

Adaptive 2.5D Visual Servoing of Kinematically Redundant Robot Manipulators^{*}

Y. Fang,¹ A. Behal,¹ W. E. Dixon,² and D. M. Dawson¹

¹Dept. of Electrical and Computer Eng., Clemson University, Clemson, SC 29634-0915

²Eng. Science and Tech. Div. - Robotics, Oak Ridge National Laboratory, 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 Conference on Decision and Control, Dec. 10-13, 2002, Las Vegas, Nevada

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

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

* This research was supported in part by the Eugene P. Wigner Fellowship Program of the Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC, for the U.S. Department of Energy (DOE) under contract DE-AC05-00OR22725 and in part by the U.S. DOE Environmental Management Sciences Program (EMSP) projects ID No. 82794 and ID No. 82797 at ORNL, by ONR Project No. N00014-00-F-0485 at ORNL, and by U.S. NSF Grants DMI-9457967, DMI-9813213, EPS-9630167, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

Adaptive 2.5D Visual Servoing of Kinematically Redundant Robot Manipulators*

Y. Fang[†], A. Behal[†], W. E. Dixon[‡], and D. M. Dawson[†]

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

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

email: yfang, abehal, ddawson@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 3D model of the target object, and (iii) parametric uncertainty in the dynamic model of the robot manipulator. Specifically, by fusing 2D image-space and 3D task-space information (i.e., 2.5D visual servoing) while actively adapting for unknown depth information, a task-space kinematic controller is developed that is proven to ensure asymptotic regulation of the position and orientation of the camera. Based on the desire to enhance the robustness of the control design, the integrator backstepping approach is then utilized to develop a joint torque control input to ensure asymptotic regulation of the position and orientation of the camera, which is held by the end-effector of a kinematically redundant robot manipulator, despite parametric uncertainty in the dynamic model of the robot. The stability of each controller is proven through a Lyapunov-based stability analysis.

1 Introduction

Motivated by the desire to enable robotic systems with a greater sense of perception and ability to operate in unstructured environments, researchers have actively investigated the use of visual servoing control systems (i.e., using information obtained from a camera system to provide position and orientation information about the robot and its environment for use in a control scheme). The results from this research can be broadly divided into Image-Based Visual Servoing (IBVS) and Position-Based Visual Servoing (PBVS) techniques. In PBVS, features are extracted from the camera image and then related to the task-space through the calibrated image Jacobian. The resulting task-space error system is then utilized by the control system. Since the control is calculated based on the task-space error system, inaccuracies in camera calibration will lead to inaccuracies in the 3-Dimensional (3D) task-space reconstruction and ultimately in the task execution. Moreover, since the image-space information is not utilized by the controller, the image features may exit the camera's field-of-view resulting in a loss of stability. In contrast to PBVS, IBVS schemes define an image-space error system that is utilized by the controller. Based on the fact that IBVS controllers servo from the image-space error, this approach ensures that the image will remain in the camera field-of-view and there is conjecture that this approach facilitates some measure of robustness to calibration errors; however, IBVS techniques have problems related to singularities in the image Jacobian, and local minimas may be reached rather than the actual desired position and orientation. In addition to the previous shortcomings, a common characteristic of many PBVS and IBVS techniques is that an accurate 3D model of the environment (or target image) is often required (see [1] for a more in-depth discussion regarding PBVS and IBVS).

Several researchers have recently developed partitioned approaches that exploit a combination of 3D task-space information and 2D image-space information to overcome many of the shortcomings of traditional PBVS and IBVS approaches. For example, in the series of papers by Malis and Chaumette (e.g., [2, 3, 15, 16])

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 3D task-space (obtained either through a given 3D model or more interestingly through a projective Euclidean reconstruction) is utilized to regulate the rotation error system while information from the 2D image-space is utilized to control the translation error system. This control approach incorporates the advantages of both PBVS and IBVS; however, many of the disadvantages of the traditional approaches are avoided [16]: (i) an accurate 3D model of the environment (or target image) is not required, (ii) the image is guaranteed to remain in the camera field-of-view, (iii) local minima can be avoided, and (iv) singularities only exist in the image Jacobian in degenerate cases. In [8], Deguchi describes how an interaction between the translation and rotation of images can result in slower transient performance due to inefficient camera motions used to reach the desired image. Based on this observation, Deguchi then proposes two algorithms to decouple the rotation and translation components using a homography and an epipolar condition. Specifically, Deguchi decomposes the translation and rotation components through a homography and states that the 2.5D controller given in [3] can be utilized, and as an alternate method, Deguchi develops a kinematic controller that utilizes task-space information to regulate the translation error and image-space information to regulate the rotation error. More recently, Corke and Hutchinson [7] developed a new hybrid image-based visual servoing scheme that decouples rotation and translation components about the z-axis from the remaining degrees of freedom to address the so called "Chaumette Conundrum," in which desirable image-space trajectories result in undesirable Cartesian trajectories. One drawback of the controllers given in [2, 3, 7, 8, 14, 15, 16] is that each of the results 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 [16], 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 compensate for the aforementioned depth information, [5] developed an adaptive kinematic controller to ensure uniformly ultimately bounded (UUB) set-point regulation of the image point errors while compensating for the unknown depth information, provided conditions on the translational velocity and the bounds on uncertain depth parameters are satisfied. In [19], 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.

Most control approaches do not account for the inevitable mismatch between the actual and desired camera translation and rotation velocity caused by the nonlinear kinematics and dynamics of the robot manipulator, and hence, reduce the problem to that of kinematic control that simply reacts to image-space errors (e.g., all of the aforementioned research efforts have taken this approach). In [6], Corke and Good presented one of the first results to highlight the advantages of incorporating the robot dynamics in the overall control design. Motivated by the results in [6], several other researchers have proposed vision-based controllers that incorporate the dynamics of the robot. Most of this research has targeted vision-based robotic systems in which the robot is constrained to move in a plane such that the optical axis of the camera remains perpendicular to the robot workspace (e.g., see [9, 12, 20, 21]). Some of the few control designs that take the robot dynamics into account for the 3D visual servoing problem are given in [4, 13].

In this paper, we relate feature points extracted from images taken from the desired and current camera position and orientation through a homography. In a similar manner as in [2, 3, 8, 14, 15, 16], we then decompose the homography into translation and rotation components. Based on this homography decomposition, we then develop a task-space kinematic controller that is inspired by [16]. Specifically, in a similar manner as in [16], the kinematic controller utilizes projected 3D task-space information to regulate the rotation error system and 2D image-space information to regulate the translation error system. Unlike the controller given in [16], the

*This research was supported in part by the Eugene P. Wigner Fellowship Program of the Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC, for the U.S. Department of Energy (DOE) under contract DE-AC05-00OR22725 and in part by the U.S. DOE Environmental Management Sciences Program (EMSP) projects ID No. 82794 and ID No. 82797 at ORNL, by ONR Project No. N00014-00-F-0485 at ORNL, and by U.S. NSF Grant DMI-9457967, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

controllers in this paper do not require an off-line learning phase to determine the unknown distance from the desired camera position to the reference plane. In contrast, we utilize Lyapunov-based control design and stability analysis techniques to develop an adaptive update law that is utilized to compensate for the unknown distance. Motivated by the desire to incorporate the effects of the robot kinematics and dynamics, we utilize the integrator backstepping technique to develop a joint torque control input that ensures regulation of the translation and rotation error systems for the camera-in-hand, while adapting for both the unknown depth information and parametric uncertainty in the dynamic model of a kinematically redundant robot manipulator.

2 Model Development

2.1 Camera Model

To facilitate the development of the camera model, we 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} . As in [16], we also consider a reference plane π that is defined by four target points $O_i \forall i = 1, 2, 3, 4$ (only 3 points are required to define a plane, however, in the subsequent analysis, 4 target points located on the plane π are required) where 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

$$\bar{m}_i = [X_i \ Y_i \ Z_i]^T \quad (1)$$

$$\bar{m}_i^* = [X_i^* \ Y_i^* \ Z_i^*]^T. \quad (2)$$

Since the task-space is projected onto the image-space, we define normalized coordinates, denoted by $m_i(t), m_i^*$, of the targets points $\bar{m}_i(t)$ and \bar{m}_i^* , respectively, 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 we make the standard assumption that $Z_i(t), Z_i^*$ are positive (i.e., the degenerate cases when $Z_i(t), Z_i^* = 0$ are not considered).

In addition to having a task-space coordinate as described previously, each target point will also have a projected pixel coordinate expressed in terms of \mathcal{F} denoted by $u_i(t), v_i(t) \in \mathbb{R}$, which are defined as elements of $p_i(t)$ as follows

$$p_i = [u_i \ v_i \ 1]^T \quad (5)$$

where the projected pixel coordinates of the target points are related to the normalized task-space coordinates by the following global invertible transformation

$$p_i = A m_i \quad (6)$$

where $A \in \mathbb{R}^{3 \times 3}$ is a known, constant, and invertible intrinsic camera calibration matrix that is explicitly defined as [15]

$$A = \begin{bmatrix} f k_u & -f k_u \cot \phi & u_0 \\ 0 & \frac{f k_v}{\sin \phi} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

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), $k_u, k_v \in \mathbb{R}$ represent camera scaling factors, $\phi \in \mathbb{R}$ is the angle between the camera axes, and $f \in \mathbb{R}$ denotes the camera focal length (see Figure 1). Similarly, the constant, desired pixel coordinates expressed in terms of \mathcal{F}^* denoted by $u_i^*, v_i^* \in \mathbb{R}$, are defined as elements of p_i^* as follows

$$p_i^* = [u_i^* \ v_i^* \ 1]^T \quad (8)$$

and can be related to the normalized coordinates m_i^* by the following relationship

$$p_i^* = A m_i^*. \quad (9)$$

Remark 1 *Since the camera is assumed to be calibrated, (i.e., the matrix A defined in (7) is known), $m_i(t)$ of (3), can be computed via the relationship of (6). In a similar manner, m_i^* can be computed by (9).*

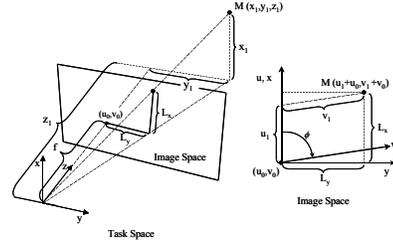


Figure 1: Relationship of the task-space coordinates (x_1, y_1, z_1) to the image-space coordinates $(u_1 + u_0, v_1 + v_0)$ of a task-space point M where (u_0, v_0) denote the image-space coordinate origin and L_x, L_y denote image-space projections for $\phi = 90^\circ$

2.2 Control Objective

The central concept behind 2.5D visual servoing is to utilize 2D information about the environment extracted from the camera image along with estimated 3D information about the environment obtained through an Euclidean reconstruction. In this paper, the Euclidean reconstruction is facilitated by comparing the current camera image to the desired camera image (i.e., solving the motion and structure from motion problem [10]) through a homography matrix [10, 16, 22]. Specifically, a projective homography $G(t) \in \mathbb{R}^{3 \times 3}$ can be utilized to relate the image points $p_i(t)$ of (6) to the image points p_i^* of (9) in the following manner [10]

$$p_i = \alpha_i G p_i^* \quad (10)$$

where $\alpha_i(t) \in \mathbb{R}$ is an unknown scaling factor defined as follows

$$\alpha_i = \frac{Z_i^*}{Z_i} \quad (11)$$

where $Z_i(t)$ and Z_i^* were defined in (1) and (3), respectively. After substituting (6) and (9) into (10) for $p_i(t)$ and p_i^* , respectively, and then premultiplying the resulting expression by A^{-1} , (10) can be rewritten as follows

$$m_i = \alpha_i H m_i^* \quad (12)$$

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

$$H = A^{-1} G A. \quad (13)$$

The Euclidean homography can be computed using (13) where $G(t)$ is determined by utilizing (10) to solve a linear system of equations using 4 pairs of points $(p_i^*, p_i(t))$ on the reference plane π , and A is defined in (7). After normalizing the projective homography $G(t) = [g(t)]_{ij} \forall i, j = 1, 2, 3$ with respect to g_{33} , 8 unknown parameters remain in $G(t)$. To solve for these unknown parameters, 4 sets of image points are required because each set of points represent 3 linear equations, resulting in 12 equations and 12 unknowns (i.e., 8 unknowns from the normalized $G(t)$ ($g_{33} = 1$) and $\alpha_i \forall i = 1, 2, 3, 4$).

By utilizing various techniques (e.g., see [10, 22]), $H(t)$ can be decomposed as follows

$$H = R + x_h n^{*T} \quad (14)$$

where the rotation matrix $R(t) \in \mathbb{R}^{3 \times 3}$, the constant unit normal from \mathcal{F}^* to π denoted by $n^* \in \mathbb{R}^3$ (where n^* is expressed in the \mathcal{F}^* coordinate frame), and the scaled translation vector denoted by $x_h(t) \in \mathbb{R}^3$ are all computed from the decomposition of $H(t)$. 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

$$x_f = x_h d^* \quad (15)$$

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

To quantify the translation mismatch between the actual and desired 3D task-space camera position, we define the translation error $e_v(t) \in \mathbb{R}^3$ as follows

$$e_v = m_e - m_e^* \quad (16)$$

where $m_e(t) \in \mathbb{R}^3$ denotes the extended coordinates [16] of an image point on π in terms of \mathcal{F} and is defined as follows

$$m_e = [m_{e1}(t) \ m_{e2}(t) \ m_{e3}(t)]^T = \begin{bmatrix} \frac{X_1}{Z_1} & \frac{Y_1}{Z_1} & \ln(Z_1) \end{bmatrix}^T \quad (17)$$

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

$$m_e^* = [m_{e1}^* \ m_{e2}^* \ m_{e3}^*]^T = \begin{bmatrix} \frac{X_1^*}{Z_1^*} & \frac{Y_1^*}{Z_1^*} & \ln(Z_1^*) \end{bmatrix}^T \quad (18)$$

where $\ln(\cdot)$ denotes the natural logarithm. Any point O_i on π can be utilized in the control development; however, to reduce the notational complexity, the image point O_1 is selected, and hence, the subscript 1 is utilized in lieu of i .

In addition to forcing the task-space camera translation error to zero, we also want to force the rotation mismatch between \mathcal{F} and \mathcal{F}^* (i.e., $R(t)$ given in (14)) to the identity matrix. To this end, we define a rotation error-like signal $e_\omega(t) \in \mathbb{R}^3$ as follows [16]

$$e_\omega = u\theta \quad (19)$$

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

$$-\pi < \theta(t) < \pi. \quad (20)$$

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

$$R = I_3 + \sin\theta [u]_\times + 2\sin^2\frac{\theta}{2}[u]_\times^2 \quad (21)$$

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

$$[\zeta]_\times \triangleq \begin{bmatrix} 0 & -\zeta_3 & \zeta_2 \\ \zeta_3 & 0 & -\zeta_1 \\ -\zeta_2 & \zeta_1 & 0 \end{bmatrix} \quad \forall \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix}. \quad (22)$$

Based on the previous development, the control objective, stated in mathematical terms, is to ensure regulation of $e_v(t)$ and $e_\omega(t)$ in the sense that

$$\lim_{t \rightarrow \infty} e_v(t), e_\omega(t) = 0. \quad (23)$$

Remark 2 After the stability analysis for Theorem 1, we prove that the results given in (23) ensure that the rotation and translation mismatch between \mathcal{F} and \mathcal{F}^* are regulated as follows

$$\lim_{t \rightarrow \infty} R(t) = I_3, \quad \lim_{t \rightarrow \infty} x_f(t) = 0. \quad (24)$$

Remark 3 To obtain $u(t)$ and $\theta(t)$ from a given rotation matrix $R(t)$, the following expressions can be utilized [11]

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

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

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

Remark 4 Based on (3), (4), (6), (16-18), and the comments in Remark 1, the first 2 elements of $e_v(t)$ are measurable. We can also prove that the third element of $e_v(t)$ is measurable. Based on the fact that $R(t)$ can be computed from (14), we can conclude from (19) and Remark 3 that $e_\omega(t)$ is measurable.

3 Task-Space Kinematic Control Development

3.1 Control Design

Based on the development given in Appendix B (also see [14, 16]), the open-loop error dynamics for $e_v(t)$ and $e_\omega(t)$ can be expressed as follows

$$\dot{e}_v = \frac{1}{d^* \gamma_2} L_v v_c + L_{(v,\omega)} \omega_c \quad (27)$$

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

where $L_v(t), L_{(v,\omega)}(t), L_\omega(t) \in \mathbb{R}^{3 \times 3}$ are measurable matrices that are defined as follows

$$L_v = \begin{bmatrix} -1 & 0 & m_{e1} \\ 0 & -1 & m_{e2} \\ 0 & 0 & -1 \end{bmatrix} \quad (29)$$

$$L_{(v,\omega)} = \begin{bmatrix} m_{e1}m_{e2} & -1 - m_{e1}^2 & m_{e2} \\ 1 + m_{e2}^2 & -m_{e1}m_{e2} & -m_{e1} \\ -m_{e2} & m_{e1} & 0 \end{bmatrix}. \quad (30)$$

$$L_\omega = I_3 - \frac{\theta}{2} [u]_\times + \left(1 - \frac{\text{sinc}(\theta)}{\text{sinc}^2\left(\frac{\theta}{2}\right)} \right) [u]_\times^2 \quad (31)$$

where

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

Given the open-loop error dynamics in (27), we design the task-space control input $v_c(t)$ as follows

$$v_c = -\gamma_2 L_v^{-1} \left[T_v e_v + \hat{d}^* L_{(v,\omega)} \omega_c \right] \quad (32)$$

where $L_v^{-1}(t)$ is given by the following expression

$$L_v^{-1} = \begin{bmatrix} -1 & 0 & -m_{e1} \\ 0 & -1 & -m_{e2} \\ 0 & 0 & -1 \end{bmatrix} \quad (33)$$

$T_v \in \mathbb{R}^{3 \times 3}$ is a constant, diagonal, positive gain matrix, and $\hat{d}^*(t) \in \mathbb{R}$ is a dynamic estimate for d^* of (15) that is generated by the following differential update law

$$\dot{\hat{d}}^* = k_0 e_v^T L_{(v,\omega)} \omega_c \quad (34)$$

where $k_0 \in \mathbb{R}$ is a positive gain constant. After substituting (32) into (27) for $v_c(t)$, the following closed-loop dynamics for $e_v(t)$ are obtained

$$\dot{e}_v = -\frac{1}{d^*} T_v e_v + \frac{\tilde{d}^*}{d^*} L_{(v,\omega)} \omega_c \quad (35)$$

where $\tilde{d}^*(t) \in \mathbb{R}$ is a parameter estimate error signal defined as

$$\tilde{d}^* = d^* - \hat{d}^*. \quad (36)$$

Based on the open-loop dynamics given in (28), we design the task-space control input $\omega_c(t)$ as follows

$$\omega_c = -T_\omega e_\omega \quad (37)$$

where $T_\omega \in \mathbb{R}^{3 \times 3}$ is a constant, diagonal, positive gain matrix. After substituting (37) into (28) for $\omega_c(t)$, the following closed-loop dynamics for $e_\omega(t)$ are obtained

$$\dot{e}_\omega = -L_\omega T_\omega e_\omega. \quad (38)$$

3.2 Stability Analysis

Theorem 1 The kinematic control input given in (32), (34), and (37) ensures that $e_v(t)$ and $e_\omega(t)$, defined in (16) and (19), respectively, are asymptotically regulated in the sense that

$$\lim_{t \rightarrow \infty} e_v(t), e_\omega(t) = 0. \quad (39)$$

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

$$V = \frac{1}{2} e_v^T e_v + \frac{1}{2} e_\omega^T e_\omega + \frac{1}{2d^*k_0} \tilde{d}^{*2}. \quad (40)$$

After taking the time derivative of (40) and then substituting (35) and (38) into the resulting expression for $\dot{e}_v(t)$ and $\dot{e}_\omega(t)$, respectively, we obtain the following expression

$$\dot{V} = e_v^T \left(-\frac{1}{d^*} T_v e_v + \frac{\tilde{d}^*}{d^*} L_{(v,\omega)} \omega_c \right) - e_\omega^T T_\omega e_\omega - \frac{1}{d^*k_0} \tilde{d}^* \dot{\tilde{d}}^* \quad (41)$$

where the following facts have been utilized

$$\dot{\tilde{d}}^* = -\dot{\hat{d}}^*, \quad e_\omega^T L_\omega = e_\omega^T. \quad (42)$$

After substituting (34) into (41) for $\hat{d}^*(t)$ and canceling common terms, the following expression is obtained

$$\dot{V} = -\frac{1}{d^*} e_v^T T_v e_v - e_\omega^T T_\omega e_\omega \leq 0; \quad (43)$$

hence, from (40) and (43), we can prove that $e_v(t), e_\omega(t), \hat{d}^*(t) \in \mathcal{L}_\infty$ and that $e_v(t), e_\omega(t) \in \mathcal{L}_2$. Since $\hat{d}^*(t) \in \mathcal{L}_\infty$, it is clear from (36) that $\tilde{d}^*(t) \in \mathcal{L}_\infty$. Based on the fact that $e_v(t), e_\omega(t) \in \mathcal{L}_\infty$, we can utilize (16), (19), (20), (29), (30), (31), and (37) to prove that $m_{e1}(t), m_{e2}(t), L_v(t), L_{(v,\omega)}(t), L_\omega(t), \omega_c(t), u(t) \in \mathcal{L}_\infty$. From the previous facts, we can now utilize (32), (34), (35), and (38) to prove that $v_c(t), \dot{\hat{d}}^*(t), \dot{e}_v(t), \dot{e}_\omega(t) \in \mathcal{L}_\infty$. Since $e_v(t), e_\omega(t) \in \mathcal{L}_\infty \cap \mathcal{L}_2$ and $\dot{e}_v(t), \dot{e}_\omega(t) \in \mathcal{L}_\infty$, we can now employ a corollary to Barbalat's Lemma [17] to conclude (39).

Remark 5 Given (23), we can utilize (19), (21), and (23) to prove that the rotation between \mathcal{F} and \mathcal{F}^* asymptotically approaches the identity matrix as shown below

$$\lim_{t \rightarrow \infty} R(t) = I_3. \quad (44)$$

From (16-18) and (23), we can prove that

$$\lim_{t \rightarrow \infty} \ln \left(\frac{Z_1(t)}{Z_1^*} \right) = 0, \quad (45)$$

and hence, from (11) and (45), we can prove that

$$\lim_{t \rightarrow \infty} \alpha_1(t) = 1. \quad (46)$$

Based on (44) and (46), we can substitute (14) into (12) to obtain the following expression

$$\lim_{t \rightarrow \infty} m_1(t) = \left(I + x_h(t) n^{*T} \right) m_1^* \quad (47)$$

From (1), (4), (12), (16-18), and (23), we can also prove that

$$\lim_{t \rightarrow \infty} m_1(t) = m_1^*; \quad (48)$$

hence, it is clear from (47) that

$$\lim_{t \rightarrow \infty} x_h(t) \left[n^{*T} m_1^* \right] = 0. \quad (49)$$

Property 1: The symmetric and positive-definite inertia matrix $M(q)$ satisfies the following inequalities

$$m_1 \|\xi\|^2 \leq \xi^T M(q) \xi \leq m_2 \|\xi\|^2 \quad \forall \xi \in \mathbb{R}^n \quad (69)$$

where m_1, m_2 are positive constants, and $\|\cdot\|$ denotes the standard Euclidean norm.

Property 2: The inertia and centripetal-Coriolis matrices satisfy the following skew symmetric relationship

$$\xi^T \left(\frac{1}{2} \dot{M}(q) - V_m(q, \dot{q}) \right) \xi = 0 \quad \forall \xi \in \mathbb{R}^n. \quad (70)$$

Property 3: The centripetal-Coriolis matrix satisfies the following relationship

$$V_m(q, \nu) \zeta = V_m(q, \zeta) \nu \quad \forall \zeta, \nu \in \mathbb{R}^n. \quad (71)$$

Property 4: The dynamics given in (68) can be linearly parameterized as shown below

$$M(q)\ddot{q} + V_m(q, \dot{q})\dot{q} + G(q) + F(\dot{q}) = W(q, \dot{q}, \ddot{q}) \psi \quad (72)$$

where $\psi \in \mathbb{R}^r$ contains the constant system parameters, and the regression matrix $W(\cdot) \in \mathbb{R}^{n \times r}$ contains known functions dependent on the signals $q(t), \dot{q}(t)$, and $\ddot{q}(t)$ (it is assumed that if the arguments of $W(\cdot)$ are bounded then $W(\cdot)$ is bounded).

Remark 7 *Since we are only concerned with revolute robot manipulators, we know that kinematic and dynamic terms denoted by $M(q), G(q), J(q)$, and $J^+(q)$ are bounded for all possible $q(t)$ (i.e., these kinematic and dynamic terms only depend on $q(t)$ as arguments of trigonometric functions). Likewise, provided $\dot{q}(t)$ is bounded, $V_m(q, \dot{q}), J(q, \dot{q})$, and $J^+(q, \dot{q})$ will be bounded.*

4.2 Backstepping Control Design

To facilitate the subsequent control development and stability analysis, we define the following composite translation and rotation error signal $e(t) \in \mathbb{R}^6$ as follows

$$e = \begin{bmatrix} e_v^T & e_\omega^T \end{bmatrix}^T. \quad (73)$$

Motivated by the desire to compensate for the mismatch between the actual and desired camera translation and rotation velocities, we add and subtract terms to the open-loop error dynamics for $e_v(t)$ and $e_\omega(t)$, developed in (27) and (28), respectively, to separate the mismatch terms from the desired velocity terms as follows

$$\dot{e} = \begin{bmatrix} \frac{1}{d^* \gamma_2} L_v v_{cd} + L_{(v, \omega)} \omega_{cd} \\ L_\omega \omega_{cd} \end{bmatrix} + \chi \quad (74)$$

where $v_{cd}(t), \omega_{cd}(t) \in \mathbb{R}^3$ denote the desired camera translation and rotation velocities, respectively, $\chi(t) \in \mathbb{R}^6$ denotes the velocity mismatch vector defined as follows

$$\chi = \begin{bmatrix} \frac{1}{d^* \gamma_2} L_v (v_c - v_{cd}) + L_{(v, \omega)} (\omega_c - \omega_{cd}) \\ L_\omega (\omega_c - \omega_{cd}) \end{bmatrix}. \quad (75)$$

Based on the results from Theorem 1, we design $v_{cd}(t), \omega_{cd}(t)$ as follows

$$v_{cd} = -\gamma_2 L_v^{-1} \left[T_v e_v + \hat{d}^* L_{(v, \omega)} \omega_{cd} \right] \quad (76)$$

$$\omega_{cd} = -T_\omega e_\omega \quad (77)$$

where $\hat{d}^*(t)$ is redesigned as follows

$$\dot{\hat{d}}^* = k_0 e_v^T L_{(v, \omega)} \omega_{cd} \quad (78)$$

and the rest of the signals of (76) and (77) are defined in the same way as in (32) and (37). After substituting (76) and (77) into (74) and exploiting the relationship given in (60), we can rewrite (74) and (75) as follows

$$\dot{e} = \begin{bmatrix} -\frac{1}{d^*} T_v e_v + \frac{\hat{d}^*}{d^*} L_{(v, \omega)} \omega_{cd} \\ -L_\omega T_\omega e_\omega \end{bmatrix} + \chi \quad (79)$$

where

$$\chi = \begin{bmatrix} \frac{1}{d^* \gamma_2} L_v \bar{R}^T (v_e - v_{cd}) + L_{(v, \omega)} \bar{R}^T (\omega_e - \omega_{cd}) \\ L_\omega \bar{R}^T (\omega_e - \omega_{cd}) \end{bmatrix} \quad (80)$$

and $v_{ed}(t), \omega_{ed}(t) \in \mathbb{R}^3$ denote the desired end-effector translation and rotation velocities, respectively, defined as follows

$$v_{fd} = \begin{bmatrix} v_{ed} \\ \omega_{ed} \end{bmatrix} = \begin{bmatrix} \bar{R} & 0_{3 \times 3} \\ 0_{3 \times 3} & \bar{R} \end{bmatrix} \begin{bmatrix} v_{cd} \\ \omega_{cd} \end{bmatrix} \quad (81)$$

where the notation $0_{i \times i}$ denotes an $i \times i$ matrix where each element is 0. Motivated by the desire to express (79) and (80) in terms of joint velocities, we utilize (60) and (62) to rewrite (80) as follows

$$\chi = -(\Pi_{m1} + \frac{1}{d^*} \Pi_{m2}) J \rho \quad (82)$$

where the auxiliary control terms $\Pi_{m1}(t), \Pi_{m2}(t) \in \mathbb{R}^{6 \times 6}$, and $\rho(t) \in \mathbb{R}^n$ are defined as follows

$$\Pi_{m1} = \begin{bmatrix} 0_{3 \times 3} & L_{(v, \omega)} \bar{R}^T \\ 0_{3 \times 3} & L_\omega \bar{R}^T \end{bmatrix}, \Pi_{m2} = \begin{bmatrix} \frac{1}{\gamma_2} L_v \bar{R}^T & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix} \quad (83)$$

$$\rho = J^+ v_{fd} + (I_n - J^+ J) g - \dot{q} \quad (84)$$

where $g(t) \in \mathbb{R}^n$ is an auxiliary signal that is constructed according to a sub-task control objective (e.g., joint-limit avoidance, obstacle avoidance), $v_{fd}(t)$ was defined in (81), and $J^+(t)$ was defined in (64). The subsequent stability analysis mandates that the sub-task control objective be formulated in such a manner that both $g(t)$ and $\dot{q}(t)$ are bounded signals.

By determining the open-loop dynamics for $\rho(t)$, we can now incorporate the robot joint torque control input. To this end, we take the time derivative of (84), pre-multiply the resulting expression by $M(q)$, and then substitute (68) for $M(q)\ddot{q}(t)$ to obtain the following expression

$$M\dot{\rho} = -V_m \rho + Y \phi + \frac{1}{d^*} J^T \Pi_{m2}^T e - \tau \quad (85)$$

where the term $\frac{J^T(t) \Pi_{m2}^T(t) e(t)}{d^*}$ has been added and subtracted to the right-hand side of (85), and the linear parameterization $Y(q, \dot{q}, \ddot{q}, g, v_f) \phi$ is defined as follows

$$Y \phi = M \left(J^+ (v_{fd} - Jg) + J^+ (\dot{v}_{fd} - \dot{J}g) + (I_n - J^+ J) \dot{q} \right) + V_m(q, \dot{q}) (J^+ v_{fd} + (I_n - J^+ J) g) + G(q) + F(\dot{q}) - \frac{1}{d^*} J^T \Pi_{m2}^T e \quad (86)$$

where $Y(q, \dot{q}, \ddot{q}, g, v_f) \in \mathbb{R}^{n \times r}$ denotes a known regression matrix, and $\phi \in \mathbb{R}^r$ denotes the constant unknown system parameters (e.g., mass, inertia, friction coefficients, desired distance of end-effector to the image plane π). The linear parameterization given in (86) is achieved by utilizing Property 4 and the fact that $\dot{v}_{fd}(t)$ can be linearly parameterized as follows

$$\dot{v}_{fd} = \frac{1}{d^*} \Omega_a + \Omega_b \quad (87)$$

where $\Omega_a(t), \Omega_b(t) \in \mathbb{R}^6$ are auxiliary measurable signals (available upon request).

Based on the error system development given in (85) and the subsequent stability analysis, we design the control torque input $\tau(t)$ as follows

$$\tau = Y \hat{\phi} + K \rho - J^T \Pi_{m1}^T e \quad (88)$$

where $K \in \mathbb{R}^{n \times n}$ is a constant, positive definite, diagonal gain matrix, and $\hat{\phi}(t) \in \mathbb{R}^r$ is a dynamic estimate for ϕ of (86) that is generated by the following differential update law

$$\dot{\hat{\phi}} = \Gamma_\phi Y^T \rho \quad (89)$$

where $\Gamma_\phi \in \mathbb{R}^{r \times r}$ is a constant, positive definite, diagonal gain matrix. After substituting (88) into (85), the following closed-loop dynamics for $\rho(t)$ are obtained

$$M\dot{\rho} = -V_m \rho + Y \hat{\phi} + J^T (\Pi_{m1} + \frac{1}{d^*} \Pi_{m2})^T e - K \rho \quad (90)$$

where $\tilde{\phi}(t) \in \mathbb{R}^r$ denotes the parameter estimation error defined as follows

$$\tilde{\phi} = \phi - \hat{\phi}. \quad (91)$$

4.3 Stability Analysis

Theorem 2 *The control law given by (76), (77), (78), (88) and (89) ensures that the 3D task-space position and orientation of the camera are asymptotically regulated to the desired position and orientation in the sense that*

$$\lim_{t \rightarrow \infty} e(t), \rho(t) = 0. \quad (92)$$

Proof: To prove Theorem 2, we define a non-negative function denoted by $V(\tilde{t}) \in \mathbb{R}$ as follows

$$V = \frac{1}{2} e^T e + \frac{1}{2k_0} \frac{\tilde{d}^{*2}}{d^*} + \frac{1}{2} \rho^T M \rho + \frac{1}{2} \tilde{\phi}^T \Gamma_\phi^{-1} \tilde{\phi}. \quad (93)$$

After taking the time derivative of (93), and then utilizing (70), (78), (79), (90), and the fact that

$$\dot{\tilde{\phi}}(t) = -\dot{\hat{\phi}}(t) \quad (94)$$

the following expression is obtained

$$\begin{aligned}
V = & e^T \begin{bmatrix} -\frac{1}{d^*} T_v e_v + \frac{\dot{d}^*}{d^*} L_{(v,\omega)} \omega_{cd} \\ -L_\omega T_\omega e_\omega \end{bmatrix} \\
& -e^T (\Pi_{m1} + \frac{1}{d^*} \Pi_{m2}) J \rho - \frac{\dot{d}^*}{d^*} \left(e_v^T L_{(v,\omega)} \omega_{cd} \right) \\
& + \rho^T \left(Y \tilde{\phi} + J^T (\Pi_{m1} + \frac{1}{d^*} \Pi_{m2})^T e - K \rho \right) - \tilde{\phi}^T Y^T \rho.
\end{aligned} \tag{95}$$

After cancelling the interconnection terms, we can rewrite (95) as follows

$$\dot{V} = -\frac{1}{d^*} e_v^T T_v e_v - e_\omega^T T_\omega e_\omega - \rho^T K \rho \tag{96}$$

where the fact that

$$e_\omega^T L_\omega T_\omega e_\omega = e_\omega^T T_\omega e_\omega \tag{97}$$

was utilized. From the structure of (93) and (96), we can prove that $e(t)$, $\dot{d}^*(t)$, $\rho(t)$, and $\hat{\phi}(t) \in \mathcal{L}_\infty$ and that $e(t)$, $\rho(t) \in \mathcal{L}_2$. By utilizing (36), (76), (77), (81), (83), (91) and similar arguments as in the proof for Theorem 1, we can also prove that $m_{e1}(t)$, $m_{e2}(t)$, $L_v(t)$, $L_{(v,\omega)}(t)$, $L_\omega(t)$, $v_{cd}(t)$, $\omega_{cd}(t)$, $\Pi_{m1}(t)$, $\Pi_{m2}(t)$, $\dot{d}^*(t)$, $v_{fd}(t)$, $\hat{\phi}(t) \in \mathcal{L}_\infty$. Given the previous boundedness arguments, (78), (79), (82), (84), and (89) can now be utilized along with the assumption that $J(q)$, $g(t) \in \mathcal{L}_\infty$ to prove that $\dot{d}(t)$, $\dot{e}(t)$, $\dot{q}(t)$, $\dot{\hat{\phi}}(t)$, $\dot{\rho}(t) \in \mathcal{L}_\infty$. Based on (86-90), the comments provided in Remark 7, and the assumption that $\dot{g}(t) \in \mathcal{L}_\infty$, we can now prove that $Y(q, \dot{q}, \dot{g}, g, v_f)$, $\dot{\rho}(t)$, $\tau(q) \in \mathcal{L}_\infty$; hence, all the closed-loop signals remain bounded. Since $e(t)$, $\rho(t) \in \mathcal{L}_\infty \cap \mathcal{L}_2$ and $\dot{e}(t)$, $\dot{\rho}(t) \in \mathcal{L}_\infty$, we can now employ a corollary to Barbalat's Lemma [17] to conclude the result given in (92).

5 Discussion and Conclusions

In this paper, a kinematic visual servoing controller is developed that ensures asymptotic regulation of the camera translation and rotation error systems while simultaneously compensating for uncertainty in the distance from the desired camera position to the stationary target plane. Specifically, by decomposing the homography into separate translation and rotation components, we were able to exploit both 2D image-space and projected 3D task-space (i.e., 2.5D visual servoing) information to construct the kinematic controller. Based on the desire to enhance the robustness of the control design, the integrator backstepping approach was utilized to incorporate the robot kinematic and dynamic models. Specifically, a joint torque control input was developed to ensure asymptotic regulation of the position and orientation of the camera held by the robot end-effector (camera-in-hand problem) of a kinematically redundant robot manipulator, despite parametric uncertainty in the dynamic model of the robot. The stability of each controller is proven through Lyapunov-based stability analysis and the performance of the controller is demonstrated through simulation results.

In the same spirit as in [8], we could utilize the development given in this paper to utilize task-space information to regulate the translation error system and image-space information to regulate the rotation error system; however, unlike the approach given in [8], we can adaptively compensate for the depth information. One of the motivating factors for developing our approach based on the work of [16], rather than [8], is that [16] only requires the controller to servo on one image point. That is, the fundamental difference in the work by [16] and [8] is that since [8] utilizes the image-space information to regulate the rotation error system, multiple image points are required to prove the stability result (i.e., rotation cannot be determined from a single point). In contrast, since [16] utilizes the image-space information to regulate the translation error system, only one image point is required in the feedback-loop.

Although the controller developed in this paper compensates for unknown depth information and parametric uncertainty in the dynamic model of the robot manipulator, uncertainty related to the camera calibration (i.e., extrinsic and intrinsic camera parameters) is not addressed. Several researchers have experimentally investigated the robustness of proposed visual servoing algorithms to calibration uncertainty; however, very few analytical treatments of the subject have been developed (with the exception of research that targets the special case in which the camera's optical axis is perpendicular to a fixed plane of the task-space, such as the fixed-camera configuration or the camera-in-hand problem when the robot is constrained to planar motion). One result that examines the effect of calibration uncertainty on the visual servoing algorithm is provided in [16]; however, the result is theoretically unsatisfying. Specifically, upon careful examination of the estimated homography matrix (i.e., Eq. (27) of [16]), it can be determined that there are more unknowns than equations; hence, additional points must be defined on the reference plane to increase the number of available equations. Unfortunately, even by defining more points on the reference plane, analytical techniques such as [10] cannot be used to determine a solution for the nonlinear system of equations (and hence, decompose the homography). Numerical techniques

could be utilized to solve the system of nonlinear equations; however, it is not guaranteed that the resulting solution would have any meaningful significance. In essence, this approach resorts back to numerically estimating the calibration parameters on-line, rather than embedding the uncertainty in the kinematic model to facilitate adaptive/robust control techniques.

References

- [1] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing", D. Kriegman, G. Hagar, and A. Morse, editors, *The Confluence of Vision and Control*, Vol. 237 of *LNCIS Series*, pp. 66-78, Springer Verlag, 1998.
- [2] 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 Int. Conf. on Rob. and Auto.*, pp. 630-635, 2000.
- [3] 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.
- [4] D.-M. Chuang, S.-C. Wu, and M.-K. Hor, "Adaptive Fuzzy Visual Servoing in Robot Control", *Proc. of the IEEE Int. Conf. on Rob. and Auto.*, pp. 811-816, 1997.
- [5] F. Conticelli and B. Allotta, "Nonlinear Controllability and Stability Analysis of Adaptive image-Based Systems", *IEEE Trans. on Rob. and Auto.*, Vol. 17, No. 2, April 2001.
- [6] P. I. Corke and M. C. Good, "Dynamic Effects in Visual Closed-Loop Systems", *IEEE Trans. on Rob. and Auto.*, Vol. 12, No. 5, pp. 671-683, October 1996.
- [7] P. I. Corke and S. A. Hutchinson, "A New Hybrid Image-Based Visual Servo Control Scheme", *Proc. of the IEEE Conf. on Decision and Control*, pp. 2521-2527, 2000.
- [8] K. Deguchi, "Optimal Motion Control for Image-Based Visual Servoing by Decoupling Translation and Rotation", *Proceedings of the Intl. Conf. on Intelligent Rob. and Systems*, B.C. Canada, pp. 705-711, Oct. 1998.
- [9] W. E. Dixon, D. M. Dawson, E. Zergeroglu, and A. Behal, "Adaptive Tracking Control of a Wheeled Mobile Robot via an Uncalibrated Camera System", *Proc. of the IEEE American Control Conf.*, Chicago, Illinois, June, 2000, pp. 1493-1497.
- [10] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment", *Int. J. of Pattern Recognition and Artificial Intelligence*, Vol. 2, No. 3, pp. 485-508, 1988.
- [11] 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, University of Colorado, Jan. 2000.
- [12] R. Kelly, "Robust Asymptotically Stable Visual Servoing of Planar Robots", *IEEE Trans. Rob. and Auto.*, Vol. 12, No. 5, pp. 759-766, Oct. 1996.
- [13] R. Kelly, R. Carelli, O. Nasisi, B. Kuchen, F. Reyes, "Stable Visual Servoing of Camera-in-Hand Robotic Systems", *IEEE Trans. on Mech.*, Vol. 5, No. 1, pp. 39-48, March 2000.
- [14] E. Malis, "Contributions à la modélisation et à la commande en asservissement visuel", Ph.D. Dissertation, University of Rennes I, IRISA, France, Nov. 1998.
- [15] E. Malis and F. Chaumette, "2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement", *Int. J. of Comp. Vision*, pp. 79-97, Vol. 37, No. 1, June 2000.
- [16] E. Malis, F. Chaumette, and S. Bodet, "2 1/2 D Visual Servoing", *IEEE Trans. on Rob. and Auto.*, pp. 238-250, Vol. 15, No. 2, April 1999.
- [17] J. J. Slotine and W. Li, *Applied Nonlinear Control*, Englewood Cliff, NJ: Prentice Hall, Inc., 1991.
- [18] M. Takegaki and S. Arimoto, "A new feedback method for dynamic control of manipulators", *Transactions of ASME J. of Dyn. Sys., Meas., and Control*, Vol. 103, pp. 119-125, June 1981.
- [19] C. J. Taylor and J. P. Ostrowski, "Robust Vision-based Pose Control", *Proc. of the IEEE Int. Conf. on Rob. and Auto.*, pp. 2734-2740, 2000.
- [20] E. Zergeroglu, D. M. Dawson, M. de Queiroz, and A. Behal, "Vision-based Nonlinear Tracking Controllers in the Presence of Parametric Uncertainty", *Proc. of the IEEE/ASME Int. Conf. on Advanced Intelligent Mech.*, pp. 854 - 859, 1999.
- [21] E. Zergeroglu, D. M. Dawson, M. de Queiroz, and S. Nagarkatti, "Robust Visual-Servo Control of Planar Robot Manipulators in the Presence of Uncertainty", *Proc. of the 38th IEEE Conf. on Decision and Control*, pp. 4137-4142, 1999.
- [22] Z. Zhang and A. R. Hanson, "Scaled Euclidean 3D Reconstruction Based on Externally Uncalibrated Cameras", *IEEE Symp. on Computer Vision*, 1995.