, Vision-based Control of Motion, IEEE Int. Conf. Robotics Automation*, May 1998.

[24] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, Inc: Englewood Cliff, NJ, 1991.

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

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