Teach by Zooming: A Camera Independent Alternative to Teach by Showing Visual Servo Control*

W. E. Dixon

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 International Conference on Intelligent Robots and Systems (IROS 2003)
Las Vegas, NV, October 27 - 31, 2003

Keywords: Multiple Camera Servoing, Teach-by-Zooming, and Uncalibrated Camera

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.

Teach by Zooming: A Camera Independent Alternative to Teach By Showing Visual Servo Control*

W. E. Dixon

Engineering Science and Technology Division - Robotics and Energetic Machines Group

Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6305

email: dixonwe@ornl.gov

Abstract

Typical visual servoing control objectives are formulated by the desire to position/orient a camera based on a reference image obtained by a priori positioning the camera in the desired location. In this paper, projective geometric techniques are used to formulate a visual servo control problem based on a cooperative camera scheme that provides an alternative to the typical "teach by showing" methodology. Specifically, by using a second camera with zoom capabilities, the proposed "teach by zooming" alternative approach eliminates the need for the camera to be a priori positioned in the desired location. To achieve this result, development is provided that illustrates how cameras with different intrinsic parameters can be used (i.e., camera independent). Control objectives are developed for cases when the intrinsic camera calibration matrices for the camera-in-hand and fixed camera are both known or both unknown.

I. Introduction

With recent advances in camera technology, computer vision, and control theory, visual servo control systems exhibit significant promise to enable autonomous systems with the ability to operate in unstructured environments. Although a vision system can provide a unique sense of perception, several technical issues have impacted the design of robust visual servo controllers. One of these issues includes the camera configuration (pixel resolution versus field-of-view). For example, for vision systems that utilize a camera in a fixed configuration (i.e., the eye-to-hand configuration), the camera is typically mounted at a large enough distance to ensure that the desired target objects will remain in the camera's view. Unfortunately, by mounting the camera in this configuration the task-space area that corresponds to a pixel in the image-space can be quite large, resulting in low resolution and noisy position measurements. Moreover, many applications are ill-suited for the fixed camera configuration. For example, a robot may be required to position the camera for close-up tasks (i.e., the camera-in-hand configuration). For the camera-in-hand configuration, the camera is naturally close to the workspace, providing for higher resolution measurements due to the fact that each pixel represents a smaller task-space area; however, the field-of-view of the camera is significantly reduced (i.e., an object may be located in the workspace but be out of the camera's view).

Several results have been recently developed to address the camera configuration issues by utilizing a cooperative camera strategy. The advantages of the cooperative camera configuration are that the fixed camera can be mounted so that the complete workspace is visible and the camera-in-hand provides a high resolution, close-up view of an object (e.g., facilitating the potential for more precise motion for robotic applications). Specifically, [7] made the first steps towards cooperatively utilizing global and local information obtained from a fixed camera and a camera-in-hand, respectively; unfortunately, to prove the stability results, the translation and rotation tasks of the controller were treated separately (i.e., the coupling terms were ignored) and the cameras were considered to be calibrated. In [2] and [3], a new cooperative visual servoing approach was developed and experimentally demonstrated to use information from both an uncalibrated fixed camera and an uncalibrated camera-in-hand to enable robust tracking by the camera-in-hand of an object moving in the task-space with an unknown trajectory. Unfortunately, a crucial assumption for this approach is that the camera and the object motion be constrained to a plane so that the unknown depth from the camera to the target remains constant. The approaches in [2], [3], and [7] are in contrast to typical multi-camera approaches, that utilize a stereo-based configuration, since stereo-based approaches typically do not simultaneously exploit local and global views of the robot and the workspace.

Another issue that has impacted the development of visual servo systems is the calibration of the camera. For example, to relate pixelized image-space information to the task-space, exact knowledge of the camera calibration parameters is required, and discrepancies in the calibration matrix result in an erroneous relationship. Furthermore, an acquired image is a function of both the task-space position of the camera and the intrinsic calibration parame-

*This research was supported in part by a subcontract to Oak Ridge National Laboratory by the Florida Department of Citrus through the University of Florida, and by the U.S. DOE Environmental Management Sciences Program (EMSP) project ID No. 82797 sponsored by the DOE Office of Biological and Environmental Research (OBER) of the DOE Office of Science (SC).

ters; hence, perfect knowledge of the camera intrinsic parameters is also required to relate the relative position of a camera through the respective images as it moves. For example, the typical visual servoing problem is constructed as a "teach by showing" (TBS) problem in which a camera is positioned at a desired location, a reference image is taken (where the normalized task-space location is determined via the intrinsic calibration parameters), the camera is moved away from the reference location, and then repositioned at the reference location under visual servo control (which requires that the calibration parameters did not change in order to reposition the camera to the same task-space location given the same image). See [1], [8], [15], and [11] for a further explanation and an overview of this problem formulation.

Unfortunately, for many practical applications it may not be possible to TBS (i.e., it may not be possible to acquire the reference image by a priori positioning the camera-in-hand to the desired location). As stated in [11], the TBS problem formulation is "camera-dependent" due to the hypothesis that the camera intrinsic parameters during the teaching stage, must be equal to the intrinsic parameters during servoing. Motivated by the desire to address this problem, the basic idea of the pioneering development in [10] and [11] is to use projective invariance to construct an error function that is invariant to the intrinsic parameters, thus enabling the control objective to be met despite variations in the intrinsic parameters. However, the fundamental idea for the problem formulation is that an error system be constructed in an invariant space, and unfortunately, as stated in [10] and [11] several control issues and a rigorous stability analysis of the invariant space approach "have been left unresolved."

Inspired by the previous issues and the development in [3], [10], and [11], this paper utilizes a cooperative camera scheme to formulate a visual control problem that does not rely on the TBS paradigm. Specifically, a calibrated camera-in-hand and a calibrated fixed camera are used to facilitate a "teach by zooming" approach to formulate a control problem with the objective to force the camera-in-hand to the exactly known Euclidean position/orientation of a virtual camera defined by a zoomed image from a fixed camera (i.e., the camera-in-hand does not need to be positioned in the desired location a priori using this alternative approach). Since the intrinsic camera calibration parameters may be difficult to exactly determine in practice, a second control objective is defined as the desire to servo an uncalibrated camera-in-hand so that the respective image corresponds to the zoomed image of an uncalibrated fixed camera. For each case, the fixed camera provides a global view of a scene that can be zoomed to provide a close-up view of the target. The reference image is then used to servo the camera-in-hand. That is, the proposed "teach by zooming" approach addresses the camera configuration issues by exploiting the wide scene view of a fixed camera and the close-up view provided by the zoom capability and the camera-in-hand.

II. Problem Formulation

To formulate an alternative to the TBS approach, the subsequent development will exploit projective geometry to relate the various camera positions. Based on the developed relationships, a Euclidean homography is developed to perform the image reconstruction to facilitate the control objective.

A. Camera Relationships

To facilitate the subsequent development, consider three orthogonal coordinate systems, denoted by $\mathcal{F}$, $\mathcal{F}_f$, and $\mathcal{F}^*$, where $\mathcal{F}$ is attached to a camera that is held by a robot end-effector, $\mathcal{F}_f$ is attached to a camera fixed in the workspace, and $\mathcal{F}^*$ is a coordinate system attached to a virtual camera that represents the constant, desired position and orientation of $\mathcal{F}$. The virtual camera coordinate system location is determined by effectively scaling the fixed camera depth along the focal axis (i.e., zooming on the target points); hence, $\mathcal{F}_f$ and $\mathcal{F}^*$ have the same rotation with respect to $\mathcal{F}$ as depicted in Fig. 1. As in [14], a reference plane π is defined by four[1] coplanar target points O_i $\forall i = 1, 2, 3, 4$ that are not collinear. Each target point on π can be projected onto the coordinate system of each camera. Based on this projection, normalized task-space coordinates, denoted by $m_i(t)$, m_{fi}, and $m_i^* \in \mathbb{R}^3$, are defined as follows

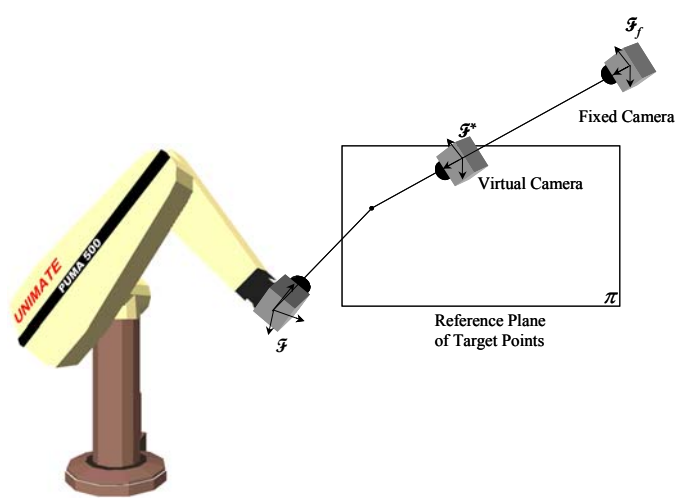

Fig. 1. Camera relationships.

$$
\begin{aligned}
m_i &\triangleq \begin{bmatrix} \frac{x_i}{z_i} & \frac{y_i}{z_i} & 1 \end{bmatrix}^T \\
m_{fi} &\triangleq \begin{bmatrix} \frac{x_{fi}}{z_{fi}} & \frac{y_{fi}}{z_{fi}} & 1 \end{bmatrix}^T \\
m_i^* &\triangleq \begin{bmatrix} \frac{x_{fi}}{z_i^*} & \frac{y_{fi}}{z_i^*} & 1 \end{bmatrix}^T
\end{aligned} \tag{1}
$$

[1]In general, only 3 points are required to define a plane, however, in the subsequent analysis, 4 coplanar target points located on the plane π are assumed to be available. The coplanar assumption is not necessary since 8 points could be used for the development if the points are not coplanar.

where the standard assumption is made that the distances from the origin of the respective coordinate frames to the reference plane along the focal axis remains positive (i.e., $z_i(t), z_{fi}, z_i^* > \varepsilon$ where ε is an arbitrarily small positive constant). From (1), it is clear that the normalized Euclidean coordinates of m_{fi} can be related to m_i^* as follows

$$m_{fi} = diag\{\frac{z_i^*}{z_{fi}}, \frac{z_i^*}{z_{fi}}, 1\} m_i^* \tag{2}$$

where diag$\{\cdot\}$ denotes a diagonal matrix with the elements of $\{\cdot\}$ along the diagonal.

In addition to having a normalized 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 is defined as elements of $p_i(t) \in \mathbb{R}^3$ as follows

$$p_i \triangleq \begin{bmatrix} u_i & v_i & 1 \end{bmatrix}^T . \tag{3}$$

The projected pixel coordinates $p_i(t)$ are related to the normalized task-space coordinates $m_i(t)$ by the following global invertible transformation (i.e., the pinhole lens model)

$$p_i = A m_i \tag{4}$$

where $A \in \mathbb{R}^{3\times 3}$ is a constant, and invertible intrinsic camera calibration matrix that is explicitly defined as [9]

$$A \triangleq \begin{bmatrix} \alpha & -\alpha \cot\phi & u_0 \\ 0 & \dfrac{\beta}{\sin\phi} & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5}$$

In (5), 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), α, $\beta \in \mathbb{R}$ represent the product of the camera scaling factors and the focal length, and $\phi \in \mathbb{R}$ is the skew angle between the camera axes. The constant, pixel coordinates, expressed in terms of $\mathcal{F}_f$ and denoted by $u_{fi}, v_{fi} \in \mathbb{R}$, acquired from the fixed camera before zooming on the target are defined as elements of $p_{fi} \in \mathbb{R}^3$ as follows

$$p_{fi} \triangleq \begin{bmatrix} u_{fi} & v_{fi} & 1 \end{bmatrix}^T . \tag{6}$$

The projected pixel coordinates defined in (6) are related to the normalized task-space coordinates $m_{fi}(t)$ by the following global invertible transformation

$$p_{fi} = A_f m_{fi} \tag{7}$$

where $A_f \in \mathbb{R}^{3\times 3}$ is a constant, and invertible intrinsic camera calibration matrix for the fixed camera defined as follows

$$A_f \triangleq \begin{bmatrix} \alpha_f & -\alpha_f \cot\phi_f & u_{0f} \\ 0 & \dfrac{\beta_f}{\sin\phi_f} & v_{0f} \\ 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

After the fixed camera is zoomed to acquire the reference image, the constant, pixel coordinates, expressed in terms of $\mathcal{F}^*$ and denoted by $u_i^*, v_i^* \in \mathbb{R}$, are defined as elements of $p_i^* \in \mathbb{R}^3$ as follows

$$p_i^* \triangleq \begin{bmatrix} u_i^* & v_i^* & 1 \end{bmatrix}^T , \tag{9}$$

and the new unknown camera calibration matrix, denoted by $A^* \in \mathbb{R}^{3\times 3}$, is defined as

$$A^* \triangleq \begin{bmatrix} \alpha^* & -\alpha^* \cot\phi_f & u_{0f} \\ 0 & \dfrac{\beta^*}{\sin\phi_f} & v_{0f} \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

A.1 Calibrated Case

Since an acquired image is a function of both the intrinsic calibration parameters and the Euclidean position of the camera, the pixel coordinates of the image acquired after zooming the fixed camera can be either related to the normalized task-space coordinates of m_{fi} or m_i^* as follows

$$p_i^* = A^* m_{fi} \qquad p_i^* = A_f m_i^* . \tag{11}$$

If the calibration matrix A_f is assumed to be known then m_{fi} can be computed from (7) and m_i^* can be computed from (11). If m_{fi} can be computed, then the first linear equation in (11) can be used to determine A^*. That is, the first relationship in (11) represents two equations with six unknowns (i.e., α^*, β^*, $\sin\phi_f$, $\cot\phi_f$, u_{0f}, v_{0f}); hence, three image points can be used to generate six equations. Moreover, if m_{fi} and m_i^* are known then two equations with two unknowns can be generated from (2) to determine z_{fi} and z_i^*; hence, the exact Euclidean coordinate of the target expressed in $\mathcal{F}_f$ and $\mathcal{F}^*$ can be determined. That is, the fixed camera can be zoomed such that the target points will have known Euclidean coordinates expressed in $\mathcal{F}_f$ and $\mathcal{F}^*$.

To force the camera-in-hand to the Euclidean position defined by $\mathcal{F}^*$, the subsequent development is based on comparisons between the pixel coordinates $p_i(t)$ and p_i^* with the control objective formulated in terms of the virtual camera position m_i^* rather than in terms of the normalized fixed camera position m_{fi}. When the intrinsic calibration matrix for both cameras are known, the control objective can be formulated in terms of the Euclidean position. However, since the intrinsic calibration matrices are different, the images will not be the same at the goal position. That is, for the calibrated camera case, the objective is to force $m_i(t)$ to m_i^*. From (4) and (11), it can be determined that when $m_i(t)$ equals m_i^* then

$$p_i = A A_f^{-1} p^*. \tag{12}$$

Remark 1: For the typical TBS approach $A = A_f$. From (12), it is clear that when the cameras are calibrated the control objective can be formulated in terms of either the Euclidean position or the image correspondence under the TBS assumption.

A.2 Uncalibrated Case

If the intrinsic calibration matrix A_f is not known then m_{fi} can not be determined and A^* cannot be computed from (11). Moreover, since m_{fi} and m_i^* are not known then (2) can not be used to determine z_{fi} and z_i^*; hence, the exact Euclidean coordinate of the target expressed in $\mathcal{F}_f$ and $\mathcal{F}^*$ can not be determined using the previous methods that rely on perfect knowledge of the intrinsic parameters. Hence, the objective must be formulated differently in the presence of uncertainty in the intrinsic calibration matrix for either camera. Specifically, in contrast to the calibrated case in which the objective is to force the Euclidean positions to align with non-matching images, for the uncalibrated case, the objective is to force the images to correspond with misaligned Euclidean positions. That is, (4) and (11) can be used to determine that if the images correspond then the Euclidean coordinates will differ according to the following expression

$$m_i = m_{di} \triangleq A^{-1} A_f m_i^* \tag{13}$$

where the new normalized Euclidean position $m_{di} \in \mathbb{R}^3$ is defined as the Euclidean mismatch. Hence, the control objective for the uncalibrated case can be formulated as the desire to force $m_i(t)$ to m_{di}.

Remark 2: From (13), it can also be determined that even in the uncalibrated case the control objective can be formulated in terms of either the Euclidean position or the image correspondence under the TBS assumption. However, from (12) and (13), it is clear that the camera calibration plays a significant role in formulating the control objective for the multi-camera TBS alternative approach developed in this paper.

B. Euclidean Reconstruction

B.1 Calibrated Case

To develop the Euclidean reconstruction, the image from the camera-in-hand can be compared to the image from the fixed camera after the zoom task. Specifically, by relating (4) with the second expression in (11), the following projective homography can be formulated

$$p_i = \lambda_i G p_i^* \ . \tag{14}$$

In (14) the ratio $\lambda_i(t) \in \mathbb{R}$ is defined as follows

$$\lambda_i \triangleq \frac{z_i^*}{z_i} \tag{15}$$

and $G(t) \in \mathbb{R}^{3\times 3}$ denotes a projective homography. From the relationship in (14), a set of 12 linear equations given by the 4 image point pairs $(p_i, p_i^*(t))$ with 3 equations per image pair can be used to determine the projective homography up to a scalar multiple. After substituting (4) and the second expression in (11) into (14), the normalized Euclidean coordinates can be related through a Euclidean homography, denoted by $H(t) \in \mathbb{R}^{3\times 3}$, as follows

$$m_i = \lambda_i \underbrace{A^{-1} G A_f}_{H} m_i^* \tag{16}$$

where $H(t)$ defines a rotation and scaled translation between $\mathcal{F}$ and $\mathcal{F}^*$ as follows [12] (see Fig. 2)

$$H = R + x_h n^{*T}. \tag{17}$$

In (17), $R(t) \in SO(3)$ denotes the rotation between $\mathcal{F}$ and $\mathcal{F}^*$, $n^* \in \mathbb{R}^3$ denotes the constant unit normal from $\mathcal{F}^*$ to π, and $x_h(t) \in \mathbb{R}^3$ denotes a translation vector from $\mathcal{F}$ to $\mathcal{F}^*$ that is scaled by the depth related parameter $d^* \in \mathbb{R}$ (see Fig. 2). If the intrinsic calibration matrices A and A_f are known (and hence, $m_i(t)$ and m_i^* are known), then various techniques can be used (e.g., see [6], [16]) to decompose the Euclidean homography to obtain $\lambda_i(t)$, $G(t)$, $H(t)$, and the rotation and translation signals $R(t)$, $x_h(t)$, and n^*.

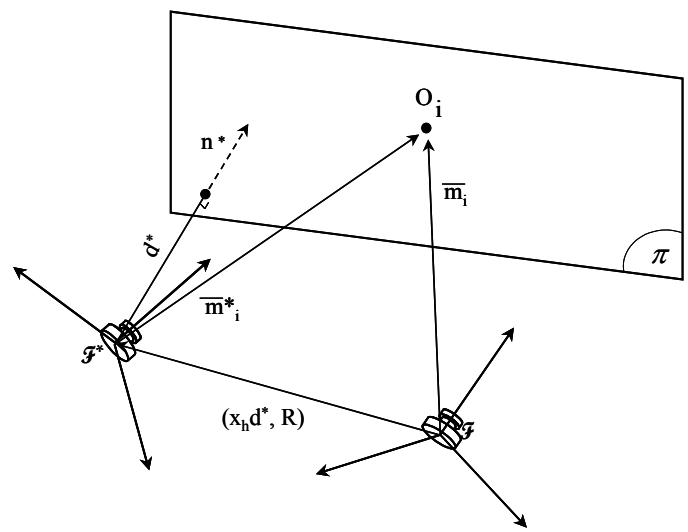

Fig. 2. Coordinate frame relationships.

B.2 Uncalibrated Case

If the intrinsic camera calibration matrices A and A_f are difficult to exactly determine, estimates of the elements of the matrices and the normalized coordinates can be utilized. As stated previously, the control objective for the uncalibrated case can be stated as the desire to force $m_i(t)$ to m_{di} introduced in (13). Motivated by the desire to recast the problem in terms of this new objective, the normalized coordinates are related through the new Euclidean homography, denoted by $H_d(t) \in \mathbb{R}^{3\times 3}$, as follows

$$m_i = \lambda_i \underbrace{A^{-1} G A}_{H_d} m_{di} \tag{18}$$

where (13) and (16) were utilized. Based on (4), (11), (18), and the subsequent development, estimates for the normalized coordinates, denoted by $\hat{m}_i(t)$, $\hat{m}_{di} \in \mathbb{R}^3$, can be generated as follows [12]

$$\hat{m}_i \triangleq \hat{A}^{-1} p_i = \tilde{A} m_i \qquad \hat{m}_{di} \triangleq \hat{A}^{-1} p_i^* = \tilde{A} m_{di} \tag{19}$$

where $\hat{A} \in \mathbb{R}^{3\times 3}$ denotes a best-guess estimate of the intrinsic camera calibration matrix A, and the calibration

error matrix $\tilde{A} \in \mathbb{R}^{3\times 3}$ is defined as follows [12]

$$\tilde{A} \triangleq \hat{A}^{-1}A = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} \\ 0 & \tilde{A}_{22} & \tilde{A}_{23} \\ 0 & 0 & 1 \end{bmatrix} \tag{20}$$

where $\tilde{A}_{11}$, $\tilde{A}_{12}$, $\tilde{A}_{13}$, $\tilde{A}_{22}$, $\tilde{A}_{23} \in \mathbb{R}$ denote unknown intrinsic calibration mismatch constants. By substituting (19) into (18) the following relationship can be obtained

$$\hat{m}_i = \lambda_i \hat{H}_d \hat{m}_{di} \tag{21}$$

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

$$\hat{H}_d \triangleq \tilde{A} H_d \tilde{A}^{-1}. \tag{22}$$

Since $\hat{m}_i(t)$ and $\hat{m}_{di}$ can be determined from (19), a set of 12 linear equations can be developed from the 4 image point pairs, and (21) can be used to solve for $\hat{H}_d(t)$. As stated in [12], provided additional information is available (e.g., at least 4 vanishing points), various techniques (e.g., see [6], [16]) can be used to decompose $\hat{H}_d(t)$ to obtain the estimated rotation and translation components as follows

$$\hat{H}_d = \hat{R} + \hat{x}_h \hat{n}^{*T} \tag{23}$$

where $\hat{R}(t) \in \mathbb{R}^{3\times 3}$ is related to $R(t)$ as follows

$$\hat{R} \triangleq \tilde{A} R \tilde{A}^{-1}, \tag{24}$$

and $\hat{x}_h(t) \in \mathbb{R}^3$, $\hat{n}^{*T} \in \mathbb{R}^3$ denote the estimate of $x_h(t)$ and n^*, respectively, and are defined as follows

$$\hat{x}_h \triangleq \gamma \tilde{A} x_h \tag{25}$$

$$\hat{n}^* \triangleq \frac{1}{\gamma} \tilde{A}^{-T} n^* \tag{26}$$

where $\gamma \in \mathbb{R}$ denotes the following positive constant

$$\gamma \triangleq \left\| \tilde{A}^{-T} n^* \right\|. \tag{27}$$

C. Control Objective and Design

C.1 Calibrated Case

The control objective for the case when the camera intrinsic calibration matrix is known can now be stated as the desire to ensure that the Euclidean coordinates of $\mathcal{F}$ are regulated to the desired and exactly known Euclidean coordinates $\mathcal{F}^*$. Based on the development given in (16), it can be shown that the control objective is achieved if the Euclidean homography $H(t)$ approaches the identity matrix and either $\lambda_i(t) \to 1$ or $m_i(t) \to m_i^*$. Mathematically, it can be shown that if $R(t) \to I_3$, and the scaled translation error $x_h(t) \to 0$, then $H(t) \to I_3$. Alternatively, it can be shown that if $R(t) \to I_3$, $\lambda_i(t) \to 1$ and $m_i(t) \to m_i^*$, then $x_h(t) \to 0$ and the control objective is achieved.

C.2 Uncalibrated Case

For the case when the intrinsic calibration matrices are not exactly known, then the control objective is to force the image from the camera-in-hand to correspond to the zoomed image of the fixed camera, or stated in terms of the normalized Euclidean coordinates, the objective is to force $m_i(t) \to m_{di}$. From (21), it can be shown that the control objective is achieved if the estimated Euclidean homography $\hat{H}_d(t)$ approaches the identity matrix and either $\lambda_i(t) \to 1$ or $\hat{m}_i(t) \to \hat{m}_{di}$. Mathematically, it can be shown that if $\hat{R}(t) \to I_3$, and the estimated scaled translation error $\hat{x}_h(t) \to 0$, then $\hat{H}(t) \to I_3$. Alternatively, it can be shown that if $\hat{R}(t) \to I_3$, $\lambda_i(t) \to 1$ and $\hat{m}_i(t) \to \hat{m}_{di}$, then $\hat{x}_h(t) \to 0$ and the control objective is achieved.

For examples of how controllers can be designed to achieve this objective see [4] and [12]. Specifically, in [4] and [12], a class of controllers are developed (under the TBS assumption) that prove asymptotic and exponential regulation, respectively, by proving that $\hat{R}(t) \to I_3$, $\alpha_i(t) \to 1$ and $\hat{m}_i(t) \to \hat{m}_i^*$. Hence, the development in this paper could be used to eliminate the TBS assumption for results such as in [4] and [12].

III. Conclusions

In this paper, an alternative approach to the TBS paradigm is provided. By examining the relationships between the coordinate systems and exploiting the use of projective geometry a mechanism is formalized that enables a different camera (e.g., a fixed camera with zoom capability) to be used to define a reference image for the visual servo control problem. The advantage of this scheme is that it enables visual servo control where the TBS mechanism is not possible and it also provides a method to use multiple cameras in a configuration to allow for a large view of the task-space while servoing using a higher resolution view. An interesting observation is made that when the intrinsic calibration matrices are known, then the control objective can be defined as the desire to servo the camera-in-hand Euclidean position to the known Euclidean coordinates of a virtual camera defined by a zoomed image. However, when the control objective is achieved, the images will not match (i.e., they will differ by a ratio of the intrinsic parameters). If the intrinsic parameters are not known exactly, then the Euclidean coordinates are unknown and the control objective is formulated on the desire to force the images to correspond and the Euclidean coordinates will differ by a ratio of the intrinsic calibration parameters. Subsequent efforts will demonstrate how the approach can be extended to encompass tracking results in which the fixed camera is used to record a desired time-varying trajectory for the camera-in-hand to track.

References

[1] P. I. Corke, "Visual Control of Robot Manipulators - A Review," *Visual Servoing: Real Time Control of Robot Manipulators Based on Visual Sensory Feedback*, K. Hashimoto (ed.), *World Scientific Series in Robotics and Automated Systems*, Vol. 7, World Scientific Press, Singapore, 1993.

[2] W. E. Dixon and L. J. Love, "Lyapunov-based Visual Servo Control for Robotic Deactivation and Decommissioning," *The 9th Biennial ANS International Spectrum Conference*, Reno, Nevada, August 2002.

[3] W. E. Dixon, E. Zergeroglu, Y. Fang, and D. M. Dawson, "Object Tracking by a Robot Manipulator: A Robust Cooperative Visual Servoing Approach," *Proceedings of the IEEE International Conference on Robotics and Automation*, Washington, DC, May 2002, pp. 211-216.

[4] Y. Fang, W. E. Dixon, D. M. Dawson, and J. Chen, "An Exponential Class of Model-Free Visual Servoing Controllers in the Presence of Uncertain Camera Calibration," *Proceedings of the 2003 IEEE Conference on Decision and Control*, accepted, to appear.

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

[6] 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).

[7] G. Flandin, F. Chaumette, and E. Marchand, "Eye-in-hand/Eye-to-hand Cooperation for Visual Servoing," *Proc. of the International Conference on Robotics and Automation*, pp. 2741-2746, 2000.

[8] S. Hutchinson, G. D. Hager, and P. I. Corke, "A Tutorial On Visual Servo Control," *IEEE Trans. Robotics and Automation*, Vol. 12, No. 5, pp. 651-670 (1996).

[9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, New York, NY: Cambridge University Press, 2000.

[10] E. Malis, "Vision-Based Control Using Different Cameras for Learning the Reference Image and for Servoing," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots Systems*, Hawaii, November 2001, pp. 1428-1433.

[11] E. Malis, "Visual servoing invariant to changes in camera intrinsic parameters," *Proceedings of the International Conference on Computer Vision*, Vancouver, Canada, July 2001, pp. 704-709.

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

[13] 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).

[14] E. Malis, F. Chaumette, and S. Bodet, "2 1/2 D Visual Servoing," *IEEE Transactions on Robotics and Automation*, Vol. 15, No. 2, pp. 238-250 (1999).

[15] *Visual Servoing: Real Time Control of Robot Manipulators Based on Visual Sensory Feedback*, K. Hashimoto (ed.), *World Scientific Series in Robotics and Automated Systems*, Vol. 7, World Scientific Press, Singapore, 1993.

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