Robotica (2007) volume 25, pp. 615–626. © 2007 Cambridge University Press
doi:10.1017/S0263574707003505 Printed in the United Kingdom

Binocular vision-based 3-D trajectory following for autonomous robotic manipulation

Wen-Chung Chang*

Department of Electrical Engineering, National Taipei University of Technology, NTUT Box 2125, Taipei 106, Taiwan, R.O.C.

(Received in Final Form: March 1, 2007. First published online: April 17, 2007)

SUMMARY

Robotic manipulators that have interacted with uncalibrated environments typically have limited positioning and tracking capabilities, if control tasks cannot be appropriately encoded using available features in the environments. Specifically, to perform 3-D trajectory following operations employing binocular vision, it seems necessary to have *a priori* knowledge on pointwise correspondence information between two image planes. However, such an assumption cannot be made for any smooth 3-D trajectories. This paper describes how one might enhance autonomous robotic manipulation for 3-D trajectory following tasks using eye-to-hand binocular visual servoing. Based on a novel encoded error, an image-based feedback control law is proposed without assuming pointwise binocular correspondence information. The proposed control approach can guarantee task precision by employing only an approximately calibrated binocular vision system. The goal of the autonomous task is to drive a tool mounted on the end-effector of the robotic manipulator to follow a visually determined smooth 3-D target trajectory in desired speed with precision. The proposed control architecture is suitable for applications that require precise 3-D positioning and tracking in unknown environments. Our approach is successfully validated in a real task environment by performing experiments with an industrial robotic manipulator.

KEYWORDS: Binocular correspondence; Binocular vision; Robot control; Task encoding; Trajectory following; Visual servoing.

1. Introduction

Robotic manipulators have been widely employed in distinct environments in industry. Sensor-based design is known to be effective in providing precision and flexibility. Among the various sensors, 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.[1,2]

Due to a variety of task requirements, robotic arms would usually need precise positioning and tracking to accomplish complicated tasks. Typical tasks of this type include contour following, trajectory tracking, as well as manufacturing-related tasks that demand motion tracking control.[3,4,5,6] Shen *et al.*[7] presented an asymptotic trajectory tracking control approach using uncalibrated eye-to-hand visual feedback. The robotic fingertip can be driven to follow a predetermined reference trajectory in the vision frame. This controller is especially useful in applications where hand-eye precalibration is not possible. Xiao *et al.*[8] developed an adaptive hybrid visual and force servoing approach to achieve a hybrid position/froce control task in an unknown environment. The manipulator was visually guided to follow a 3-D trajectory with six-axis force measurements in addition to an eye-to-hand vision sensor. Akella[9] introduced an adaptive position-tracking control approach for planar robotic manipulator with eye-to-hand visual feedback. All vision sensors above are in fixed camera configuration, which are feasible for systems requiring both the manipulator and the environment to be observed by the camera at the cost of online or offline camera calibration. On the other hand, eye-in-hand configuration has the advantage of formulating problems in task frame but it generally allows only local workspace due to the restricted camera field of view. Gangloff and Mathelin[10] proposed an eye-in-hand visual servoing of a six-DOF manipulator for unknown 3-D profile following. The profile has an unknown curvature, but its cross section is known. In[11], Bettini *et al.* presented the design and implementation of a vision-assisted planar manipulation system using virtual fixtures at millimeter to micrometer scales. An eye-in-hand vision sensor provides the reference trajectory for path following tasks. These earlier works on trajectory following did not seem to have considered using a binocular vision system to observe an unknown smooth 3-D target trajectory. In fact, to perform 3-D trajectory following, control with only eye-to-hand binocular vision is a challenging problem since no pointwise correspondence information between the two image planes is available.

In this paper, a tool is mounted on the end-effector of a rigid robotic manipulator. The control task is to drive the tool tip to precisely follow a visually determined 3-D target trajectory using an approximately calibrated binocular vision system. The paper is organized as follows. The preliminary definitions and problem formulation are described in Section 2. Section 3 then presents the proposed encoded error and control approach for point-to-trajectory positioning tasks. Moreover, the modified encoded error and control law for point-to-trajectory tracking tasks are introduced in Section 4. Section 5 reports the experimental results and Section 6 the

* Corresponding author. E-mail: wchang@ee.ntut.edu.tw

conclusion, where the significance of the experiments and the perspectives of the research are addressed.

2. Preliminaries

2.1. Notation

Let bold uppercase denote matrix, calligraphic alphabet denote set, bold lowercase denote vector, prime denote transpose, superscripts + denote pseudoinverse, and leading superscripts identify the space a quantity is written in. Moreover:

- $\mathcal{X}, \mathcal{V}, {}^{I}\mathcal{X}$: the robot workspace $\mathcal{X} \subset \mathbb{R}^3$, the binocular-camera field of view $\mathcal{V} \subset \mathbb{R}^3$, and the binocular image space ${}^{I}\mathcal{X} \triangleq \mathbb{R}^2 \oplus \mathbb{R}^2$;
- $\boldsymbol{r}, \boldsymbol{r}^*$: the position of the robot and the target in $\mathcal{X}$;
- ${}^{I}\boldsymbol{r}, {}^{I}\boldsymbol{r}^*$: the position of the robot and the target in ${}^{I}\mathcal{X}$;
- $\boldsymbol{u} \in \mathbb{R}^3$: the robot control input;
- $\boldsymbol{c}_n \in \mathbb{R}^3$, f_n: the position of the optical center and focal length of camera n respectively, $n = 1, 2$;
- SO^3: special orthogonal group of order 3;
- $[\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{z}_n]' \in SO^3$: the rotation matrices of camera n, $n = 1, 2$;
- $\boldsymbol{o}_n$: the image centers in the image plane of camera n, $n = 1, 2$;
- $G : \mathcal{V} \to {}^{I}\mathcal{X}$: binocular-camera model;
- ${}^{I}\boldsymbol{r}_n \in \mathbb{R}^2$: the robot position in the image plane of camera n, ${}^{I}\boldsymbol{r}_n \triangleq [{}^{I}r_{nx}, {}^{I}r_{ny}]'$, $n = 1, 2$;
- ${}^{I}\boldsymbol{r}_n^* \in \mathbb{R}^2$: the target position in the image plane of camera n, ${}^{I}\boldsymbol{r}_n^* \triangleq [{}^{I}r_{nx}^*, {}^{I}r_{ny}^*]'$, $n = 1, 2$;
- ${}^{I}\boldsymbol{r}_{d_n} \in \mathbb{R}^2$: the minimal-distance desired position in the image plane of camera n, ${}^{I}\boldsymbol{r}_{d_n} \triangleq [{}^{I}r_{d_{nx}}, {}^{I}r_{d_{ny}}]'$, $n = 1, 2$.

2.2. System description

The problem of interest is to perform 3-D unknown trajectory following control of a robotic manipulator employing only binocular visual feedback as illustrated in Fig. 1. The goal of the autonomous task is to drive the tool tip of a robotic manipulator to a visually determined 3-D smooth trajectory and continue tracking the trajectory in desired speed with precision. The robotic manipulator is equipped with a binocular vision system that does not need to be precisely calibrated. The target 3-D trajectory is assumed unknown and arbitrary. A visual detection and tracking system provides the evolution of the binocular image projections of the tool tip and the target trajectory in real time. This visual information allows the robotic manipulator to perform 3-D trajectory following tasks.

To perform such an autonomous 3-D trajectory following task employing only binocular vision, one would actually find difficulty when trying to encode such a task due to the fact that no precise pointwise binocular correspondence information can be easily observed. By *pointwise binocular correspondence* we mean that the projection of each point in a smooth 3-D trajectory onto binocular image space

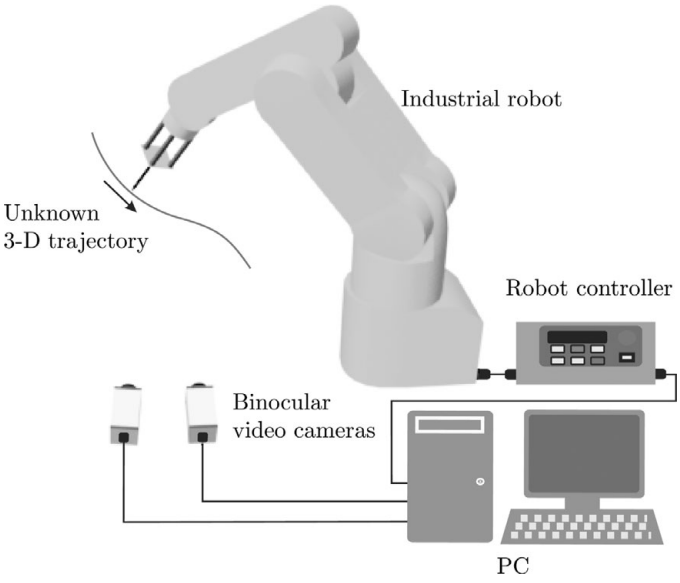

Fig. 1. Configuration of the proposed visual servoing system capable of autonomously positioning and trajectory following in 3-D space.

can be identified precisely. Therefore, 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.[12–15]

The problem of interest is to control the position of the robotic tool tip in a prescribed *workspace* $\mathcal{X} \subset \mathbb{R}^3$ using data observed by *approximately calibrated* binocular video cameras. Specifically, it consists of driving the tool tip to target positions in $\mathcal{X}$ determined by a 3-D trajectory $\mathcal{C}$ in the binocular-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 $\mathbb{R}^3$.

2.3. Problem formulation

The tool tip in $\mathcal{X}$ together with the target trajectory in $\mathcal{V}$ are seen in the binocular image space ${}^{I}\mathcal{X} \triangleq \mathbb{R}^2 \oplus \mathbb{R}^2$ through a fixed but imprecisely known, continuously differentiable, readout function or *perspective projection*[16] camera model $G : \mathcal{V} \to {}^{I}\mathcal{X}$, which describes the binocular vision system. Specifically, camera coordinate directions of the binocular vision system are established as follows: for camera 1, $\boldsymbol{x}_1$ points to the right and $\boldsymbol{y}_1$ points downward in image plane of camera 1, and $\boldsymbol{z}_1 \triangleq \boldsymbol{x}_1 \times \boldsymbol{y}_1$ points outward along the camera optical axis. Camera coordinate directions for camera 2 are established similarly. Thus, G is the nonlinear function, which maps from $\mathcal{V}$ to ${}^{I}\mathcal{X}$ and is defined as

$$G(\boldsymbol{r}) \triangleq \begin{bmatrix} f_1 \frac{\boldsymbol{x}_1'\boldsymbol{r}}{\boldsymbol{z}_1'\boldsymbol{r}} \\ f_1 \frac{\boldsymbol{y}_1'\boldsymbol{r}}{\boldsymbol{z}_1'\boldsymbol{r}} \\ f_2 \frac{\boldsymbol{x}_2'(\boldsymbol{r}+\boldsymbol{l})}{\boldsymbol{z}_2'(\boldsymbol{r}+\boldsymbol{l})} \\ f_2 \frac{\boldsymbol{y}_2'(\boldsymbol{r}+\boldsymbol{l})}{\boldsymbol{z}_2'(\boldsymbol{r}+\boldsymbol{l})} \end{bmatrix}, \quad \boldsymbol{l} \triangleq \boldsymbol{c}_1 - \boldsymbol{c}_2 \tag{1}$$

where $\boldsymbol{r}$ is the robot position in $\mathcal{X}$ relative to $\boldsymbol{c}_1$.

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

The position of the tool tip in the binocular image space is a measured output ${}^I\boldsymbol{r}$ related to $\boldsymbol{r}$ by the formula

$$ {}^I\boldsymbol{r} = G(\boldsymbol{r}). \tag{2} $$

Similarly, the target position in the binocular image space is an output ${}^I\boldsymbol{r}^*$ related to the target or desired set-point $\boldsymbol{r}^*$ by the formula

$$ {}^I\boldsymbol{r}^* = G(\boldsymbol{r}^*). \tag{3} $$

The velocity-controlled robotic manipulator is assumed to admit a simple kinematic model of the form

$$ \dot{\boldsymbol{r}} = \boldsymbol{u} \tag{4} $$

where $\boldsymbol{u}$ is a velocity control vector taking values in $\mathbb{R}^3$. Differentiating Eq. (2) with respect to time, we have

$$ {}^I\dot{\boldsymbol{r}} = \boldsymbol{J}(\boldsymbol{r})\,\boldsymbol{u} \tag{5} $$

where the Jacobian of the nonlinear map G is defined as

$$ \boldsymbol{J}(\boldsymbol{r}) \triangleq \frac{\partial G(\boldsymbol{r})}{\partial \boldsymbol{r}} = \begin{bmatrix} \frac{f_1}{z'_{1\boldsymbol{r}}}\left(\boldsymbol{x}'_1 - \frac{x'_{1\boldsymbol{r}}}{z'_{1\boldsymbol{r}}}\boldsymbol{z}'_1\right) \\ \frac{f_1}{z'_{1\boldsymbol{r}}}\left(\boldsymbol{y}'_1 - \frac{y'_{1\boldsymbol{r}}}{z'_{1\boldsymbol{r}}}\boldsymbol{z}'_1\right) \\ \frac{f_2}{z'_2(\boldsymbol{r}+\boldsymbol{l})}\left(\boldsymbol{x}'_2 - \frac{x'_2(\boldsymbol{r}+\boldsymbol{l})}{z'_2(\boldsymbol{r}+\boldsymbol{l})}\boldsymbol{z}'_2\right) \\ \frac{f_2}{z'_2(\boldsymbol{r}+\boldsymbol{l})}\left(\boldsymbol{y}'_2 - \frac{y'_2(\boldsymbol{r}+\boldsymbol{l})}{z'_2(\boldsymbol{r}+\boldsymbol{l})}\boldsymbol{z}'_2\right) \end{bmatrix}. \tag{6} $$

There are two subtasks in the 3-D trajectory following task of interest. One is *point-to-trajectory positioning* and the other is *point-to-trajectory tracking*. The point-to-trajectory positioning control problem is to synthesize a feedback control law $\boldsymbol{u}$ using only binocular visual measurements such that

$$ \boldsymbol{r} \longrightarrow \text{ some } \boldsymbol{r}^* \in \mathcal{C}. \tag{7} $$

As regards the point-to-trajectory tracking control problem, one needs to synthesize a binocular visual feedback control law $\boldsymbol{u}$ such that $\boldsymbol{r}$ precisely tracks the smooth 3-D trajectory $\mathcal{C}$ in desired speed.

Generally speaking, one would expect G to be at least an injective function in order for a stereo vision system to provide the requisite information for positioning and tracking tasks. We will assume that this is so. Clearly then, driving $\boldsymbol{r}$ to $\boldsymbol{r}^*$ is equivalent to driving ${}^I\boldsymbol{r}$ to ${}^I\boldsymbol{r}^*$ if ${}^I\boldsymbol{r}$ and ${}^I\boldsymbol{r}^*$ can both been observed. In this case, existing result has demonstrated effective control approaches.[17] However, in case the pointwise correspondence information for the 3-D target trajectory cannot be obtained, which further implies that one would not be able to determine ${}^I\boldsymbol{r}^*$, existing control approaches would fail. Therefore, this paper is concerned with how one might define effective encoded errors and control laws to achieve precise positioning and tracking when the binocular camera model G is not known precisely.

In the sequel, a novel encoded error is defined. Based on such an error, an image-based controller is proposed, which can guarantee accomplishment of the point-to-trajectory positioning task using only available binocular visual information. In addition, another image-based control law capable of driving a modified encoded error to zero for point-to-trajectory tracking task is presented.

3. Point-to-trajectory positioning

To precisely drive the tool tip $\boldsymbol{r}$ to any set-point $\boldsymbol{r}^*$ in an unknown smooth 3-D trajectory $\mathcal{C} \subset \mathcal{X}$ using a binocular vision system that is only approximately calibrated, one must define a positioning error without assuming pointwise binocular correspondence. Unlike traditional set-point problems, the choice of an encoded error for such a positioning task is not clear.

3.1. Encoding the positioning task

Due to the fact that the binocular correspondence information about the desired set-point cannot be obtained (i.e., binocular image points ${}^I\boldsymbol{r}_1^*$ and ${}^I\boldsymbol{r}_2^*$ that correspond to the same 3-D point $\boldsymbol{r}^*$ cannot be identified), ${}^I\boldsymbol{r}^*$ cannot be determined. Thus, the following intuitive encoded error defined in the binocular image space cannot be applied to the considered task.

$$ \boldsymbol{e}(t) \triangleq {}^I\boldsymbol{r}(t) - {}^I\boldsymbol{r}^*. \tag{8} $$

Instead, the following encoded error is proposed, which does not need any pointwise binocular correspondence information about the desired set-point.

$$ \boldsymbol{e}(t) \triangleq {}^I\boldsymbol{r}(t) - {}^I\boldsymbol{r}_d(t). \tag{9} $$

Specifically,

$$ \begin{bmatrix} \boldsymbol{e}_1(t) \\ \boldsymbol{e}_2(t) \end{bmatrix} = \begin{bmatrix} {}^I\boldsymbol{r}_1(t) - {}^I\boldsymbol{r}_{d_1}(t) \\ {}^I\boldsymbol{r}_2(t) - {}^I\boldsymbol{r}_{d_2}(t) \end{bmatrix} \tag{10} $$

where ${}^I\boldsymbol{r}_{d_n}, n = 1, 2,$ are the minimal-distance desired positions in image space of camera n, modeled by G_n, and are defined as

$$ {}^I\boldsymbol{r}_{d_n}(t) \triangleq \arg\left\{ \min_{{}^I\boldsymbol{r}_c \in G_n(\mathcal{C})} \|{}^I\boldsymbol{r}_n(t) - {}^I\boldsymbol{r}_c\| \right\}, \quad n = 1, 2. \tag{11} $$

Obviously, the proposed encoded error is zero if and only if the tool tip touches the target trajectory $\mathcal{C}$ in binocular image planes. This further implies that ${}^I\boldsymbol{r}_{d_1}$ and ${}^I\boldsymbol{r}_{d_2}$ correspond to a physical point in $\mathcal{C}$, and the desired positioning task is precisely accomplished. For those system configurations satisfying Assumption 3.1, this can actually be made possible.

Assumption 3.1. The binocular vision system and the smooth 3-D trajectory $\mathcal{C}$ are configured such that any two points in the 3-D trajectory are not coplanar with the optical centers of binocular cameras $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$.

Clearly, the assumption stated above requires that any *epipolar plane*[18] containing a point in the 3-D trajectory must not intersect with this 3-D trajectory at any other points. In

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

other words, the projection of $\mathcal{C}$ onto each of the binocular image planes can be well defined by a function $x = g(y)$, where x and y are the horizontal and vertical coordinates, respectively. Therefore, based on Assumption 3.1, it follows that

$$\boldsymbol{e} = 0 \implies \boldsymbol{e}_d = 0 \quad \text{and} \quad \boldsymbol{e}_b = 0 \tag{12}$$

where

$$\boldsymbol{e}_{d_n}(t) \triangleq {}^I\boldsymbol{r}_n(t) - {}^I\boldsymbol{r}_n^*,\ \boldsymbol{e}_{b_n}(t) \triangleq {}^I\boldsymbol{r}_n^* - {}^I\boldsymbol{r}_{d_n}(t), \qquad n = 1, 2 \tag{13}$$

and

$$\boldsymbol{e}(t) = \boldsymbol{e}_d(t) + \boldsymbol{e}_b(t) = \begin{bmatrix} \boldsymbol{e}_{d_1}(t) + \boldsymbol{e}_{b_1}(t) \\ \boldsymbol{e}_{d_2}(t) + \boldsymbol{e}_{b_2}(t) \end{bmatrix}. \tag{14}$$

This is due to the fact that as long as the controlled point coincides with the trajectory in both the binocular image planes, it physically contacts the trajectory at a single point. If this is not so, the controlled point might stay at a 3-D point that is away from the trajectory but in the plane containing the optical centers $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ while itself touching the trajectory in binocular image space. Thus, $\boldsymbol{e} = 0$ and there is no correspondence between the two contacting points in the trajectory on the binocular image planes. However, this contradicts with the assumption that any plane containing the optical centers $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ does not intersect with the 3-D trajectory at more than one point.

3.2. Image-based control design

In the light of the image-based encoded error defined in Eq. (9), it follows from Eq. (5) that

$$\dot{\boldsymbol{e}} = \boldsymbol{J}(\boldsymbol{r})\,\boldsymbol{u} - {}^I\dot{\boldsymbol{r}}_{\mathrm{d}}(t) \tag{15}$$

where

$${}^I\dot{\boldsymbol{r}}_{\mathrm{d}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \arg\left\{ \min\limits_{{}^I\boldsymbol{r}_c \in G_1(\mathcal{C})} \| {}^I\boldsymbol{r}_1(t) - {}^I\boldsymbol{r}_c \| \right\} \\ \arg\left\{ \min\limits_{{}^I\boldsymbol{r}_c \in G_2(\mathcal{C})} \| {}^I\boldsymbol{r}_2(t) - {}^I\boldsymbol{r}_c \| \right\} \end{bmatrix} \quad \text{and}$$

$$\| {}^I\dot{\boldsymbol{r}}_{\mathrm{d}}(t) \| \le \alpha < \infty. \tag{16}$$

The reason why ${}^I\dot{\boldsymbol{r}}_{\mathrm{d}}(t)$ is finite is due to the fact that the trajectory of the controlled point and the target trajectory to follow are both smooth and, thus, the selected set-points, ${}^I\boldsymbol{r}_{\mathrm{d}_i}(t)$, $i = 1, 2$, must also be smooth.

Based on the process model defined in Eq. (15) and the novel encoded error defined in Eq. (9), the following image-based control law is proposed.

$$\boldsymbol{u} = -k\eta({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_d)([\boldsymbol{Q} \circ G^{-1}]({}^I\boldsymbol{r}))([\boldsymbol{J} \circ G^{-1}]({}^I\boldsymbol{r}))'\boldsymbol{e} \tag{17}$$

where

$$\eta({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_{\mathrm{d}}) \triangleq \upsilon_{\max}(\max\{k \| ([\boldsymbol{Q} \circ G^{-1}]({}^I\boldsymbol{r})) ([\boldsymbol{J} \circ G^{-1}]({}^I\boldsymbol{r}))'\boldsymbol{e} \|, \upsilon_{\max}\})^{-1} \tag{18}$$

is a saturation function, $k > 0$ is a constant gain, $\upsilon_{\max}$ is the maximum speed of the robot, $\boldsymbol{Q}(\cdot)_{3\times 3}$ is a symmetric positive definite matrix,[1] and $G^{-1} : {}^I\mathcal{X} \to \mathcal{V}$ is a continuously differentiable left inverse of G. The control law in Eq. (17) can drive $\boldsymbol{e}$ to zero exponentially.

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

$$\boldsymbol{u} = -k\hat{\eta}({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_d)\big([\boldsymbol{Q} \circ G_q^{-1}]({}^I\boldsymbol{r})\big)\big([\boldsymbol{J}_q \circ G_q^{-1}]({}^I\boldsymbol{r})\big)'\boldsymbol{e} \tag{19}$$

where

$$\hat{\eta}({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_d) \triangleq \upsilon_{\max}\big(\max\big\{k \big\| \big([\boldsymbol{Q} \circ G_q^{-1}]({}^I\boldsymbol{r})\big) \big([\boldsymbol{J}_q \circ G_q^{-1}]({}^I\boldsymbol{r})\big)'\boldsymbol{e} \big\|, \upsilon_{\max}\big\}\big)^{-1} \tag{20}$$

and $\boldsymbol{J}_q(\boldsymbol{r}) \triangleq \frac{\partial G_q(\boldsymbol{r})}{\partial \boldsymbol{r}}$. The control law in Eq. (19) can still drive $\boldsymbol{e}$ to zero exponentially provided that G_q were a good enough approximate model of G.

3.3. Stability analysis

The exponential stability of the point-to-trajectory positioning system employing visual feedback control law in Eq. (17) can be established in Theorem 3.1.

Theorem 3.1. (Exponential stability) Let

$$\mathcal{B} \triangleq \{{}^I\boldsymbol{r} \in G(\mathcal{X}) \mid \| {}^I\boldsymbol{r} - {}^I\boldsymbol{r}^* \| \le \rho\} \tag{21}$$

such that $\mathcal{S} \triangleq \{\boldsymbol{r} \in \mathcal{X} \mid G(\boldsymbol{r}) \in \mathcal{B}\}$ *is bounded and* G *is injective on* $\mathcal{S}$. *Suppose that there exists a positive number* ν_1 *satisfying* $\nu_1 < 1$ *and such that for every* ${}^I\boldsymbol{r}, {}^I\boldsymbol{r}^* \in \mathcal{B}$

$$\| \boldsymbol{e}_b \| \le \nu_1 \| \boldsymbol{e}_d \|^{\dagger}. \tag{22}$$

For the system defined by Eqs. (14)–(17), there exist positive numbers a *and* λ *which depend only on* $\mathcal{S}$ *such that if* $\boldsymbol{r}(0), \boldsymbol{r}^* \in \mathcal{S}$,

$$\| \boldsymbol{r}(t) - \boldsymbol{r}^* \| \le a \cdot e^{-\lambda t} \cdot \| \boldsymbol{r}(0) - \boldsymbol{r}^* \|, \qquad \forall t \ge 0. \tag{23}$$

Proof of Theorem 3.1. Define a Lyapunov function candidate

$$V = \| \boldsymbol{e}_d \|^2. \tag{24}$$

Let $\boldsymbol{w}$ in ${}^I\mathcal{X}$ be defined such that $\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d$ and $\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{w}$ are orthogonal in $\mathbb{R}^3$ and $\mathrm{span}\{\boldsymbol{w}, \boldsymbol{e}_d\} = \mathrm{span}\{\boldsymbol{e}_b, \boldsymbol{e}_d\}$. $\boldsymbol{e}_b$ can thus be expressed as

$$\boldsymbol{e}_b = \alpha \boldsymbol{e}_d + \beta \boldsymbol{w} \tag{25}$$

[1] For example, $[\boldsymbol{Q} \circ G^{-1}]({}^I\boldsymbol{r})$ can be chosen as the commonly used pseudoinverse $(([\boldsymbol{J} \circ G^{-1}]({}^I\boldsymbol{r}))'([\boldsymbol{J} \circ G^{-1}]({}^I\boldsymbol{r})))^{-1}$.

† Note that the vision system is in incorrect correspondence situations when $\boldsymbol{e}_d = -\boldsymbol{e}_b$.

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

where (α, β) is the coordinate of $\boldsymbol{e}_b$ with respect to the basis $\{\boldsymbol{w},\boldsymbol{e}_d\}$. For the positioning task, ${}^I\boldsymbol{r}^*$ is the set-point. Therefore, one can see from Eq. (13) that

$$\dot{\boldsymbol{e}}_b(t) = -{}^I\dot{\boldsymbol{r}}_d(t). \tag{26}$$

Differentiating V with respect to time, we have

$$\begin{aligned}
\dot{V} &= -2k\eta \boldsymbol{e}_d' \boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}'\boldsymbol{e} && (27)\\
&= -2k\eta\big(\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\|^2 + \boldsymbol{e}_d'\boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}'\boldsymbol{e}_b\big) && (28)\\
&\le -2k\eta\big(\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\|^2 - |\boldsymbol{e}_d'\boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}'\boldsymbol{e}_b|\big) && (29)\\
&= -2k\eta\big(\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\|^2 - |\boldsymbol{e}_d'\boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}'(\alpha\boldsymbol{e}_d + \beta\boldsymbol{w})|\big) && (30)\\
&= -2k\eta\big(\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\|^2 - \big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\| \cdot \big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'(\alpha\boldsymbol{e}_d)\big\|\big). && (31)
\end{aligned}$$

Due to the fact that $\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d$ and $\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{w}$ are orthogonal in $\mathbb{R}^3$, it follows from Eq. (25) that

$$\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'(\alpha\boldsymbol{e}_d)\big\| \le \big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_b\big\|. \tag{32}$$

In the light of Eq. (22), it thus follows that

$$\dot{V} \le -2k\eta(1-\nu_1)\big\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}'\boldsymbol{e}_d\big\|^2 \le 0. \tag{33}$$

From Eq. (33), we see that $\boldsymbol{y}$ stays inside the ball $\mathcal{B}$. In other words, solution $\boldsymbol{r}$ never leaves $\mathcal{S}$. Thus, $\mathcal{S}$ is a bounded invariant set. Hence, $\boldsymbol{r}$ is bounded. Meanwhile, for all $t \ge 0$

$$\int_0^t |\dot{V}(\tau)|\,d\tau = -\int_0^t \dot{V}(\tau)\,d\tau = V(0) - V(t) \le V(0). \tag{34}$$

We see that $\dot{V} \in \mathcal{L}_1$. Furthermore, one can check that $\ddot{V}$ is piecewise continuous and bounded. Therefore, we conclude that

$$\dot{V} \longrightarrow 0 \quad \text{as} \quad t \to \infty \tag{35}$$

and thus

$$\boldsymbol{J}'(\boldsymbol{r})\,\boldsymbol{e}_d \longrightarrow 0 \quad \text{as} \quad t \to \infty \tag{36}$$

since $\boldsymbol{Q}(\boldsymbol{r})$ is a symmetric positive definite matrix. From *binocular visual constraint*[17] we know that if $\boldsymbol{e}_d$ is in the null space of $\boldsymbol{J}'(\boldsymbol{r})$, it must be true that $\boldsymbol{r} = \boldsymbol{r}^*$ for all $\boldsymbol{r}, \boldsymbol{r}^* \in \mathcal{S}$. Therefore, we see that

$$\boldsymbol{r}(t) \longrightarrow \boldsymbol{r}^* \quad \text{as} \quad t \to \infty. \tag{37}$$

Moreover, it can be shown[17] that there exists a positive number μ, which depends only on $\mathcal{S}$ such that for every $\boldsymbol{r}, \boldsymbol{r}^* \in \mathcal{S}$

$$\|\boldsymbol{J}'(\boldsymbol{r})(G(\boldsymbol{r}) - G(\boldsymbol{r}^*))\| \ge \mu\|G(\boldsymbol{r}) - G(\boldsymbol{r}^*)\|. \tag{38}$$

Hence, it follows from Eq. (33) that

$$\dot{V} \le -2k\eta(1-\nu_1)\sigma^2\|\boldsymbol{J}'(\boldsymbol{r})\,\boldsymbol{e}_d\|^2 \le -2k\eta(1-\nu_1)\sigma^2\,\mu\,V \tag{39}$$

where $\sigma > 0$ is the least singular value of $\boldsymbol{Q}^{1/2}(\boldsymbol{r})$, $\boldsymbol{r} \in \mathcal{S}$.

Therefore, $\|\boldsymbol{e}_d\|$ goes to zero exponentially fast. By injectivity of G on $\mathcal{S}$, one thus concludes that there exist positive numbers a, λ, which depend only on $\mathcal{S}$ such that if $\boldsymbol{r}(0), \boldsymbol{r}^* \in \mathcal{S}$,

$$\|\boldsymbol{r}(t) - \boldsymbol{r}^*\| \le a \cdot e^{-\lambda t} \cdot \|\boldsymbol{r}(0) - \boldsymbol{r}^*\|, \qquad \forall t \ge 0. \tag{40}$$

□

The feedback control law in Eq. (19), in which the camera model is assumed to be known approximately, is employed to the system described by Eqs. (2), (4), (9), and (11). Theorem 3.2 states that if the approximate camera model is close enough to the true model and the initial positioning error is small enough, the introduced image-based control law can still guarantee exponential convergence and precise positioning.

Theorem 3.2. (Robustness) *Suppose that there exists a positive number*

$$\nu < \tilde{\eta}\sigma^2\mu \tag{41}$$

and such that for every ${}^I\boldsymbol{r} \in \mathcal{B}$

$$\delta({}^I\boldsymbol{r}) \le \nu \tag{42}$$

where

$$\tilde{\eta} \triangleq \min_{{}^I\boldsymbol{r}\in\mathcal{B}}\{\hat{\eta}({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_d)\} \tag{43}$$

$\mu > 0$ *is defined such that Eq. (38) holds for every* $\boldsymbol{r}, \boldsymbol{r}^* \in \mathcal{S}$, $\sigma > 0$ *is the least singular value of* $\boldsymbol{Q}^{1/2}(\boldsymbol{r})\forall \boldsymbol{r} \in \mathcal{S}$, *and*

$$\begin{aligned}
\delta({}^I\boldsymbol{r}) \triangleq\ & \nu_1\big\|([\boldsymbol{J}\circ G^{-1}]({}^I\boldsymbol{r}))\big(\big[\boldsymbol{Q}\circ G_q^{-1}\big]({}^I\boldsymbol{r})\big)\big(\big[\boldsymbol{J}_q\circ G_q^{-1}\big]({}^I\boldsymbol{r})\big)'\big\|\\
&+\big\|([\boldsymbol{J}\circ G^{-1}]({}^I\boldsymbol{r}))(([\boldsymbol{Q}\circ G^{-1}]({}^I\boldsymbol{r}))([\boldsymbol{J}\circ G^{-1}({}^I\boldsymbol{r})]'\\
&-\big(\big[\boldsymbol{Q}\circ G_q^{-1}\big]({}^I\boldsymbol{r})\big)\big(\big[\boldsymbol{J}_q\circ G_q^{-1}\big]({}^I\boldsymbol{r})\big)'\big)\big\|. \qquad (44)
\end{aligned}$$

The solution to the system defined by Eqs. (2), (4), (9), (11), (19), (21), and (22) exists globally and $\boldsymbol{r} \longrightarrow \boldsymbol{r}^*$ *exponentially if* $\|\boldsymbol{e}_d(0)\| \le h\,\rho$ *for every* $h \in [0, 1)$.

Proof of Theorem 3.2. In the light of Eqs. (15), (19), and (26), one can obtain the following equation by differentiating Eq. (14) with respect to time.

$$\begin{aligned}
\dot{\boldsymbol{e}}_d(t) = &-k\hat{\eta}\boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}'\boldsymbol{e}_d + k\hat{\eta}\boldsymbol{J}\big(\boldsymbol{Q}\boldsymbol{J}' - \big[\boldsymbol{Q}\circ G_q^{-1}\big]\\
&\big[\boldsymbol{J}_q'\circ G_q^{-1}\big]\big)\boldsymbol{e}_d - k\hat{\eta}\boldsymbol{J}\big[\boldsymbol{Q}\circ G_q^{-1}\big]\big[\boldsymbol{J}_q'\circ G_q^{-1}\big]\boldsymbol{e}_b
\end{aligned} \tag{45}$$

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

By differentiating Eq. (24) with respect to time, one can see that

$$\|\boldsymbol{e}_d\|\frac{d\|\boldsymbol{e}_d\|}{dt} = \boldsymbol{e}_d'\dot{\boldsymbol{e}}_d. \tag{46}$$

It thus follows from Eqs. (45) and (46) that

$$\begin{aligned}\|\boldsymbol{e}_d\|\frac{d\|\boldsymbol{e}_d\|}{dt} = &-k\hat{\eta}\left\|\boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{J}\boldsymbol{e}_d\right\|^2 + k\hat{\eta}\boldsymbol{e}_d'\boldsymbol{J}\\ &\left(\boldsymbol{Q}\boldsymbol{J}' - \left[\boldsymbol{Q}\circ G_q^{-1}\right]\left[\boldsymbol{J}_q'\circ G_q^{-1}\right]\right)\boldsymbol{e}_d\\ &- k\hat{\eta}\boldsymbol{e}_d'\boldsymbol{J}\left[\boldsymbol{Q}\circ G_q^{-1}\right]\left[\boldsymbol{J}_q'\circ G_q^{-1}\right]\boldsymbol{e}_b.\end{aligned} \tag{47}$$

From Eq. (38) and recalling that $\sigma > 0$ is the least singular value of $\boldsymbol{Q}^{\frac{1}{2}}(\boldsymbol{r})$, $\boldsymbol{r} \in \mathcal{S}$, we have

$$\begin{aligned}\frac{d\|\boldsymbol{e}_d\|}{dt} \le &-\hat{\eta}\sigma^2\mu\|\boldsymbol{e}_d\| + \hat{\eta}\frac{\boldsymbol{e}_d'}{\|\boldsymbol{e}_d\|}\\ &\boldsymbol{J}\left(\boldsymbol{Q}\boldsymbol{J}' - \left[\boldsymbol{Q}\circ G_q^{-1}\right]\left[\boldsymbol{J}_q'\circ G_q^{-1}\right]\right)\boldsymbol{e}_d\\ &-\hat{\eta}\frac{\boldsymbol{e}_d'}{\|\boldsymbol{e}_d\|}\boldsymbol{J}\left[\boldsymbol{Q}\circ G_q^{-1}\right]\left[\boldsymbol{J}_q'\circ G_q^{-1}\right]\boldsymbol{e}_b.\end{aligned} \tag{48}$$

Observing the fact from Eq. (20) that $\hat{\eta}({}^I\boldsymbol{r}, {}^I\boldsymbol{r}_d) \in (0, 1]$ for every ${}^I\boldsymbol{r} \in \mathcal{B}$, the definitions of $\tilde{\eta}$ in Eq. (43) and δ in Eq. (44), and the condition in Eq. (22), one can further see that

$$\frac{d\|\boldsymbol{e}_d\|}{dt} \le -k\tilde{\eta}\sigma^2\mu\|\boldsymbol{e}_d\| + k\delta\|\boldsymbol{e}_d\|. \tag{49}$$

By using the variation of constant formula one concludes that

$$\begin{aligned}\|\boldsymbol{e}_d(t)\| \le &\|\boldsymbol{e}_d(0)\|e^{-k\tilde{\eta}\sigma^2\mu t}\\ &+ \int_0^t e^{-k\tilde{\eta}\sigma^2\mu(t-\tau)}k\delta({}^I\boldsymbol{r}(\tau))\|\boldsymbol{e}_d(\tau)\|\,d\tau.\end{aligned} \tag{50}$$

Multiplying both sides by $e^{k\tilde{\eta}\sigma^2\mu t}$, we have

$$e^{k\tilde{\eta}\sigma^2\mu t}\|\boldsymbol{e}_d(t)\| \le \|\boldsymbol{e}_d(0)\| + \int_0^t e^{k\tilde{\eta}\sigma^2\mu\tau}k\delta({}^I\boldsymbol{r}(\tau))\|\boldsymbol{e}_d(\tau)\|\,d\tau. \tag{51}$$

Applying the Bellman–Gronwall Lemma,[19] one further concludes that

$$e^{k\tilde{\eta}\sigma^2\mu t}\|\boldsymbol{e}_d(t)\| \le \|\boldsymbol{e}_d(0)\|e^{\int_0^t k\delta({}^I\boldsymbol{r}(\tau))d\tau}. \tag{52}$$

Therefore,

$$\|\boldsymbol{e}_d(t)\| \le \|\boldsymbol{e}_d(0)\|e^{-\int_0^t k\left(\tilde{\eta}\sigma^2\mu - \delta({}^I\boldsymbol{r}(\tau))\right)d\tau}, \qquad \forall\, t \ge 0. \tag{53}$$

It can be shown that if $\|\boldsymbol{e}_d(0)\| \le h\rho$, $\|\boldsymbol{e}_d(t)\| < \rho$ for all $t \ge 0$. In fact, by contradiction assume that there exists some time $\bar{t}$ for which $\|\boldsymbol{e}_d(t)\|$ gets to be larger or equal to ρ. That is,

$$\|\boldsymbol{e}_d(\bar{t})\| = \rho \quad \text{and} \quad \|\boldsymbol{e}_d(t)\| < \rho, \qquad \forall\, t \in [0, \bar{t}). \tag{54}$$

Hence, one can see that

$$\delta({}^I\boldsymbol{r}(\tau)) \le \nu, \qquad \forall\, t \in [0, \bar{t}). \tag{55}$$

From Eqs. (53) and (55), one concludes that

$$\|\boldsymbol{e}_d(\bar{t})\| \le \|\boldsymbol{e}_d(0)\|e^{-\int_0^{\bar{t}} k\left(\tilde{\eta}\sigma^2\mu - \nu\right)d\tau} \le \|\boldsymbol{e}_d(0)\| \le h\rho < \rho \tag{56}$$

which contradicts the definition of $\bar{t}$. Moreover, note that

$$\|\boldsymbol{e}_d(t)\| < \rho, \forall\, t \ge 0 \quad \Longrightarrow \quad \delta({}^I\boldsymbol{r}(\tau)) \le \nu,\ \forall\, t \ge 0. \tag{57}$$

In the light of Eqs. (53) and (57), and the fact that $\nu < \tilde{\eta}\sigma^2\mu$, one concludes that

$$\|\boldsymbol{e}_d(t)\| \le \|\boldsymbol{e}_d(0)\|e^{-\int_0^t k\left(\tilde{\eta}\sigma^2\mu - \nu\right)d\tau} \to 0 \quad \text{as} \quad t \to \infty. \tag{58}$$

□

The fact that $\|\boldsymbol{e}_d(t)\|$ decays to zero exponentially is equivalent to say that ${}^I\boldsymbol{r}$ coincides with ${}^I\boldsymbol{r}^*$ in the binocular image space. By injectivity of G on $\mathcal{S}$ and Assumption 3.1, one further assures that the tip of the robot tool has reached a point in the target trajectory with precision.

4. Point-to-trajectory tracking

In addition to precisely positioning the tool tip 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. The encoded error for the tracking task is modified from the positioning encoded error by demanding required tracking trajectory in one of the two binocular image planes. Under such a newly defined image-based task encoding approach, feedback control laws can be synthesized for driving the tool tip to track visually determined smooth 3-D target trajectories in desired speed with precision. Even if the vision system is not exactly calibrated, precise tracking can still be guaranteed by the proposed image-based feedback control law.

4.1. Target trajectory generation

For the control task that requires the tool tip to continue tracking the smooth 3-D trajectory $\mathcal{C}$ after completing the positioning task, the encoded error defined in Eq. (9) cannot be directly applied since the tool tip is already on the trajectory. In order to precisely track smooth 3-D trajectories, the projection of the 3-D trajectory onto the image space must be identified visually before the trajectory following command can be determined in the image plane.

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

For the trajectory following task illustrated in Fig. 1, the target trajectory in image space can be extracted by the vision system. Let g_n denote the function that relates the components of the projection of $\boldsymbol{r}^*$ onto the image plane of camera n by

$$ {}^I r^*_{n_x} = g_n\big({}^I r^*_{n_y}\big), \qquad n = 1, 2. \tag{59} $$

Given the desired vertical $\boldsymbol{y}$-component of the tracking velocity along the target trajectory, ${}^I\dot{\boldsymbol{r}}^*_{n_y}(t) = v^*(t)$, the horizontal $\boldsymbol{x}$-component can be computed by

$$ {}^I\dot{r}^*_{n_x}(t) = \frac{dg_n}{d^I r^*_{n_y}} {}^I\dot{r}^*_{n_y}(t), \qquad n = 1, 2. \tag{60} $$

Therefore, the desired trajectory in the image space can be defined as

$$ {}^I\boldsymbol{r}^*_n(t) = \begin{bmatrix} {}^I r^*_{n_x}(0) + \displaystyle\int_0^t \frac{dg_n}{d^I r^*_{n_y}} {}^I v^*(\tau)\, d\tau \\ {}^I r^*_{n_y}(0) + \displaystyle\int_0^t {}^I v^*(\tau)\, d\tau \end{bmatrix}, \qquad n = 1, 2. \tag{61} $$

Similarly, with the given desired tracking speed, $v^*(t)$, along the projection of the 3-D trajectory onto the image plane, the vertical $\boldsymbol{y}$-component of the tracking velocity can be computed by

$$ {}^I\dot{\boldsymbol{r}}^*_{n_y}(t) = v^*(t) \left(1 + \left(\frac{dg_n}{d^I r^*_{n_y}} \right)^2 \right)^{-1/2}, \qquad n = 1, 2. \tag{62} $$

Hence, the horizontal $\boldsymbol{x}$-component of the tracking velocity in image space can be obtained by Eq. (60). Meanwhile, the desired trajectory in the image space can thus be defined by

$$ {}^I\boldsymbol{r}^*_n(t) = \begin{bmatrix} {}^I r^*_{n_x}(0) + \displaystyle\int_0^t \frac{dg_n}{d^I r^*_{n_y}} {}^I v^*(\tau) \left(1 + \left(\frac{dg_n}{d^I r^*_{n_y}} \right)^2 \right)^{-\frac{1}{2}} d\tau \\ {}^I r^*_{n_y}(0) + \displaystyle\int_0^t {}^I v^*(\tau) \left(1 + \left(\frac{dg_n}{d^I r^*_{n_y}} \right)^2 \right)^{-\frac{1}{2}} d\tau \end{bmatrix}, \qquad n = 1, 2. \tag{63} $$

4.2. Trajectory following control design

In order to perform the required trajectory following tasks, the encoded errors can be defined by modifying the encoded error for positioning tasks introduced in Eq. (9).

$$ \boldsymbol{e}_1(t) = {}^I\boldsymbol{r}_1(t) - {}^I\boldsymbol{r}^*_1(t) \tag{64} $$

$$ \boldsymbol{e}_2(t) = {}^I\boldsymbol{r}_2(t) - {}^I\boldsymbol{r}_{d_2}(t) \tag{65} $$

where ${}^I\boldsymbol{r}^*_1(t)$ is the visually determined target trajectory in the image plane of camera 1 and ${}^I\boldsymbol{r}_{d_2}(t)$ is defined in Eq. (11). Differentiating Eqs. (64) and (65) with respect to time, the differential kinematic equations can be seen as

$$ \dot{\boldsymbol{e}}_1(t) = \boldsymbol{J}_1(\boldsymbol{r})\boldsymbol{u}(t) - {}^I\dot{\boldsymbol{r}}^*_1(t) \tag{66} $$

$$ \dot{\boldsymbol{e}}_2(t) = \boldsymbol{J}_2(\boldsymbol{r})\boldsymbol{u}(t) - {}^I\dot{\boldsymbol{r}}_{d_2}(t). \tag{67} $$

To make the tracking problem tractable, the encoded errors in Eqs. (64) and (65) introduced above do not require any *a priori* knowledge on binocular pointwise correspondence of the target trajectory. The tracking control problem is to develop a feedback control law capable of driving both $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ to zero exponentially. By injectivity of the binocular camera model G and Assumption 3.1, $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ both driven to zero exponentially implies that $\boldsymbol{r}(t)$ precisely tracks $\mathcal{C}$ with zero steady-state error.

Nominal control laws with feedforward compensation for the two tasks modelled by Eqs. (66) and (67) are introduced, respectively, as

$$ \boldsymbol{u}(t) = \boldsymbol{J}_1^+(\boldsymbol{r})\big(-\lambda_1 \boldsymbol{e}_1(t) + {}^I\dot{\boldsymbol{r}}^*_1(t)\big) \tag{68} $$

$$ \boldsymbol{u}(t) = \boldsymbol{J}_2^+(\boldsymbol{r})\big(-\lambda_2 \boldsymbol{e}_2(t) + {}^I\dot{\boldsymbol{r}}_{d_2}(t)\big). \tag{69} $$

Due to the fact that the desired trajectories in each of the binocular image planes do not have pointwise correspondence and the two nominal control laws in fact disturb each other, appropriate constraints must be employed before integrating these two feedback control laws for the purpose of driving $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ to zero simultaneously.[20] Let

$$ \boldsymbol{P} = \boldsymbol{I} - \boldsymbol{J}_1^+ \boldsymbol{J}_1 \tag{70} $$

be the projector onto the null space of the Jacobian $\boldsymbol{J}_1$ whose pseudoinverse $\boldsymbol{J}_1^+ \triangleq \boldsymbol{J}_1'(\boldsymbol{J}_1\boldsymbol{J}_1')^{-1}$. The control law that drives the encoded error $\boldsymbol{e}_2$ to zero while maintaining control action in the direction not affecting the control law in Eq. (68) is then given by

$$ \boldsymbol{u}(t) = \boldsymbol{P}\,(\boldsymbol{J}_2\boldsymbol{P})^+ \big(-\lambda_2 \boldsymbol{e}_2 + {}^I\dot{\boldsymbol{r}}_{d_2}(t)\big). \tag{71} $$

In the light of Eqs. (68) and (71), the image-based feedback control law capable of driving both encoded errors, $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, to zero exponentially is proposed as

$$ \boldsymbol{u}(t) = \boldsymbol{J}_1^+(\boldsymbol{r})\big(-\lambda_1 \boldsymbol{e}_1(t) + {}^I\dot{\boldsymbol{r}}^*_1(t)\big) + \boldsymbol{P}(\boldsymbol{J}_2\boldsymbol{P})^+ \big(-\lambda_2 \boldsymbol{e}_2 + {}^I\dot{\boldsymbol{r}}_{d_2}(t) - \boldsymbol{J}_2\boldsymbol{J}_1^+\big(-\lambda_1 \boldsymbol{e}_1(t) + {}^I\dot{\boldsymbol{r}}^*_1(t)\big)\big). \tag{72} $$

Specifically, the system modeled by differential kinematics in Eqs. (66) and (67), when closed by the feedback connection in Eqs. (72), appears to be exponentially stable. That is,

$$ \dot{\boldsymbol{e}}_1(t) = -\lambda_1 \boldsymbol{e}_1 \tag{73} $$

$$ \dot{\boldsymbol{e}}_2(t) = -\lambda_2 \boldsymbol{e}_2. \tag{74} $$

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

Moreover, note that

$$\boldsymbol{e}(t) = \begin{bmatrix} \boldsymbol{e}_1(t) \\ \boldsymbol{e}_2(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_{d_1}(t) \\ \boldsymbol{e}_{d_2}(t) + \boldsymbol{e}_{b_2}(t) \end{bmatrix} \tag{75}$$

where

$$\boldsymbol{e}_{d_n}(t) \triangleq {}^I\boldsymbol{r}_n(t) - {}^I\boldsymbol{r}_n^*(t), \quad \boldsymbol{e}_{b_n}(t) \triangleq {}^I\boldsymbol{r}_n^* - {}^I\boldsymbol{r}_{d_n}(t), \ \ n = 1, 2. \tag{76}$$

It follows directly from Assumption 3.1 and Eq. (12) that

$$\boldsymbol{e}_{d_1}(t), \boldsymbol{e}_{d_2}(t), \text{ and } \boldsymbol{e}_{b_2}(t) \to 0 \quad \text{as} \quad t \to \infty. \tag{77}$$

Therefore,

$${}^I\boldsymbol{r}_n(t) \to {}^I\boldsymbol{r}_n^*(t), \quad n = 1, 2, \quad \text{as} \quad t \to \infty \tag{78}$$

and thus

$$\boldsymbol{r}_n(t) \to \boldsymbol{r}_n^*(t), \quad n = 1, 2, \quad \text{as} \quad t \to \infty \tag{79}$$

which is equivalent to say that the 3-D trajectory following task has been accomplished with precision.

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

$$\boldsymbol{u}(t) = \boldsymbol{J}_{1_q}^+({}^I\boldsymbol{r})\left(-\lambda_1\boldsymbol{e}_1(t) + {}^I\dot{\boldsymbol{r}}_1^*(t)\right) + \boldsymbol{P}_q(\boldsymbol{J}_{2_q}\boldsymbol{P}_q)^+ \left(-\lambda_2\boldsymbol{e}_2 + {}^I\dot{\boldsymbol{r}}_{d_2}(t) - \boldsymbol{J}_{2_q}\boldsymbol{J}_{1_q}^+\left(-\lambda_1\boldsymbol{e}_1(t) + {}^I\dot{\boldsymbol{r}}_1^*(t)\right)\right) \tag{80}$$

where $\boldsymbol{J}_{n_q}({}^I\boldsymbol{r}) \triangleq [\frac{\partial G_{n_q}}{\partial \boldsymbol{r}} \circ G_q^{-1}]({}^I\boldsymbol{r})$, $n = 1, 2$, and $\boldsymbol{P}_q \triangleq \boldsymbol{I} - \boldsymbol{J}_{1_q}^+\boldsymbol{J}_{1_q}$. The control law in Eq. (80) can still drive $\boldsymbol{e}$ to zero exponentially provided that the binocular model imprecision is small, which is equivalent to say that G_q is a good enough approximate model of G.

Fig. 2. Experimental setup.

5. Experiments

5.1. Experimental setup

As illustrated in Fig. 2, experiments in real-time conditions are conducted on a smooth 3-D trajectory, which is generated by projecting a laser line onto an arbitrary-shaped acrylic fiber 3-D object. A Pentium IV 2.8 GHz PC running Windows XP for image processing and control law computation is connected via RS-232 serial link to a Mitsubishi industrial robot RV-1A. Two Sony KMTV-63F1ND CCD video cameras, placed 300 mm in front of the robot with 230 mm baseline, are linked to a ADLink RTV-24 PCI image capture board that grab binocular images of the observed scene. The two CCD cameras are particularly configured not to have parallel coordinate systems, and are only roughly calibrated offline to demonstrate the robust performance of the proposed visual feedback control law. The pixel pitch and the focal length of the cameras are 0.0185 and 2.5 mm, respectively. The travelling distance along the 3-D trajectory is around 500 mm. The sample rate of the real-time trajectory following control system is 12 Hz.

5.2. Experimental tasks

Typical offline calibration procedure needs to be performed for only approximate estimates of camera parameters.

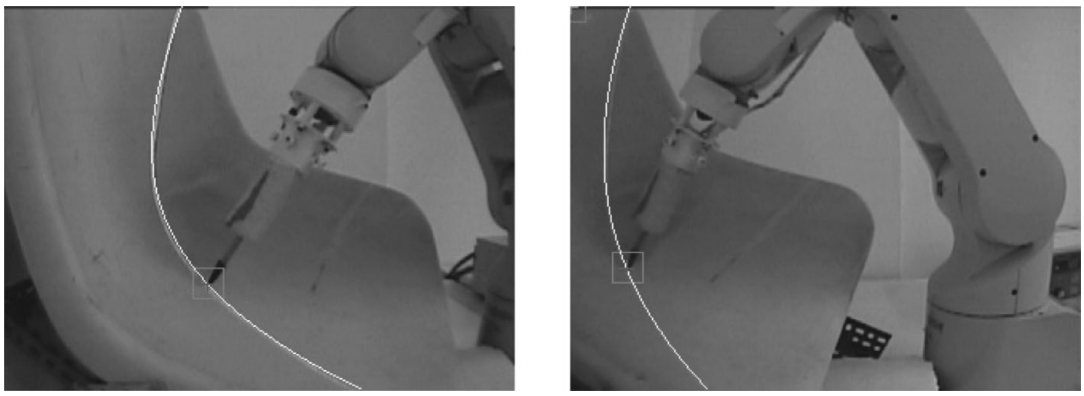

Fig. 3. Observed binocular images from visual detecting and tracking system showing visually determined trajectory and tracking windows for the tool feature.

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

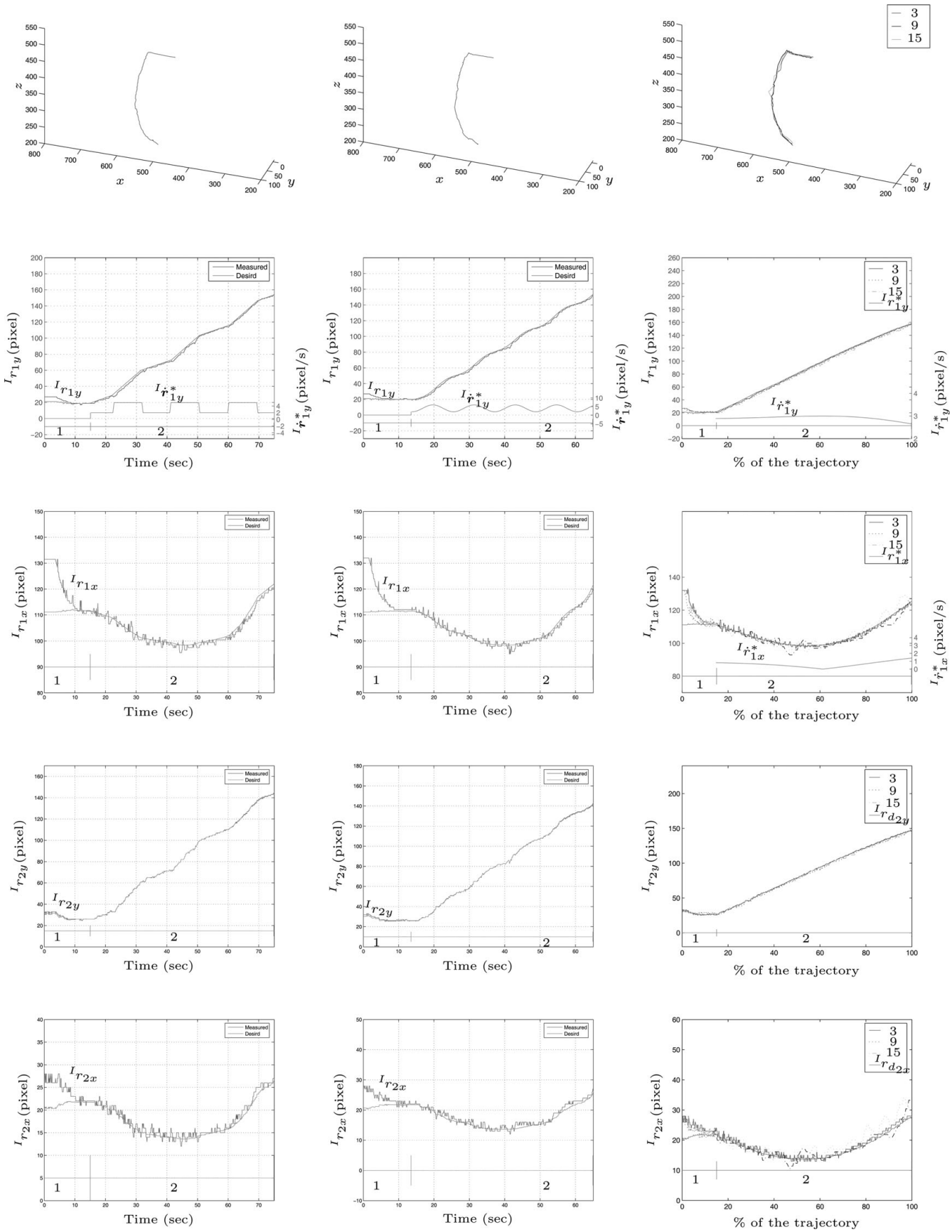

Fig. 4. Positioning and tracking of a manipulator on an unknown smooth 3-D trajectory with set-points velocity command (*left column figures*), time-varying velocity command (*middle column figures*), and three different fixed speed commands (pixel/s)(*right column figures*).

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

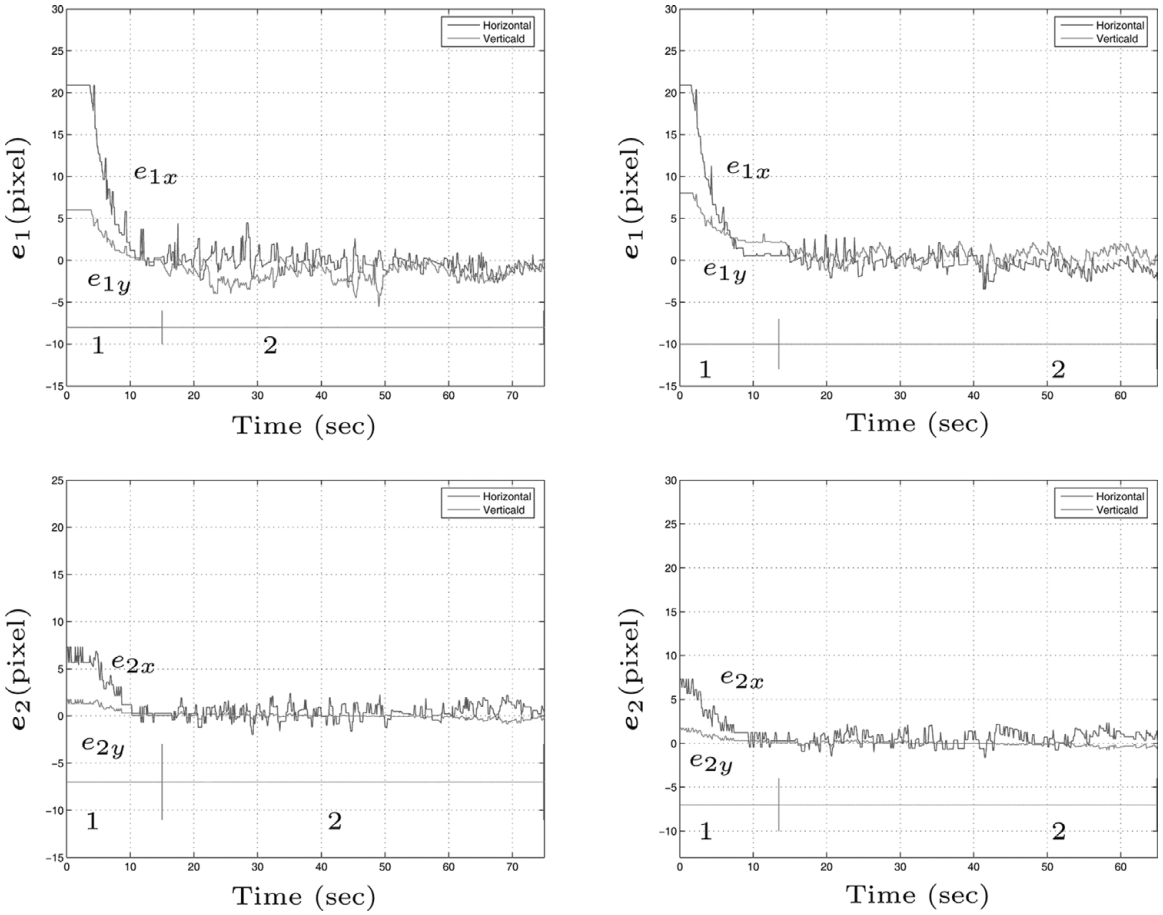

Fig. 5. Positioning and tracking errors of a manipulator on an unknown smooth 3-D trajectory with set-points velocity command (*left column figures*) and time-varying velocity command (*right column figures*).

Meanwhile, the initial position of the tool must be manipulated in open loop in order for the tool feature to appear in the field of view of binocular cameras. Successive stages in the autonomous positioning and tracking control of an arbitrarily smooth 3-D trajectory are as follows.

Stage 1. Autonomously perform *point-to-trajectory positioning* task to drive the tool tip to the arbitrarily smooth 3-D trajectory with precision by visual feedback control law in Eq. (19).

Stage 2. Autonomously perform *point-to-trajectory tracking* task to drive the tool tip to continue tracking the arbitrarily smooth 3-D trajectory in desired speed, either set-points or sinusoidal reference signals, using visual feedback control law in Eq. (80).

Note that by further specifying desired set-point in one of the binocular image plane, the task performed in Stage 1 can drive the tool tip to a particular point in the 3-D trajectory.

5.3. Visual detecting and tracking

During the initialization stage, one can locate and distinguish features of interest in binocular image planes using *thresholding* and *connected component labeling*. Once all features of interest have been extracted, the use of a computational efficient algorithm for continuing tracking moving features is critical in real-time operation. Window-based tracking technique[2] is adopted in the experiments to visually track the robot feature in real time as illustrated in Fig. 3. There are two tracking windows, both 10×10 in pixels, for the tool feature point ${}^I\boldsymbol{r}_1$ and ${}^I\boldsymbol{r}_2$. The positions of these two tracking windows in binocular images are updated based on current positions of those features of interest. As regards the unknown arbitrary target trajectory, *least squares algorithm* is applied to generate polynomial functions, g_1 and g_2, for the observed trajectories in binocular image planes.

5.4. Experimental results

To validate the proposed point-to-trajectory positioning and tracking control approaches, experiments have been performed for the Mitsubishi industrial robotic manipulator to precisely follow an unknown 3-D trajectory generated by projecting a laser line onto an arbitrary-shaped acrylic-fiber object. In Stage 1 for all experiments, the tool tip was driven to the 3-D trajectory to accomplish point-to-trajectory positioning task with precision. For the point-to-trajectory tracking task, two types of vertical velocity commands on the left image plane ${}^I\dot{r}^*_{1y}$, 0.1 Hz square wave (high 6, low 2) and 0.1 Hz sin wave (high 5, low 2), are employed in Stage 2. Meanwhile, three different tracking speed commands specified on the left image plane $\|{}^I\dot{\boldsymbol{r}}^*_1\|$, 3, 9, and 15 pixel/s, are also tested in Stage 2 to see how fast the tracking system could operate while compromising with tracking precision.

Figure 4 shows experimental measurements of a robotic task following a 3-D smooth trajectory with either

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

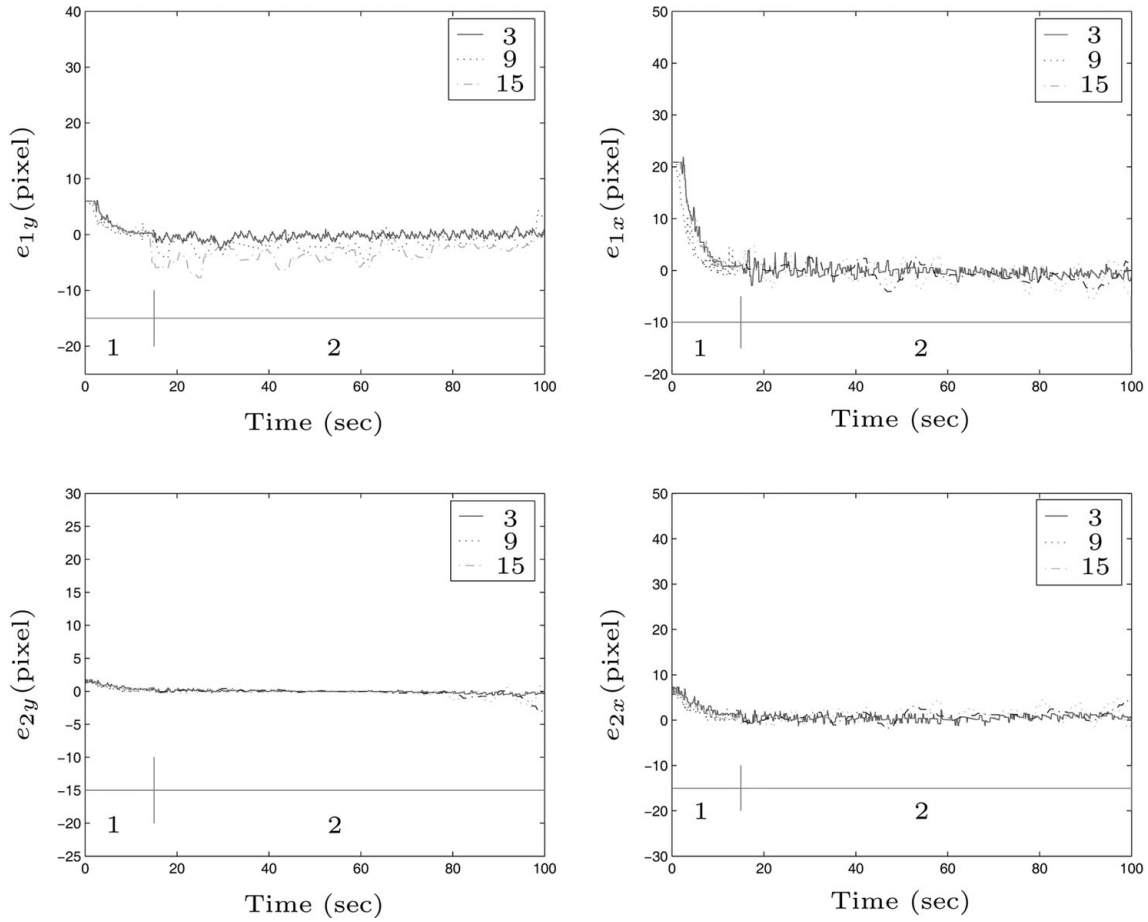

Fig. 6. Positioning and tracking errors of a manipulator on an unknown smooth 3-D trajectory with three different fixed speed commands (pixel/s).

set-points (left column figures) or *sinusoidal* (middle column figures) velocity commands. In these figures, one can see the responses of the vertical image component ${}^{I}r_{ny}, n = 1, 2$, the horizontal image component ${}^{I}r_{nx}, n = 1, 2$, and the resulting 3-D trajectories of the tool tip. With the tracking velocity specified in the left image plane, the positioning and tracking errors for the four directions in binocular image planes, e_{nx} and $e_{ny}, n = 1, 2$, are presented in Fig. 5. Meanwhile, comparison of three different tracking speeds, i.e., $\|{}^{I}\dot{\boldsymbol{r}}_1^*\| = 3, 9,$ and 15 pixel/s, is provided in Fig. 4 (*right column figures*). Figure 6 shows the positioning and tracking errors during these trajectory following tasks. In these experiments, the proposed vision-based control approach has demonstrated satisfactory performance on point-to-trajectory positioning and tracking control even though imprecisely calibrated binocular vision system is employed.

6. Conclusion

In this paper, point-to-trajectory positioning and tracking tasks in uncalibrated environments are considered in a binocular eye-to-hand configuration. Under the proposed approach, a family of 3-D trajectory following tasks, employing approximately calibrated binocular video cameras, can be encoded and accomplished with precision. In particular, an autonomous robotic manipulation system capable of following unknown 3-D trajectories in desired speed has been presented, analyzed, and validated in experiments. The PC-based system enables a robotic manipulator to interact with unknown environments and precisely perform 3-D trajectory following tasks in real time.

Compared with earlier works, this research starts with the selection of novel encoded errors for point-to-trajectory positioning and tracking tasks. These encoded errors being zero imply that the original tasks have been accomplished with precision. In particular, the encoded errors are computed purely based on the binocular vision sensor without assuming pointwise binocular correspondence. This effectively reduces the need to consider complicated epipolar constraints that rely on precise calibration of camera parameters. A simple proportional visual servo controller with feedforward compensation, which is capable of driving the encoded errors to zero, is proposed. The proposed method can be readily employed to existing velocity-controlled robotic manipulators for satisfactory performance in industrial applications. It can also be easily extended to other types of machine instruments and objects for manufacturing purposes.

There are a number of challenging issues that remain unresolved in our future work. In order to further perform dexterous manipulation on a 3-D unknown surface, one needs to actively project certain patterns onto the surface in order to visually determine the pose of the tool relative to the surface in a systematic way. Moreover, one could seek to encode

https://doi.org/10.1017/S0263574707003505 Published online by Cambridge University Press

hybrid force and pose control tasks using multicamera vision and low-cost single-axis force sensors. More importantly, such a seemingly novel approach should further allow robotic manipulators to perform hybrid tasks in unstructured real world with precision.

Acknowledgments

This work was supported by the National Science Council of Taiwan, R.O.C., under grants NSC-90-2213-E-027-026 and NSC-94-2213-E-027-006.

References

1. A. Castaño and S. A. Hutchinson, "Visual compliance: Task-directed visual servo control," *IEEE Trans. Robot. Autom.* **10**(3), 334–342 (Jun. 1994).
2. S. A. Hutchinson, G. D. Hager and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.* **12**(5), 651–670 (Oct. 1996).
3. H.-P. Huang and N. H. McClamroch, "Time-optimal control for a robotic contour following problem," *IEEE Trans. Robot. Autom.* **4**(2), 140–149 (Apr. 1988).
4. J. Baeten and J. De Schutter, "Hybrid vision/force control at corners in planar robotic-contour following," *IEEE/ASME Trans. Mechatron.* **7**(2), 143–151 (Jun. 2002).
5. J. F. Reyes and L. E. Chiang, "Image-to-space path planning for a scara manipulator with single color camera," *Robotica* **21**(3), 245–254 (Jun. 2003).
6. W.-C. Chang, "Hybrid force and vision-based contour following of planar robots," *J. Intell. Robot. Syst.* **47**(3), 215–237 (Nov. 2006).
7. Y. Shen, D. Sun, Y.-H. Liuand and K. Li, "Asymptotic trajectory tracking of manipulators using uncalibrated visual feedback," *IEEE Trans. Mechatron.* **8**(1), 87–98 (Mar. 2003).
8. Di Xiao, B. K. Ghosh, Ning Xi and T. J. Tarn, "Sensor-based hybrid position/force control of a robot manipulator in an uncalibrated environment," *IEEE Trans. Control Syst. Technol.* **8**(4), 635–645 (Jul. 2000).
9. M. R. Akella, "Vision-based adaptive tracking control of uncertain robot manipulators," *IEEE Trans. Robot.* **21**(4), 747–753 (Aug. 2005).
10. J. A. Gangloff and M. F. de Mathelin, "Visual servoing of a 6-dof manipulator for unknown 3-D profile following," *IEEE Trans. Robot. Autom.* **18**(4), 511–520 (Aug. 2002).
11. A. Bettini, P. Marayong, S. Lang, A. M. Okamura and G. D. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Trans. Robotics* **20**(6), 953–966 (Dec. 2004).
12. W.-C. Chang, Vision-Based Control of Uncertain Systems *Ph.D. Thesis* (New Haven, CT: Yale University, Dec. 1997).
13. W.-C. Chang, J. P. Hespanha, A. S. Morse and G. D. Hager, "Task Re-encoding in Vision-Based Control Systems", *Proceedings of IEEE Conference on Decision and Control*, San Diego, California **1** (Dec. 1997) pp. 48–53.
14. W.-C. Chang and A. S. Morse, "Six Degree-of-Freedom Task Encoding in Vision-Based Control Systems," *Proceedings of the 14th World Congress of International Federation of Automatic Control*, Beijing, China **1** (Jul. 1999) pp. 311–316.
15. J. P. Hespanha, Z. Dodds, G. D. Hager and A. S. Morse, "What tasks can be performed with an uncalibrated stereo vision system?" *Int. J. Comput. Vis.* **35**(1), 65–85 (1999).
16. B. K. P. Horn. *Robot Vision* 8th ed. *The MIT Electrical Engineering and Computer Science Series* (McGraw-Hill, New York, 1986).
17. W.-C. Chang, "Precise positioning of binocular eye-to-hand robotic manipulators," *J. Intell. Robot. Syst.* (Feb. 27, 2007) (published online).
18. O. Faugeras *Three-Dimensional Computer Vision: A Geometric Viewpoint* (MIT Press, Cambridge, Massachusetts, 1993).
19. E. D. Sontag. "*Mathematical Control Theory*, *Texts in Applied Mathematics* **6** (Springer-Verlag, New York, 1990).
20. B. Siciliano and J.-J. E. Slotine, "A general framework for managing multiple tasks in highly redundant robotic systems," *Proceedings of IEEE International Conference on Advanced Robotics* (May 1991) pp. 1211–1216.