STEREO VISION-BASED TRAJECTORY FOLLOWING WITHOUT CORRESPONDENCE INFORMATION

Wen-Chung Chang

Department of Electrical Engineering
National Taipei University of Technology
Taipei 106, Taiwan, R.O.C.
E-Mail: wchang@ee.ntut.edu.tw

Keywords: Robot control; Task encoding; Trajectory following; Stereo vision; Visual servoing.

Abstract

This article addresses the visual servoing of a rigid robotic manipulator under fixed camera configuration. A tool is mounted on the end-effector of a robot which can be controlled by automatic visual feedback. The control goal is to drive the tip of the tool to follow a visually determined three dimensional target trajectory by using a two-camera vision system without assuming any *pointwise* correspondence information. Based on a novel encoded error, an image-based control law is proposed to achieve *precise* positioning of the tool in the absence of measurement noise and pointwise correspondence information. With an online falsification algorithm, one can further achieve automatic trajectory following using an approximately calibrated two-camera vision system. The control strategy is successfully validated in simulations on arbitrary line and contour following.

1 Introduction

Robotic manipulators have been widely employed in distinct environment in industry. Sensor-based design is known to be effective in providing precision and flexibility. Among a variety of sensing devices, vision is capable of measuring, recognizing, and object tracking in open workspace. Recent progress in computing makes vision a much more popular sensor. Therefore, vision-based control of robots has been an active research field and is being integrated into industrial applications [1, 8, 4, 2, 3].

Issues concerning task encoding in vision-based control systems have recently been discussed in [5]. It is shown that in the absence of measurement noise *precise* positioning sometimes can possibly be achieved despite camera model imprecision, just as it is in the case of a conventional set-point control system with a loop-integrator and fixed exogenous reference.

In this paper, a tool is mounted on the end-effector of a rigid robotic manipulator. The control task is to drive the tip of the robot tool to follow a visually determined target trajectory using an approximately calibrated two-camera vision system. Image-based encoding is chosen for defining a novel error in the two-camera vision-based control system. Based on the encoded error, an image-based control law is proposed to guarantee asymptotical convergence provided that camera modeling error is small.

The aim of this paper is to address issues in design of a robotic control system for three-dimensional trajectory following employing two-camera vision under fixed configuration. Observe that establishing *pointwise* correspondence on an arbitrary visually determined target trajectory, which might be a smooth curve without any sharp features like corners, is a hard problem. In fact, in order to make pointwise correspondence possible, one must employ epipolar constraint [6, 9]. But, without pointwise correspondence information, it seems that one cannot precisely position and follow this trajectory. However, a novel task encoding approach is proposed in this research without assuming pointwise correspondence information. Only knowledge of correspondence in the left and right images for the target trajectory as a whole is assumed. In Section 2 we briefly describe the vision/control system of interest and the control problem we are concerned with. In Section 3, we propose an effective image-based control law that assures asymptotical convergence. Simulation results are given in Section 4.

2 System description

An interesting control task that requires stereo vision is to precisely drive an observed point to any position in a line or a smooth contour. One would actually find difficulty when trying to encode such a task due to the fact that no precise pointwise correspondence information can be easily observed for such a line or a smooth contour. That is, one cannot precisely determine a set-point in either image space or Cartesian space to form an encoded error for controller design such that the encoded error being zero implies the required task being accomplished with precision.

In the problem we are concerned with, a tool is mounted on the end-effector of a robotic manipulator whose tip position can thus be controlled by performing rotational motion at each joint. For the purpose of driving the tip of the robotic tool to follow a target trajectory, two video cameras under fixed configuration is used as a vision sensor to observe the controlled point, the tip of the robotic tool, and the target trajectory. The problem of interest is to design an image-based controller using only visual information from the two cameras to achieve precision control. The system configuration used to perform autonomous positioning and trajectory following in three-dimensional space is shown in Figure 1. The problem of

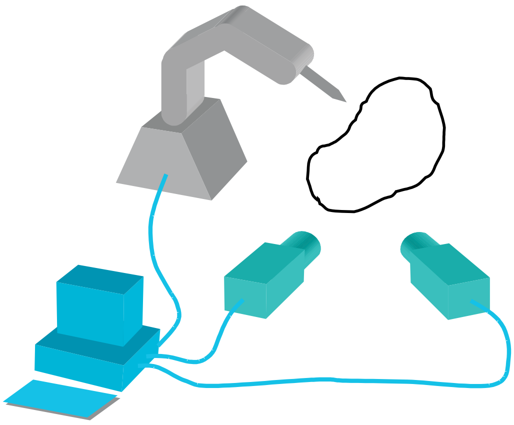

Figure 1: Configuration of the proposed visual servoing system capable of autonomously positioning and trajectory following in 3-D space

interest is to control the position of the tip of a robotic tool in a prescribed *workspace* $\mathcal{X} \subset \mathbb{R}^3$ using data observed by two *approximately calibrated* video cameras. Specifically, it consists of driving the tool tip to target positions in $\mathcal{X}$ determined by a contour $\mathcal{C}$ in the two-camera field of view $\mathcal{V} \subset \mathbb{R}^3$. The observed data consists of the tool tip position in $\mathcal{X}$ as well as target trajectory which appear in $\mathcal{V}$. Invariably $\mathcal{X} \subset \mathcal{V}$ and both $\mathcal{X}$ and $\mathcal{V}$ are compact subsets of $\mathbb{R}^3$.

The tool tip together with the target trajectory in $\mathcal{X}$ are seen in the image space $\mathcal{Y} \triangleq \mathbb{R}^2 \oplus \mathbb{R}^2$ through a fixed but imprecisely known, continuously differentiable, readout function or *perspective projection* [7] camera model $G : \mathcal{V} \rightarrow \mathcal{Y}$ which describes the two-camera vision system. Thus the position of the tool tip in the image space is a measured output y related to r by the formula

$$y = G(r) \tag{1}$$

where r is a state vector whose components are the Cartesian coordinates of the tool tip in $\mathcal{X}$. Similarly, the target position in the image space is an output y^* related to the target or desired set-point r^* by the formula

$$y^* = G(r^*).$$

In this robot positioning problem, the robot is assumed to admit a simple kinematic model of the form

$$\dot{r} = u$$

where r is a state vector whose components are the Cartesian coordinates of the robot in $\mathcal{X}$, and u is a control vector taking values in $\mathbb{R}^3$.

Generally for a stereo vision system one would expect G to be at least an injective function. We will assume that this is so. Clearly then, driving r to r^* is equivalent to driving y to y^* if y and y^* can both been observed. In this case, existing result has demonstrated effective control approaches [4]. However, in case the pointwise correspondence information for the target line or contour cannot be obtained, which further implies that one would not be able to determine y^*, existing control approaches would fail. Therefore, what this paper is concerned with is how one might define an effective encoded error and control laws to achieve precise positioning

$$r \longrightarrow \text{some } r^* \in \mathcal{C}$$

when the camera model G is not known precisely.

Moreover, one would like to see if it is possible to continue following the contour $\mathcal{C}$ after precisely reaching a point in $\mathcal{C}$.

2.1 Camera model

The symbols employed in camera model are listed in Table 1. In the sequel, prime denotes transpose.

$\mathcal{W}$	The robot coordinate system. A frame attached to the non-moving base of the robot.
c_1, c_2	The position of the optical centers of camera 1 and 2 relative to W respectively. $c_1, c_2 \in \mathbb{R}^3$.
f_1, f_2	The focal length of camera 1 and 2 respectively.
$[\mathbf{i}_1, \mathbf{j}_1, \mathbf{k}_1]'$ $[\mathbf{i}_2, \mathbf{j}_2, \mathbf{k}_2]'$	The rotation matrices of camera 1 and 2 respectively. $\{\mathbf{i}_1, \mathbf{j}_1, \mathbf{k}_1\}$ and $\{\mathbf{i}_2, \mathbf{j}_2, \mathbf{k}_2\}$ are sets of orthogonal unit vectors in $\mathbb{R}^3$.
r	The position in $\mathcal{W}$ of the vector between the location of the tool tip and c_1.
u	The robot control input. $u \in \mathbb{R}^3$.
y_1, y_2	The tool tip positions in the two cameras' image planes respectively. $y_1, y_2 \in \mathbb{R}^2$.

Table 1: Table of symbols for cameras

Camera coordinate directions are established as follows: for camera $n (n = 1, 2)$, $\mathbf{i}_n$ points to the right and $\mathbf{j}_n$ points downward in camera n's image plane, and $\mathbf{k}_n \triangleq \mathbf{i}_n \times \mathbf{j}_n$ points outward along the camera optical axis.

Using *perspective projection* [7] camera model, the nonlinear function G in (1) which maps from $\mathcal{V}$ to $\mathcal{Y}$ can be defined as follows.

$$G(r) \triangleq \begin{bmatrix} f_1 \frac{\mathbf{i}_1' r}{\mathbf{k}_1' r} \\ f_1 \frac{\mathbf{j}_1' r}{\mathbf{k}_1' r} \\ f_2 \frac{\mathbf{i}_2' (r+l)}{\mathbf{k}_2' (r+l)} \\ f_2 \frac{\mathbf{j}_2' (r+l)}{\mathbf{k}_2' (r+l)} \end{bmatrix}, \quad l \triangleq c_1 - c_2$$

2.2 Controlled process

Differentiating (1) with respect to time, we have

$$\dot{y} \quad = \quad J(r)\, u \tag{2}$$

where the Jacobian of the nonlinear map G is defined as

$$J(r) \triangleq \frac{\partial G(r)}{\partial r} = \begin{bmatrix} \frac{f_1}{\mathbf{k}_1' r}\left(\mathbf{i}_1' - \frac{\mathbf{i}_1' r}{\mathbf{k}_1' r}\mathbf{k}_1'\right) \\ \frac{f_1}{\mathbf{k}_1' r}\left(\mathbf{j}_1' - \frac{\mathbf{j}_1' r}{\mathbf{k}_1' r}\mathbf{k}_1'\right) \\ \frac{f_2}{\mathbf{k}_2'(r+l)}\left(\mathbf{i}_2' - \frac{\mathbf{i}_2'(r+l)}{\mathbf{k}_2'(r+l)}\mathbf{k}_2'\right) \\ \frac{f_2}{\mathbf{k}_2'(r+l)}\left(\mathbf{j}_2' - \frac{\mathbf{j}_2'(r+l)}{\mathbf{k}_2'(r+l)}\mathbf{k}_2'\right) \end{bmatrix}.$$

In the sequel, a novel encoded error is defined. Based on such an error, an image-based controller is proposed that will drive r to some $r^* \in \mathcal{C}$ asymptotically using only visual information from a two-camera vision system.

3 Positioning of the robot tool

The control task considered in this paper is to precisely position the tip of the robot tool in a prescribed workspace $\mathcal{W}$ using visual information observed from a two-camera vision system that is approximately calibrated. That is, one wants to drive the tool tip r to a visually determined target position r^* in a trajectory $\mathcal{C} \in \mathcal{W}$. Moreover, r is to be controlled to follow the trajectory $\mathcal{C}$. The challenging problem considered in the research is that no pointwise correspondence information is given about the trajectory $\mathcal{C}$. Due to the fact that r and r^* can never be precisely measured, one must consider using their projections in another space where they can be directly measured. A typical choice is the two-camera image space. Moreover, the original task should be re-encoded in such a space to guarantee accomplishment of the control task.

3.1 Task encoding

An intuitive choice is to define the encoded error in the image space as

$$e = y - y^*$$

where y and y^* are the image coordinates of r and r^* in $\mathcal{Y}$ respectively. However, due to the fact that the corresponding information about the desired set-point cannot be obtained, y^* cannot be determined. One must encode this task in a way that would guarantee the accomplishment of the original task without assuming any pointwise correspondence information.

For the control task that requires precisely positioning the tool tip in the contour $\mathcal{C}$, the following encoded error is proposed which does not need any pointwise corresponding information about the desired set-point.

$$e_{4\times 1}(t) \equiv y(t) - y_d(t) = \begin{bmatrix} y_1(t) - y_{d_1}(t) \\ y_2(t) - y_{d_2}(t) \end{bmatrix} \tag{3}$$

where

$$y_{d_i}(t) \equiv \arg\left\{ \min_{y_{d_i} \in G_i(\mathcal{C})} \|y_i(t) - y_{d_i}\| \right\}, i = 1, 2. \tag{4}$$

Since the desired control task is accomplished if and only if the tool tip touches with the target trajectory $\mathcal{C}$ *in both image planes* which further implies that the *final values* of y_{d_1} and y_{d_2} correspond to a physical point in $\mathcal{C}$, one can be sure that this simple image-based encoding approach will ensure precise positioning of tool tip using two-camera vision. Moreover, one should keep in mind that this image-based encoding is invariant on camera models. That is, the knowledge of camera calibration parameters is not required.

3.2 Falsification of admissible set-points

In addition to precisely positioning the tip of robot tool in an observed target trajectory $\mathcal{C}$ without given pointwise correspondence information, a more complicated task with potential applications is to further drive the tool tip to follow the trajectory. Since the pointwise correspondence information is again not given, this tracking problem is seemingly a challenging one. A falsifier is proposed to eliminate set-points that have been reached thereby generating over time a descending chain of subsets of admissible set-points.

Geometrically, the positioning task demands that the visual distance in both image planes be zero. Algebraically, this requirement takes the form

$$H_i(y_i(t), y_{d_i}) = 0,\ i = 1, 2$$

where

$$H_i(y_i(t), y_{d_i}) \triangleq \|y_i(t) - y_{d_i}\|.$$

As a consequence of the algebraic constraint, one can see that $H_i(y_i(t), y_{d_i}) = 0,\ i = 1, 2$ for the set-point (y_{d_1}, y_{d_2}) given $(y_1(t), y_2(t))$ just in case the tip of robot tool coincides with a point in the trajectory $\mathcal{C}$ in image space. In principle this set-point should therefore be excluded in order to continue following the trajectory $\mathcal{C}$. Of course any noise in the system could foil such a drastic falsification policy, so what we shall consider instead is something a little less severe. The idea is to discard set-point pairs for which $H_1(y_1(t), y_{d_1}) < \epsilon$ and $H_2(y_2(t), y_{d_2}) < \epsilon$ where ϵ is some small, prespecified number. This idea is heuristic and can be justified via simulation.

We now describe in more detail the falsifier. For each pair $y_1(t)$ and $y_2(t)$, let $\mathcal{F}(t) \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ denote the *falsified set*.

$$\mathcal{F}(t) \triangleq \big\{(y_{d_1}, y_{d_2}) | H_1(y_1(t), y_{d_1}) < \epsilon \text{ and } H_2(y_2(t), y_{d_2}) < \epsilon,\ y_{d_i} \in G_i(\mathcal{C}),\ i = 1, 2\big\}$$

The descending chain of subsets mentioned before, can now be defined recursively by the formulas.

$$\mathcal{M}(0) \triangleq \mathcal{M}_1(0) \times \mathcal{M}_2(0) = G_1(\mathcal{C}) \times G_2(\mathcal{C}) \tag{5}$$

and

$$\mathcal{M}(t_{j+1}) \triangleq \mathcal{M}(t_j) - \mathcal{F}(t_j) \cap \mathcal{M}(t_j) \tag{6}$$

where $\mathcal{M}(t_j) - \mathcal{F}(t_j) \cap \mathcal{M}(t_j)$ is the complement of $\mathcal{F}(t_j) \cap \mathcal{M}(t_j)$ in $\mathcal{M}(t_j)$. As defined, the $\mathcal{M}(t_j)$ clearly form the descending chain.

$$G_1(\mathcal{C}) \times G_2(\mathcal{C}) = \mathcal{M}(0) \supset \mathcal{M}(t_1) \supset \mathcal{M}(t_2) \supset \cdots \supset \mathcal{M}(t_i) \supset \cdots$$

In order to reduce computation, one can actually perform (6) by reducing candidate set-points according to the following formula.

$$F(t_j) \cap \mathcal{M}(t_j) = \{(y_{d_1}, y_{d_2}) | H_1(y_1(t_j), y_{d_1}) < \epsilon \text{ and } H_2(y_2(t_j), y_{d_2}) < \epsilon, \, (y_{d_1}, y_{d_2}) \in \mathcal{M}(t_j)\}$$

Moreover, the set of admissible set-points $\mathcal{M}$ cannot be reduced unless the controlled tip point reaches any of its elements. This further implies that the set $\mathcal{M}$ might converge to the empty set if the controlled tip point has reached all admissible set-points under appropriate control action. That is, the contour following task has been accomplished with precision in the order of ϵ. Note in addition, that the computation of the $\mathcal{M}(t_i)$ can be carried out by a causal algorithm because the only real-time data needed to evaluate $\mathcal{M}(t_i)$ at each time t_i are $\mathcal{M}(t_{i-1})$ and $y(t_{i-1})$. Meanwhile, due to the nonzeroness of ϵ, the set of admissible set-points $\mathcal{M}(t)$ could be computed at discrete times to avoid infinite falsifications {in finite time}.

In the light of the admissible set of set-points $\mathcal{M}(t)$ at time t (5) and (6), the encoded error (3) can be further modified as follows.

$$e_{4\times 1}(t) = \begin{bmatrix} y_1(t) - y_{d_1}(t) \\ y_2(t) - y_{d_2}(t) \end{bmatrix} \tag{7}$$

where

$$y_{d_i}(t) = \arg \left\{ \min_{y_{d_i} \in \mathcal{M}_i(t)} \|y_i(t) - y_{d_i}\| \right\}, i = 1, 2.$$

3.3 Image-based control approach

In the light of the image-based encoded error defined in (7), it follows from (2) that

$$\dot{e} \quad = \quad J(r)\, u - \Delta(t) \tag{8}$$

where

$$\Delta(t) \equiv \frac{d}{dt} \begin{bmatrix} \arg \left\{ \min_{y_{d_1} \in \mathcal{M}_1(t)} \|y_1 - y_{d_1}\| \right\} \\ \arg \left\{ \min_{y_{d_2} \in \mathcal{M}_2(t)} \|y_2 - y_{d_2}\| \right\} \end{bmatrix} \leq \alpha < \infty$$

because of the fact that the trajectory of the controlled point and the target trajectory to follow are both smooth and thus the selected set-point, $y_{d_i}(t), i = 1, 2$, must also be smooth.

Based on the process model defined in (8), the following image-based control law is proposed.

$$u = -\eta(y, y_d) \cdot \Big(\left[Q \circ G^{-1} \right](y) \Big) \cdot \Big(\left[J \circ G^{-1} \right](y) \Big)' \cdot e \tag{9}$$

where $\eta(y, y_d) \triangleq \frac{\mathrm{v}_{\max}}{\max\left\{ \left\| \left([Q \circ G^{-1}](y) \right) \cdot \left([J' \circ G^{-1}](y) \right) \cdot e \right\|, \frac{\mathrm{v}_{\max}}{k} \right\}}$ is a saturation function, k is a constant gain, $\mathrm{v}_{\max}$ is the maximum speed of the robot, $Q(\cdot)_{3\times 3}$ is a symmetric positive definite matrix[2], and $G^{-1} : \mathcal{Y} \to \mathcal{V}$ is a continuously differentiable left inverse of G. The control law (9) can drive e to zero exponentially.

If G is modelled only approximately by some perspective projection function $G_q : \mathcal{V} \to \mathcal{Y}$, with G_q^{-1} a continuously differentiable left inverse of G_q, the following feedback law is proposed.

$$u = -\hat{\eta}(y, y_d) \cdot \Big(\left[Q \circ G_q^{-1} \right](y) \Big) \cdot \Big(\left[J_q \circ G_q^{-1} \right](y) \Big)' \cdot e \tag{10}$$

where $\hat{\eta}(y, y_d) \triangleq \frac{\mathrm{v}_{\max}}{\max\left\{ \left\| \left([Q \circ G_q^{-1}](y) \right) \cdot \left([J_q \circ G_q^{-1}](y) \right) \cdot e \right\|, \frac{\mathrm{v}_{\max}}{k} \right\}}$ and $J_q(r) \triangleq \frac{\partial G_q(r)}{\partial r}$. The control law (10) can still drive e to zero exponentially provided that G_q were a good enough approximate model of G.

3.4 Stability analysis

The stability analysis for the set-point problem can basically be performed as follows. Let y_1^* and y_2^* be the set-points in the two image planes that cannot be directly observed due to unknown pointwise correspondence. Defining $e_{d_i}(t) = y_i(t) - y_i^*$ and $e_{b_i}(t) = y_i^* - y_{d_i}(t)$, one can see that

$$e(t) = \begin{bmatrix} y_1(t) - y_{d_1}(t) \\ y_2(t) - y_{d_2}(t) \end{bmatrix} = \begin{bmatrix} e_{d_1}(t) + e_{b_1}(t) \\ e_{d_2}(t) + e_{b_2}(t) \end{bmatrix} = e_d(t) + e_b(t) \tag{11}$$

In the light of (8) and (10), one can obtain the following equation by differentiating (11).

$$\dot{e}_d(t) = \dot{e}(t) - \dot{e}_b(t) = -\hat{\eta} J Q J_q' e_d - \hat{\eta} J Q J_q' e_b - \Delta(t) - \dot{e}_b$$

According to the proposed algorithm for selecting y_d one can see that e_b, defined in the image planes, is smooth and finite. Therefore, there exists a finite positive number β such that

$$\dot{e}_d(t) \quad \leq \quad -\eta J Q J' e_d + \beta$$

[2] For example, $\left[Q \circ G^{-1} \right](y)$ can be chosen as $\left(\left(\left[J \circ G^{-1} \right](y) \right)' \left(\left[J \circ G^{-1} \right](y) \right) \right)^{-1}$.

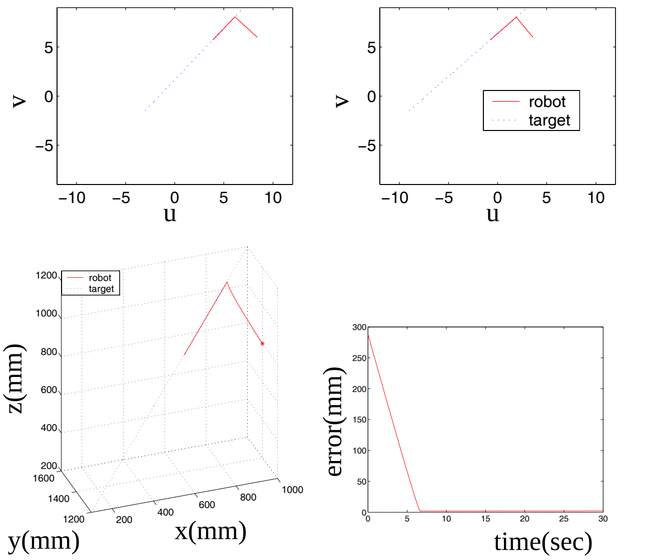

Figure 2: From approaching the line to touching and following the line-[exact calibration]: stereo image trajectories (*upper*), 3-D trajectories (*bottom left*), and distance to target trajectory (*bottom right*)

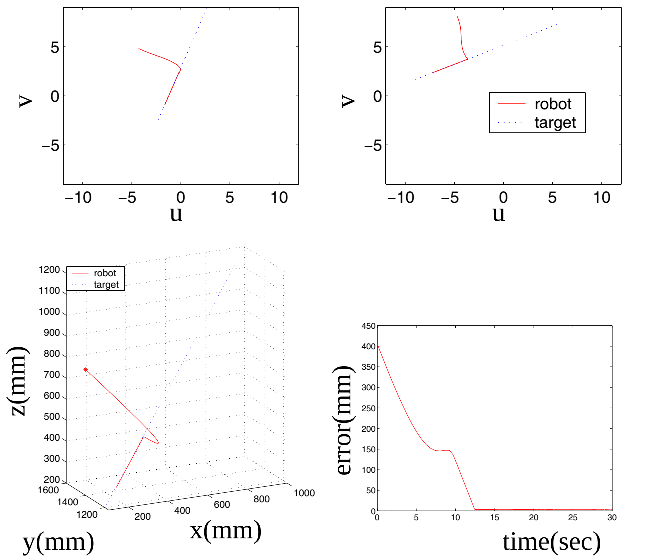

Figure 3: From approaching the line to touching and following the line-[cameras rotated inwards 20 degrees]: stereo image trajectories (*upper*), 3-D trajectories (*bottom left*), and distance to target trajectory (*bottom right*)

In the light of the *Visual Constraints* in [4], one thus concludes that

$$e_d(t) \quad \rightarrow \quad 0 \tag{12}$$

exponentially fast. This is equivalent to say that y coincides with y^* in the image space. Based on the epipolar constraint, one further assure that the tip of the robot tool has reached a point in the target trajectory. Hence, by virtue of (4) and (12) one can see that

$$e_b(t) \rightarrow 0 \text{ and } e(t) \rightarrow 0.$$

4 Simulations

The proposed robotic task is to drive the tip of a robot tool to follow a target trajectory visually determined by a two-camera vision system. The observed target trajectory can be any arbitrary line or contour in the three-dimensional space.

The proposed system is evaluated through two typical target trajectories. One is a straight line and the other a circle both in the three-dimensional space and can be observed by the two cameras. Based on the proposed control law, the tool tip is driven to a point in the target trajectory and continue to follow this trajectory.

In the unit of millimeter, The two cameras each with a focal length of 12 are positioned at $[300\,400\,300]'$ and $[700\,400\,300]'$ with respect to the world frame respectively. Both optical axes are oriented to point along the y-axis of the world frame. The trajectories of the target and the tool tip are plotted as dotted and solid lines respectively. The star indicates the initial position of the tool tip. In the following two target trajectories, the proposed control algorithm is further verified to demonstrate its robustness and performance. Specifically, the two cameras are rotated inwards 20 degrees. No deteriorated performance can be observed when the two cameras are shifted. In all simulations, the distance to target trajectory converges to zero exponentially. In fact, due to the saturation function $\hat{\eta}$ defined in (10), the tool tip is driven to the target trajectory at almost constant maximum velocity 50 mm/sec. Moreover, it stays within 1 millimeter while following the trajectory at about the speed of 15 mm/sec by setting k, $\mathrm{v}_{\max}$, and ϵ to be 20, 50, and 0.02 respectively. Simulation results are illustrated in Figure 2 to 5.

5 Conclusion

In this paper, we have proposed an effective image-based control law for driving an observed robotic controlled point to follow a target trajectory in three-dimensional space. The controlled point is first driven to a point, specified in only one image or not, in the target trajectory and then continue to follow this trajectory. In order to follow the arbitrary trajectory in 3-D space, a falsifier is proposed that is capable of eliminating admissible set-points. Based on a novel encoded error, an image-based control law is proposed. This control approach together with the falsification of admissible set-points achieve trajectory following using two-camera vision without assuming pointwise corresponding information. This encoding approach avoids getting into complicated epipolar constraints that relies on precise calibration of camera parameters. It is based on a simple and effective rule that can be performed on the observed images from two-camera vision. Moreover, only approximately calibrated two-camera vision is needed in our approach to precisely position the controlled point in real-time. The target trajectory is visually determined allowing the pro-

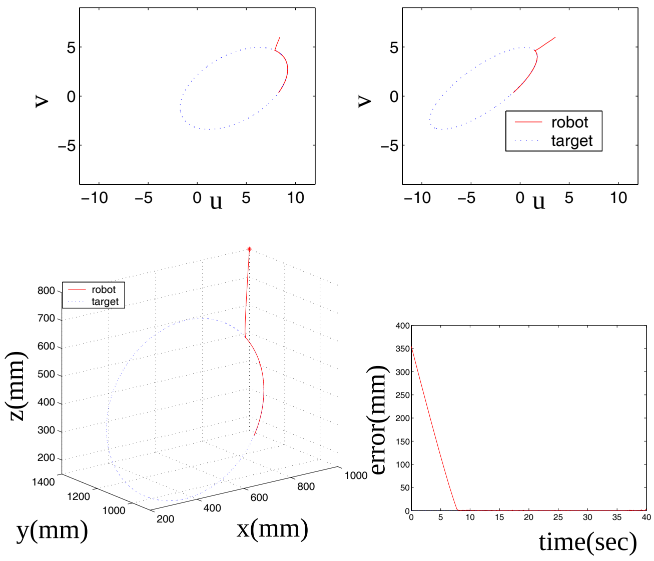

Figure 4: From approaching the circle to touching and following the circle-[exact calibration]: stereo image trajectories (*upper*), 3-D trajectories (*bottom left*), and distance to target trajectory (*bottom right*)

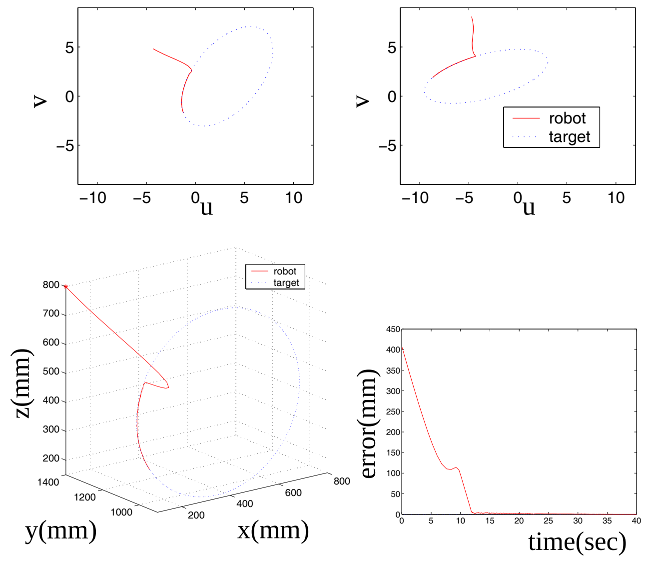

Figure 5: From approaching the circle to touching and following the circle-[cameras rotated inwards 20 degrees]: stereo image trajectories (*upper*), 3-D trajectories (*bottom left*), and distance to target trajectory (*bottom right*)

posed system capable of performing autonomous control tasks.

The proposed control law that is robust with respect to the camera modelling error can asymptotically stabilize the set-point control system. Furthermore, due to the fact that the encoded error being zero implies the controlled point having reached a point in the target trajectory, one can perform automatic trajectory following without assuming pointwise correspondence information.

The basic results presented here are to be extended in a number of ways. Although the simulations only validate the control approach for planar target trajectories in 3-D space, one can actually perform the same controller to precisely drive the controlled point to follow a three-dimensional contour again without assuming pointwise correspondence information. One can also arbitrarily select a particular set-point in one of the two image planes and then apply the same encoding approach in the other image plane to reach the desired set-point. Comparison with other control laws that are synthesized based on Cartesian-based or other encoding is also of interest. It is likely that these issues will form the basis of a class of adaptive algorithms which do not require the knowledge of camera calibration parameter values and correspondence information.

Acknowledgements

The author was supported by the National Science Council, R.O.C. under grant NSC 90-2213-E-027-026.

References

[1] Andrés Castaño and Seth A. Hutchinson. Visual compliance: Task-directed visual servo control. *IEEE Trans. Robot. Automat.*, 10(3):334–342, June 1994.

[2] Wen-Chung Chang. Visual servoing of rigid robots with modified cartesian-based encoding. In *Proc. of the 2000 American Control Conference*, Chocago, U.S.A., June 2000. IEEE.

[3] Wen-Chung Chang. Real-time vision-based contour following with laser pointer. In *Proc. of the 2003 International Conference on Robotics and Automation*, Taipei, Taiwan, R.O.C., September 2003.

[4] Wen-Chung Chang and A. S. Morse. Exponentially stable positioning of a rigid robot using stereo vision. In *Proc. of the 1999 International Conference on Robotics and Automation*, Detroit, MI, U.S.A., May 1999.

[5] Wen-Chung Chang and A. S. Morse. Six degree-of-freedom task encoding in vision-based control systems. In *Proc. of the 14th World Congress of IFAC*, volume B, pages 311–316, Beijing, China, July 1999. International Federation of Automatic Control.

[6] Olivier Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.

[7] Horn. *Robot Vision*. The MIT Electrical Engineering and Computer Science Series. McGraw-Hill, New York, 8 edition, 1986.

[8] Seth A. Hutchinson, Gregory D. Hager, and Peter I. Corke. A tutorial on visual servo control. *IEEE Trans. Robot. Automat.*, 12(5):651–670, October 1996.

[9] Quang-Tuan Luong and Oliver D. Faugeras. Camera calibration, scene motion and structure recovery from point correspondences and fundamental matrices. *Int. J. of Computer Vision*, 17(12), December 1995.