Robust Vision-based Pose Control

Camillo J. Taylor James P. Ostrowski
General Robotics and Active Sensory Perception (GRASP) Laboratory
University of Pennsylvania, 3401 Walnut Street, Philadelphia, PA 19104-6228
E-mail: {cjtaylor, jpo}@grip.cis.upenn.edu

Abstract

In this paper the problem of controlling the spatial position and orientation of a robotic platform based on the image data obtained from a video camera mounted on that platform is considered. More specifically, we propose control laws that will cause the robot to achieve and maintain a fixed position and orientation with respect to a set of feature points in the scene.

We demonstrate analytically that the proposed control scheme is globally convergent even in the presence of large calibration errors in both the intrinsic parameters of the camera and in the extrinsic parameters which relate the frame of reference of the camera to the body frame of the robot platform which is being controlled. Furthermore no a'priori knowledge about the structure of the scene is assumed.

1 Introduction

The use of camera-based techniques to control robotic systems has seen a significant rise in popularity recently. This field, known as *visual servoing*, has been aided by the faster speeds and lower costs of modern microprocessors, coupled with the general availability of high quality cameras. The wealth of information available in image data and the flexibility of the sensor combine to make it an attractive option for a control input if the right algorithms can be developed.

Using monocular systems Papanikolopoulos, Nelson, and Khosla [16] have developed a system that can track full 3-D motions by utilizing a Jacobian-based adaptive controller that estimates the depth parameters online. More recent work by Papanikolopoulos has extended these results using optical flow to estimate depth parameters for use in uncalibrated environments [15]. Espiau, Rives, et al. [3, 17, 18] proposed approaches that also utilize the image Jacobian. They use the *interaction screw* to describe the relationship between robot motion and image feature motion. They have implemented these systems, for example, in controlling mobile robots based on landmarks of known geometries. These techniques, however, also require estimation of the depth of features in the scene.

For stereo camera systems Hager, Chang and Morse [6], Hollinghurst and Cipolla [10] and Horaud, Dornaika and Espiau [11] have all investigated the issues involved in controlling robotic manipulators based on the image data acquired with uncalibrated or coarsely calibrated rigs. This work is important in the current context, since it established that image-based techniques can have a proven insensitivity to calibration parameters [7].

In this paper we consider the problem of controlling the position and orientation of a robotic platform based on image data obtained from an onboard camera. The control laws that are proposed generate translational and angular velocities that drive the robot to a fixed position and orientation with respect to a set of features in the scene. We make the assumption that the features are derived from rigid objects, such as walls, tables, and other fixtures. Thus, they can move with respect to the robot, but not with respect to each other.

The proposed control schemes make use of well established techniques for computing estimates for the relative orientation of two camera positions from a set of feature correspondences. As in other visual servoing schemes, the goal position of the platform is actually specified indirectly in terms of the image measurements at the desired pose. The use of relative orientation for pose control has also been proposed by Basri, Rivlin and Shimshoni [1] Malis, Chaumette and Boudet [13, 14], Soatto and Perona [19] and Deguchi [2]. Malis, Chaumette and Boudet and have also considered the issue of robustness to intrinsic calibration parameters. This work improves on those results by formally demonstrating robustness to *both* intrinsic and extrinsic calibration errors. Unlike other approaches our analysis proceeds by proposing a Lya-

punov function over the pose parameters as opposed to analyzing the behavior of a task function defined in terms of the image measurements. Among other advantages, this analysis leads to a simpler characterization of the types of calibration errors that can be tolerated by the control scheme.

In previous work [20] a pair of control laws which separately regulated the translational and rotational velocities of a platform were analyzed. It was shown that, when used in isolation, these controllers were able to regulate the translational and rotational offsets respectively to zero even in the presence of large calibration errors. This paper extends those results by demonstrating analytically that a combined controller which regulates both the translational and angular velocities of the platform *simultaneously* is, in fact, stable in the presence of calibration errors and will drive the robot to the desired configuration.

It is important to note that the control laws presented here differ from previous work in that they do not require the estimation of an image Jacobian or interaction screw, used to relate the robot's body velocity to changes in image measurements. Since this image Jacobian depends upon the positions of the features with respect to the camera, control schemes which rely on this approach must either estimate the Jacobian online or assume that this matrix is nominally fixed at some known value. No such estimates or assumptions are required in this method.

The techniques described in this paper are applicable to a wide range of positioning tasks. They could be used on mobile platforms such as blimps, helicopters, underwater vehicles, space based robots or terrestrial vehicles to guide the robot to a desired position and orientation with respect to one or more target objects. The schemes could also be used for so-called "eye in hand" servoing applications to position the end effector of a robot arm with respect to a work piece based on the image data obtained from a camera mounted on the robot's gripper.

In Section 2 the relative orientation problem is briefly discussed. Section 3 describes the consequences of calibration errors in the intrinsic and extrinsic calibration parameters. Section 4 presents an analysis of the pose control problem, proposes a control scheme for solving this task and provides conditions under which this strategy is known to be convergent. The experimental results that have been obtained with this method on a six degree of freedom robot arm are presented in Section 5 while Section 6 presents some of the conclusions that have been drawn from this research.

2 Relative Orientation

The problem of computing the relative position of two cameras from a set of point correspondences in two images has been well studied in both the computer vision and photogrammetry literatures. Most of the approaches to solving this problem proceed by exploiting the epipolar constraint, which relates the position of the projection of a point feature in one image to its projection in the second image.

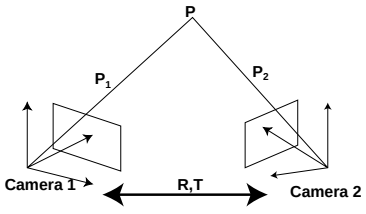

Figure 1: The basic geometry of the relative orientation problem

The basic imaging geometry of the relative orientation problem is shown in Figure 1. Let $\mathbf{P}_1$ denote the coordinates of a point $\mathbf{P}$ in the scene with respect to a coordinate frame of reference centered at the focus of projection of the first camera. The coordinates of that same point $\mathbf{P}$ with respect to a coordinate frame associated with the second camera are given by the expression $\mathbf{P}_2 = R^T(\mathbf{P}_1 - \mathbf{T})$ where $R \in SO(3)$ and $\mathbf{T} \in \Re^3$ denote the relative orientation and position of the second frame with respect to the first. The rotation matrix R can also be expressed in terms of an axis of rotation $\mu \in \Re^3, \|\mu\| = 1$ and an angle θ as follows $R = \exp(\theta J(\mu))$ where

$$J(\mu) = \begin{pmatrix} 0 & -\mu_z & \mu_y \\ \mu_z & 0 & -\mu_x \\ -\mu_y & \mu_x & 0 \end{pmatrix}. \quad (1)$$

Assuming a perspective projection model, the relationship between the coordinates of a point with respect to the camera frame $\mathbf{P} = (X, Y, Z)^T$ and the projective coordinates of the image of that point on the focal plane, $\mathbf{p} = (u, v, 1)^T$ can be written as follows:

$$\mathbf{p} \propto A\mathbf{P}, \quad (2)$$

where $A \in \Re^{3\times 3}$ is an invertible upper triangular matrix that represents the *intrinsic parameters* of the camera.

In our application, camera frames 1 and 2 refer to two different positions of the same camera and we will assume that the intrinsic parameters of the camera

are fixed, hence A is a constant matrix. This means that the projective coordinates of the projection of the point into the two images, $\mathbf{p}_1$ and $\mathbf{p}_2$, are given by the following equations:

$$\mathbf{p}_1 \propto A\mathbf{P}_1 \quad (3)$$

$$\mathbf{p}_2 \propto A\mathbf{P}_2 = AR^T(\mathbf{P}_1 - \mathbf{T}) \quad (4)$$

From these two equations one can derive the *epipolar constraint* which is given by

$$\mathbf{p}_1^T A^{-T} J(\mathbf{T}) R A^{-1} \mathbf{p}_2 = 0, \quad (5)$$

The matrix $F \propto A^{-T}J(\mathbf{T})RA^{-1}$ is termed the *fundamental matrix* [4]. A number of algorithms have been proposed for estimating this fundamental matrix from a set of point correspondences [12, 8, 9].

Note that the fundamental matrix F can also be rewritten in terms of a matrix H_∞ and a vector e. That is:

$$F \propto J(e)H_\infty \quad (6)$$

where $H_\infty \propto ARA^{-1}$ is refered to as the collineation of the plane at infinity [5] and the vector $e \propto A\mathbf{T}$ denotes the homogeneous coordinates of the epipole in the image.

The matrix H_∞ cannot, in general, be recovered solely from point correspondences. It can, however, be estimated if additional information, such as the correspondences between vanishing points in the two images, can be recovered. In the sequel, we will assume that sufficient information is available in the image measurements to produce estimates for both H_∞ and e.

This estimate for H_∞ can be used to discount the effects of camera rotation on the measurements for the locations of feature points in the image. That is, if (u_i, v_i) denote the coordinates of a point feature in the image then the location where that feature would have appeared if the camera had translated to its current position without rotating, $(\acute{u}_i, \acute{v}_i)$, is given by the following expression.

$$\begin{pmatrix} \acute{u}_i \\ \acute{v}_i \\ 1 \end{pmatrix} \propto H_\infty \begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix} \quad (7)$$

By measuring the disparity between a vector of these corrected image measurements, $\acute{\mathbf{y}} = (\acute{u}_1, \acute{v}_1, \acute{u}_2, \acute{v}_2, ...)^T$, and the value of this vector at the target position $\mathbf{y_0}$ we obtain an error function, $d(\mathbf{T}) = \|\acute{\mathbf{y}}(\mathbf{T}) - \mathbf{y_0}\|$. This function depends only on the translational disparity, $\mathbf{T}$. In the sequel this disparity function will be used to scale the translational velocity control signal. It is assumed that the features being tracked are not in a degenerate configuration so that the disparity function will evaluate to zero only at $\mathbf{T} = \mathbf{0}$.

From the matrix H_∞ we can recover an estimate for the rotational displacement between the two camera frames by noting that this matrix takes the form of a rotation matrix under a similarity transform. Since similarity transformations preserve eigenvalues it follows that after an appropriate scale factor is applied, the eigenvalues of H_∞ should be 1, $e^{i\theta}$ and $e^{-i\theta}$, and its trace should be $1 + 2\cos\theta$ (recall that $R = \exp(\theta J(\mu))$). Furthermore the eigenvector of H_∞ corresponding to the unit eigenvalue will be $\eta = A\mu$.

If the intrinsic parameter matrix of the camera, A, is known then it is a simple matter to compute estimates for the axis of rotation, μ, and the direction of translation, $\mathbf{T}$, from η and e respectively. Note that it is not possible to recover the magnitude of the translation vector. In practise, however, our estimates for the intrinsic parameters of the camera will contain errors. The relationship between the estimate for the intrinsic parameters, A_{est}, and the actual values, A, can be expressed as follows:

$$A = A_{est}\tilde{A} \quad (8)$$

where $\tilde{A}$ is an unknown invertible upper triangular matrix.

This implies that we will only be able to estimate the axis of rotation and direction of translation up to this unknown transformation $\tilde{A}$. That is, by applying A_{est}^{-1} to η and e we can estimate the unit vectors, $\frac{\bar{A}\mu}{\|\bar{A}\mu\|}$ and $\frac{\bar{A}T}{\|\bar{A}T\|}$.

2.1 Controlling Pose

The objective of the proposed visual servoing scheme is to drive the disparity between the robot's current position and desired configuration to zero. As shown in Figure 1 this disparity is characterized by two components, the rotational disparity, $R \in SO(3)$, and the translational disparity, $\mathbf{T} \in \Re^3$. Similarly, the *body velocity* of the platform, that is, its instantaneous velocity with respect to its current pose, can be characterized in terms of its angular velocity $\omega \in \Re^3$ and its translational velocity $\mathbf{v} \in \Re^3$.

In order to accomplish this positioning task based on the estimates that can be obtained from the image measurements we propose the following control laws for regulating the body velocity of the platform (see [20] for further motivation of these control laws).

$$\omega_r = -\sin\theta \frac{\tilde{A}\mu}{\|\tilde{A}\mu\|} \tag{9}$$

$$\mathbf{v}_r = -d(\mathbf{T}) \frac{\tilde{A}\mathbf{T}}{\|\tilde{A}\mathbf{T}\|} \tag{10}$$

In practise the relationship between the camera frame and the robot's body frame can be difficult to measure accurately. In this case there will be some discrepancy between the frame in which the body velocities, ω_r and $\mathbf{v}_r$, are applied and the camera frame as shown in Figure 2. Let these unknown rotational and translational displacements be denoted by $\tilde{R}$ and $\tilde{\mathbf{T}}$, respectively, which we will term the *extrinsic calibration errors.*

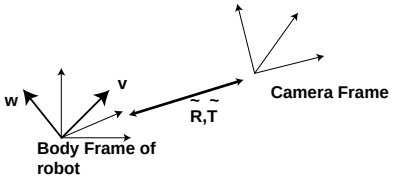

Figure 2: This figure shows the relationship between the cameras frame of reference and the body frame of the platform to which the rigid body velocities, $\mathbf{v}_r$ and ω_r, are applied

In this case the body velocity of the camera, (ω_c, v_c), and the commanded body velocity of the robot, (ω_r, v_r), will be related by an adjoint transformation as shown in Equation (11).

$$\begin{pmatrix} \omega_c \\ v_c \end{pmatrix} = \begin{pmatrix} \tilde{R} & 0 \\ J(\tilde{\mathbf{T}})\tilde{R} & \tilde{R} \end{pmatrix} \begin{pmatrix} \omega_r \\ v_r \end{pmatrix} \tag{11}$$

Notice that the presence of the $J(\tilde{\mathbf{T}})\tilde{R}$ term in the adjoint transformation points to a coupling between the rotational and translational velocities of the camera which complicates the stability analysis. If $\tilde{\mathbf{T}}$ were zero then it would be possible to decouple the pose control problem into separate rotational and translational control tasks.

2.2 Proof of Asymptotic Stability

Having characterized both the intrinsic and extrinsic calibration errors we can now state our main result.

Theorem 1 *If the matrix $\tilde{R}\tilde{A}$ is positive definite then the control laws given in equations 9 and 10 will regulate the platform to the desired pose asymptotically.*

Proof: In order to prove that the proposed control laws lead to an asymptotically stable system we consider the Lyapunov function $\mathcal{L}(R, \mathbf{T})$ given in equation 12

$$\begin{aligned} \mathcal{L}(R,\mathbf{T}) &= \|R - I\|^2 + \gamma\|\mathbf{T}\|^2 \\ &= tr((R-I)^T(R-I)) + \gamma\|\mathbf{T}\|^2 \\ &= 2(3 - tr(R)) + \gamma\|\mathbf{T}\|^2 \end{aligned} \tag{12}$$

where γ is a positive constant and $\|R - I\|$ refers to the Frobenius norm of the matrix.

We will investigate the behavior of this function on the domain $\{R, \mathbf{T} | R = \exp(\theta\mu), \theta < \pi, \|\mathbf{T}\| < \delta\}$. This set includes practically all displacements from the target position since the scalar δ can be set to any positive value without affecting the correctness of the following proof.

Taking the derivative of this Lyapunov function with respect to time yields the following expression:

$$\begin{aligned} \dot{\mathcal{L}}(R,\mathbf{T}) &= 2tr(RJ(\omega_c)) + 2\gamma\mathbf{T}^T v_c \\ &= 2(\sin\theta\mu^T\omega_c + \gamma\mathbf{T}^T v_c) \\ &= 2(\sin\theta\mu^T\ \gamma\mathbf{T}^T) \begin{pmatrix} \omega_c \\ v_c \end{pmatrix} \\ &= 2(\sin\theta\mu^T\ \gamma\mathbf{T}^T) \begin{pmatrix} \tilde{R} & 0 \\ J(\tilde{\mathbf{T}})\tilde{R} & \tilde{R} \end{pmatrix} \begin{pmatrix} \omega_r \\ v_r \end{pmatrix} \\ &= -2\Big(\frac{\sin^2\theta\mu^T\tilde{R}\tilde{A}\mu}{\|\tilde{A}\mu\|} + \frac{\gamma\sin\theta\mathbf{T}^T J(\tilde{\mathbf{T}})\tilde{R}\tilde{A}\mu}{\|\tilde{A}\mu\|} \\ &\quad + \frac{\gamma d(\mathbf{T})\mathbf{T}^T\tilde{R}\tilde{A}\mathbf{T}}{\|\tilde{A}\mathbf{T}\|}\Big) \end{aligned}$$

If we let a represent the smallest eigenvalue of the symmetric part of $\tilde{R}\tilde{A}$, b denote the largest singular value of the matrix $J(\tilde{\mathbf{T}})\tilde{R}\tilde{A}$ and s and t denote $\sin\theta$ and $\|\mathbf{T}\|$ respectively then we obtain the following inequality:

$$\begin{aligned} \dot{\mathcal{L}}(R,\mathbf{T}) &< -2\left(\frac{as^2}{\|\tilde{A}\mu\|} - \frac{\gamma bst}{\|\tilde{A}\mu\|} + \frac{\gamma d(\mathbf{T})at^2}{\|\tilde{A}\mathbf{T}\|}\right) \\ &< -2(s\ \ t) \begin{pmatrix} \frac{a}{\|\tilde{A}\mu\|} & -\frac{\gamma b}{2\|\tilde{A}\mu\|} \\ -\frac{\gamma b}{2\|\tilde{A}\mu\|} & \frac{\gamma d(\mathbf{T})a}{\|\tilde{A}\mathbf{T}\|} \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} \end{aligned}$$

Recognizing the right hand side of the expression above as a quadratic form we can conclude that this value will be less than zero for all non zero values of $\sin\theta$ and $\|\mathbf{T}\|$ if the two by two matrix in this expression is positive definite. Since the matrix $\tilde{R}\tilde{A}$ is assumed to be positive definite, we can conclude that the minimum eigenvalue a and, hence, the diagonal

entries of the matrix will be greater than zero. This implies that the matrix itself will be positive definite if its determinant, given in Equation 13 , evaluates to a positive number for all R and $\mathbf{T}$.

$$\frac{a}{\|\tilde{A}\mu\|}\frac{\gamma d(\mathbf{T})a}{\|\tilde{A}\mathbf{T}\|} - \left(\frac{\gamma b}{2\|\tilde{A}\mu\|}\right)^2 > 0 \qquad (13)$$

Rearranging Equation (13) yields.

$$\gamma < \|\tilde{A}\mu\| \left(\frac{a}{b}\right)^2 \left(\frac{d(\mathbf{T})}{\|\tilde{A}\mathbf{T}\|}\right) \qquad (14)$$

To recap, we have demonstrated that the derivative of the Lyapunov function is locally negative definite on the set $\{R, \mathbf{T} | R = \exp(\theta\mu), \theta < \pi, \|\mathbf{T}\| < \delta\}$ if there exists a γ such that the inequality given in Equation 14 is satisfied for all $R, \mathbf{T}$ in the set. It remains then to show that the expression given on the right hand side of the inequality can be bounded below by a finite number. We proceed by bounding the constituents $\|\tilde{A}\mu\|$ and $\frac{d(\mathbf{T})}{\|\tilde{A}\mathbf{T}\|}$ since a and b are constants. The expression $\|\tilde{A}\mu\|$ is bounded below by the square root of the minimum eigenvalue of the positive definite matrix $\tilde{A}^T\tilde{A}$since $\|\mu\| = 1$ by definition.

For any given positive value of ϵ it is possible to find a finite lower bound for the ratio $\frac{d(\mathbf{T})}{\|\tilde{A}\mathbf{T}\|}$ over the set of all $\mathbf{T}$ such that $\|\mathbf{T}\| > \epsilon$ since the both the quantities $d(\mathbf{T})$ and $\|\tilde{A}\mathbf{T}\|$ are required to be greater than zero for all non zero $\mathbf{T}$ and $\|\mathbf{T}\|$ cannot exceed δ by definition.

This implies that we need only consider the limiting behavior of this ratio as $\|\mathbf{T}\| \to 0$. By expanding $d(\mathbf{T}) = \|\acute{\mathbf{y}}(\mathbf{T}) - \mathbf{y_0}\|$ in a Taylor series and rewriting $\mathbf{T}$ as $\alpha\hat{\mathbf{T}}$ where $\|\hat{T}\| = 1$ and $\alpha \in \Re$, we obtain the following approximation for $\frac{d(\mathbf{T})}{\|\tilde{A}\mathbf{T}\|}$ in the neighborhood of $\mathbf{T} = \mathbf{0}$.

$$\frac{d(\mathbf{T})}{\|\tilde{A}\mathbf{T}\|} = \frac{\|\acute{\mathbf{y}}(\mathbf{T}) - \mathbf{y_0}\|}{\|\tilde{A}\mathbf{T}\|} \qquad (15)$$

$$\approx \frac{\|(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})\mathbf{T}\|}{\|\tilde{A}\mathbf{T}\|} \qquad (16)$$

$$\approx \frac{\|\alpha(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})\hat{\mathbf{T}}\|}{\|\alpha\tilde{A}\hat{\mathbf{T}}\|} \qquad (17)$$

$$\approx \frac{\sqrt{\hat{\mathbf{T}}^T(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})^T(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})\hat{\mathbf{T}}}}{\sqrt{\hat{\mathbf{T}}^T\tilde{A}^T\tilde{A}\hat{\mathbf{T}}}} \qquad (18)$$

Where $\bigtriangledown\acute{\mathbf{y}}_\mathbf{0}$ represents the Jacobian of the vector $\acute{\mathbf{y}}(T)$ with respect to the translation parameter $\mathbf{T}$ evaluated at $T = \mathbf{0}$ The denominator of the expression in equation 18 can be bounded above by the square root of the largest eigenvalue of $\tilde{A}^T\tilde{A}$ while the numerator can be bounded below by the square root of the smallest eigenvalue of the matrix $(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})^T(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})$. This value will be greater than zero as long as the image Jacobian $(\bigtriangledown\acute{\mathbf{y}}_\mathbf{0})$ has full rank. This will be true as long as the configuration of feature points being tracked is non-degenerate as we have already assumed.

In summary, the preceding analysis demonstrates that a finite constant, γ, can be found such that the derivative of the Lyapunov function given in Equation 12 is locally negative definite. This in turn implies that the control law regulates the pose to $(R, \mathbf{T}) = (I, \mathbf{0})$ asymptotically [21].

In practise the requirement that $\tilde{R}\tilde{A}$ be positive definite can be satisfied rather easily. Considering the terms individually, $\tilde{A}$ will be positive definite as long as its diagonal elements are positive and $\tilde{R}$ will be positive as long as it denotes a rotation of less than 90 degrees. Interestingly, this stability analysis does not place any constraints on the magnitude of the error parameter $\tilde{T}$.

3 Experimental Results

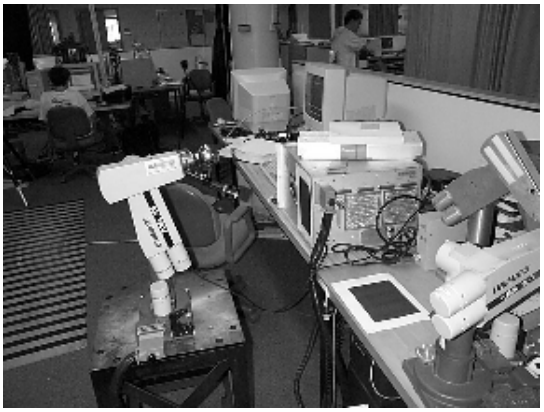

Figure 3: Pose control experiments were carried out on a Puma 260 robot arm outfitted with a video camera.

In order to evaluate the efficacy of the proposed servoing technique experiments were carried out using the Puma 260 robot shown in Figure 3. The relative orientation of a camera mounted on the robot's end effector was estimated by tracking a set of targets shown in Figure 4. Since this set of targets contained four pairs of parallel lines the system was able to estimate both the fundamental matrix F and the collineation

of the plane at infinity H_∞ directly without requiring any knowledge about the dimensions of the targets. No attempt was made to estimate the intrinsic or extrinsic calibration parameters of the camera.

Figure 4: This figure shows an image obtained from the camera mounted on the robot arm. The relative orientation of the camera with respect to the target position was obtained by tracking the corners of the black rectangles in the image. The white crosses are used to mark both the current locations of these features in this image and the target image configuration.

One of the reasons for choosing this platform for our experiments was the fact that the joint encoders of the robot provided an independent measure for the position of the robot manipulator which was used to measure the positioning accuracy of the scheme. Note however, that we were only able to estimate the position of the end effector and *not* the position of the center of projection of the camera because of the unknown extrinsic calibration parameter errors, $\tilde{R}$ and $\tilde{\mathbf{T}}$.

Three experiments were carried out to investigate the robustness of the technique to variations in the intrinsic and extrinsic parameters of the camera. Each experiment consisted of a series of trials where the robot manipulator was started at one position and then deliberately displaced from this target pose. The control law was then invoked to guide the manipulator back to the initial pose and the disparity between the position of the manipulator at convergence and the initial pose was measured. The control law was terminated when the disparity between the image measurements reflected in $d(\mathbf{T})$ dropped below a specified value.

In the first set of experiments the matrix of initial parameters, A, was set to $\begin{pmatrix} 100 & 0 & 220 \\ 0 & 100 & 340 \\ 0 & 0 & 1 \end{pmatrix}$. Note that these values were chosen arbitrarily since the actual calibration parameters were unknown. Over the sequence of five trials the mean rotational and translational errors at convergence were 0.50 degrees and 4.5 millimeters while the median rotational and translational errors were 0.46 degrees and 3.8 millimeters.

In the second set of experiments the camera was deliberately displaced from its original configuration by approximately 30 degrees and 10 centimeters. A new set of trials was carried out without changing any of the parameters in the program. Over this sequence of 6 trials the mean rotational and translational error at convergence were 2.23 degrees and 6.7 millimeters while the median translational and rotational errors were 0.54 degrees and 5.4 millimeters.

In the final set of experiments the matrix of intrinsic parameters A was changed to $\begin{pmatrix} 200 & 0 & 320 \\ 0 & 200 & 240 \\ 0 & 0 & 1 \end{pmatrix}$ note that this represents a substantial change in the parameters since the scale factors were doubled and the center of projection was displaced by over 100 pixels. Over this sequence of 6 trials the mean rotational and translational error at convergence were 4.70 degrees and 6.5 millimeters while the median translational and rotational errors were 3.9 degrees and 6.0 millimeters.

Note that in each case the controller was able to guide the manipulator end effector to within a centimeter of the desired pose.

4 Conclusions

This paper describes an approach to controlling the pose of a robotic platform using image measurements obtained from a camera rigidly mounted on that platform. The techniques make use of well-established techniques for computing the relative orientation of two camera positions from feature correspondences between the two images. These techniques do not require any a'priori knowledge of the locations of the features in the scene nor do they attempt to estimate such information online.

The techniques differ from methods based on pose estimation since they do not attempt to measure the location of the platform with respect to any particular target object in the scene. They also differ from other visual servoing techniques since no attempt is made to estimate the image Jacobian that relates control

actions to changes in the image measurements.

An analysis has been presented which demonstrates that the proposed control schemes are convergent in the face of large calibration errors in both the intrinsic and extrinsic parameters. From this analysis, we are able to derive a precise characterization of the magnitudes of the errors that can be tolerated.

These analytical results have been verified by experiments which indicate that the proposed pose control scheme is able to position the robot system accurately even in the presence of large calibration errors. These results demonstrate that it is possible to design systems with accurate positioning capabilities without having to expend much time and effort obtaining accurate calibration information about the system.

Acknowledgments

This work was supported by the NSF under grant number IRI-9711834, and DoD MURI DAAH04-96-1-0007

References

[1] Ronen Basri, E. Rivlin, and I. Shimshoni. Visual homing: surfing on the epipoles. In *International Conference on Computer Vision*, 1998.

[2] Koichiro Deguchi. Optimal motion control for image-based visual servoing by decoupling translation and rotation. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, page 705, 1998.

[3] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313–326, June 1992.

[4] Olivier Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.

[5] Olivier Faugeras. Stratification of three-dimensional vision: projective, affine and metric representations. *Journal of the Optical Society of America A*, 12(3):465, March 1995.

[6] G. Hager, W. Chang, and A.S. Morse. Robot feedback control based on stereo vision: Towards calibration-free hand-eye coordination. *IEEE Control Systems Magazine*, 15(1):30–39, 1995.

[7] Gregory D. Hager. Calibration-free visual control using projective invariance. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 1009–1015, 1995.

[8] Richard E. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Second European Conference on Computer Vision*, page 579, May 1992.

[9] Richard I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(6):580, June 1997.

[10] Nicholas Hollinghurst and Roberto Cipolla. Uncalibrated stereo hand-eye coordination. *Image and Vision Computing*, 12(3):187–192, April 1994.

[11] R. Horaud, F. Dornaika, and B. Espiau. Visually guided object grasping. *IEEE Trans. on Robotics and Automation*, 14(4):525–533, August 1998.

[12] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, September 1981.

[13] Ezio Malis, Francois Chaumette, and Sylvie Boudet. 2d 1/2 visual servoing stability analysis with respect to camera calibration errors. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, page 691, 1998.

[14] Ezio Malis, Francois Chaumette, and Sylvie Boudet. Positioning of a coarse-calibrated camera with respect to an unkown object by 2d 1/2 visual servoing. In *IEEE Int. Conf. on Robotics and Automation*, page 1352, 1998.

[15] N. P. Papanikolopoulos and P. K. Khosla. Eye-in-hand robotic tasks in uncalibrated environments. *IEEE Trans. Robotics and Automation*, 13(6):903–914, December 1997.

[16] N. P. Papanikolopoulos, B. J. Nelson, and P. K. Khosla. Six degree-of-freedom hand/eye visual tracking with uncertain parameters. *IEEE Trans. Robotics and Automation*, 11(5):725–732, October 1995.

[17] R. Pissard-Gibollet and P. Rives. Applying visual servoing techniques to control a mobile hand-eye system. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 166–171, Nagoya, Japan, May 1995.

[18] D. Simon, K. Kapellos, and B. Espiau. Design of control procedures for a free-floating underwater manipulation system. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 2158–2165, Albuquerque, April 1997.

[19] S. Soatto and P. Perona. Structure-independent visual motion control on the essential manifold. In *Proc. of the IFAC Symposium on Robot Control (SYROCO), Capri, Italy*, pages 869–876, Sept 1994.

[20] Camillo J. Taylor, James P. Ostrowski, and Sang-Hack Jung. Robust visual servoing based on relative orientation. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 574–580, June 1999.

[21] M. Vidyasagar. *Nonlinear Systems Analysis*. Prentice Hall, 2nd edition, 1993.

[22] W. J. Wilson, C. C. Williams Hulls, and Graham S. Bell. Relative end-effector control using Cartesian position based visual servoing. *IEEE Trans. Robotics and Automation*, 12(5):684–696, October 1996.