

リアプノフ法で保証された3次元追跡 Eye-Vergence ビジュアルサーボ実験の周波数応答

○于 福佳、松本 紘明、矢納 陽、見浪 護（岡山大学）
宋 薇（上海大学）

Frequency Response Experiments of Lyapunov-guaranteed 3-D Visual Servoing Experiments with Eye-Vergence

*Fujia Yu Hiroaki Matsumoto Akira Yanou Mamoru Minami (Okayama University)
Wei Song (Shanghai University)

Abstract– In this paper, we use on-line 1-step genetic algorithm (1-step GA) to recognize the 6-D object in a much bigger space comparing with other recognition methods, then we guarantee the convergence of this method by Lyapunov theorem. But visual servoing methods for hand-eye configuration have been presented so far seems to be vulnerable for tracking ability since it may lose a moving target. Our proposal to solve this problem is that the controller for visual servoing of the hand and the eye-vergence should be separated independantly based on decoupling each other. Base on this prerequisite the eye-vergence system to track target object in camera in view sight (trackability) can be much faster than conventional visual servoing with fixed cameras.

Key Words: Visual servoing, 1-step GA, Lyapunov theorem

1 Introduction

Comparing with the visual servoing systems whose cameras are static to the world, the hand-eye visual servoing systems are easy to become unstable because of the disturb from the motion of the end-effector. Otherwise, the hand-eye visual servoing systems can change its angle of view easily. In our research, we use the hand-eye system with two cameras.

On the other hand, a fixed-hand-eye system has some disadvantages, making the observing ability deteriorated depending on the relative geometry of the camera and the target. Such as: the robot cannot observe the object well when it is near the cameras (Fig. 1 (a)), small intersection of the possible sight space of the two cameras (Fig. 1 (b)), and the image of the object cannot appear in the center of both cameras, so we could not get clear image information of target and its periphery, reducing the pose measurement accuracy (Fig. 1 (c)). To solve the problems above, in this paper, we give the cameras an ability to rotate themselves to see target at center of the images. There is no research using such rotatable hand-eye system as far as we know. Thus it is possible to change the pose of the cameras in order to observe the object better, as it is shown in Fig. 2, enhancing the measurement accuracy in trigonometric calculation and peripheral distortion of camera lens by observing target at the center of lens. Moreover, recent researches on visual servoing are limited generally in a swath of tracking an object while keeping a certain constant distance^{3), 4), 5)}. But the final objective of visual servoing lies in approaching the end-effector to a target and then work on it, like grasping. In this case, the desired relation between the cameras and the object is time varying, so such rotational camera system in Fig. 2 is required to keep suitable viewpoint all the time during the visual servoing application.

In visual servoing application, it is important to keep the

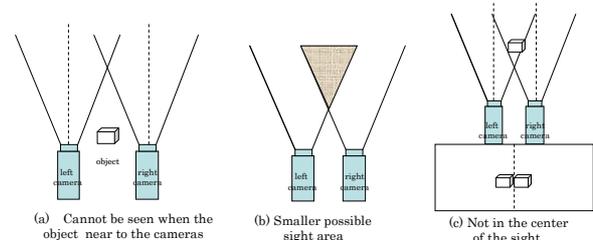


Fig. 1: Disadvantage of fix camera system

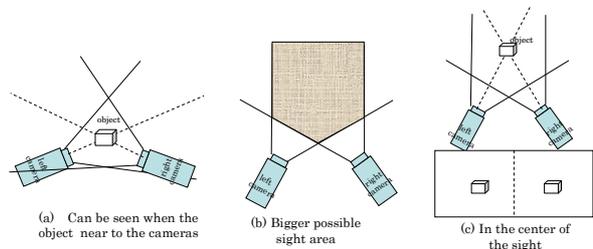


Fig. 2: Advantage of Eye-vergence system

object in the visual eye sight to make the visual feed back to keep stable closed-loop dynamical motion. If the camera lose the sight of target, its pose cannot be measured, that means, the visual feedback is cut, and the robot may fall in some unexpected motion, being dangerous. As it is shown in Fig. 3 (a), in visual servoing system the cameras can keep staring at the object at first in (a), but when the target moves so fast that the manipulator can not catch up the speed of the target because of the big mass of whole manipulator itself, then the object may disappear in the sight of the cameras, resulting in that the visual feedback of the system is cut as shown in (b), loosing feedback information that appears most dangerous. So in visual servoing system it is very important to keep the camera tracking

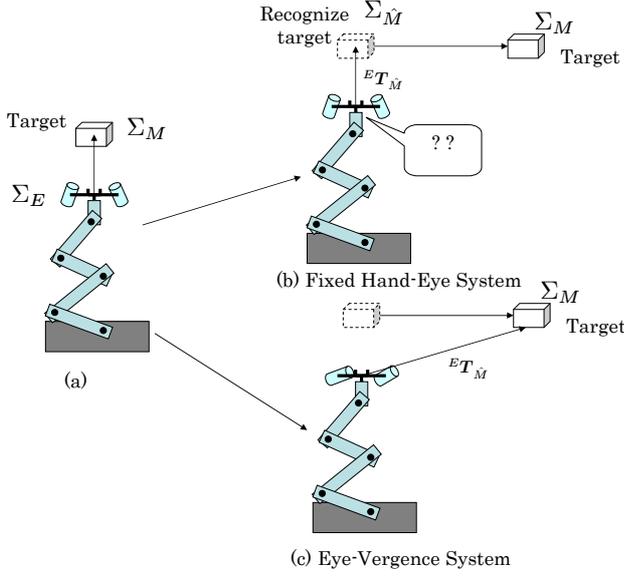


Fig. 3: Dynamics advantage of Eye-vergence system

the target. A system with high tracking ability also has better security and validity. To realize this stable tracking ability against quick and unknown motion of the target, we propose to control the cameras and the manipulator separately. Because of the small mass and inertia moment of the cameras, it can track the target better, as in Fig. 3 (c), like animals tracks target with eye motion before rotate their heads to the target to improve dynamical tracking ability.

To evaluate the observation of the camera, we put forward a concept of trackability. This concept has been used in ⁶⁾, where trackability is defined as a kinematic function of singular value of Jacobian matrix connecting hand's velocities and angular joint velocities, ignoring the relationship between the hand and the target objects, including the both dynamical motion of the target and the manipulator, which seems to be essential for evaluating the eye-vergence visual servoing. Then we define a new concept of trackability to evaluate our visual servoing.

As shown in Fig.4, the proposed method includes two loops: a loop for conventional visual servoing that direct a manipulator toward a target object and an inner loop for active motion of binocular camera for accurate and broad observation of the target object. We set relatively high gain to the eye-vergence controller to put the priority to the 3D pose tracking to improve the system trackability.

2 On-line Evolutionary Recognition

2.1 3-D Model-based Matching

We use a model-based matching method to recognize a target object in a 3-D searching area. A solid models is located in Σ_E , its position and orientation are determined by six parameters, $\psi = [r^T, \epsilon^T]^T$, where $r = [x, y, z]^T$, $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3]^T$. Here, the target's orientation is represented by unit quaternion ¹⁾, which has an advantage that can represent the orientation of a rigid body without sin-

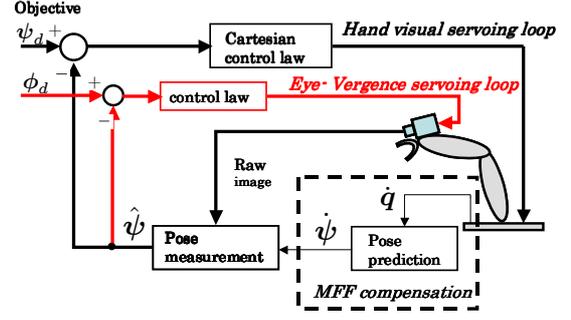


Fig. 4: Hand & Eye Visual servo system

gularities, when $-\pi < \theta < \pi$ (θ is defined below). The unit quaternion, is defined as

$$Q = \{\eta, \epsilon\}, \quad (1)$$

where

$$\eta = \cos \frac{\theta}{2}, \quad \epsilon = \sin \frac{\theta}{2} \mathbf{k}, \quad (2)$$

here, $\mathbf{k}(\|\mathbf{k}\| = 1)$ is the rotation axis and θ is the rotation angle. η is called the scalar part of the quaternion while ϵ is called the vector part of the quaternion. They are constrained by

$$\eta^2 + \epsilon^T \epsilon = 1. \quad (3)$$

In (3) η can be calculated by ϵ , so we just use three parameters ϵ to represent an orientation.

The left and right input images from the stereo cameras are directly matched by the left and right searching models, which are projected from 3-D model onto 2-D image plane. The matching degree of the model to the target can be estimated by a correlation function between them as $F(\psi)$ by using the color information of the target. Please refer to ⁹⁾ for a detailed definition of $F(\psi)$. When the searching models fit to the target objects being imaged in the right and left images, $F(\psi)$ gives the maximum value. Therefore the 3-D object's position/orientation measurement problem can be converted to a searching problem of ψ that maximizes $F(\psi)$. We solve this optimization problem by 1-step GA method that will be explained in the next section.

2.2 GA-based On-line Recognition "1-step GA"

Theoretically optimal pose $\psi_{max}(t)$ that gives the highest peak of $F(\psi(t))$ is defined as

$$\psi_{max}(t) = \{\psi(t) \mid \max_{\psi \in \mathbf{L}} F(\psi(t))\} \quad (4)$$

where \mathbf{L} represents 6-DoF searching space of $x, y, z, \epsilon_1, \epsilon_2, \epsilon_3$.

An individual of GA is defined as $\psi_i^j(t)$, which means the i -th gene ($i = 1, 2, \dots, p$) in the j -th generation, to search $\psi_{max}(t)$. Denote $\psi_{max}^{GA}(t)$ to be the maximum among the p genes of $\psi_i^j(t)$ in GA process,

$$\psi_{max}^{GA}(t) = \{\psi_i^j(t) \mid \max_{\psi_i^j \in \mathbf{L}} F(\psi_i^j(t))\}. \quad (5)$$

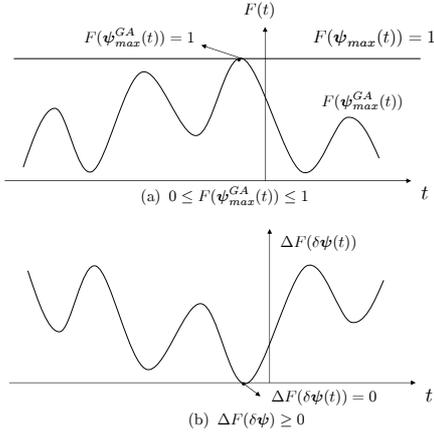


Fig. 5: $\Delta F(\delta\psi(t))$ is positive definite

In fact we cannot always guarantee the best individual of GA $\psi_{max}^{GA}(t)$ should coincide with the theoretically optimal pose $\psi_{max}(t)$, because the number of GA's individuals is not infinite. The difference between $\psi_{max}(t)$ and $\psi_{max}^{GA}(t)$ is denoted as

$$\delta\psi(t) = \psi_{max}(t) - \psi_{max}^{GA}(t). \quad (6)$$

And the difference between $F(\psi_{max}(t))$ and $F(\psi_{max}^{GA}(t))$ is denoted as

$$\Delta F(\delta\psi(t)) = F(\psi_{max}(t)) - F(\psi_{max}^{GA}(t)), \quad (7)$$

Since $F(\psi_{max}(t)) \geq F(\psi_{max}^{GA}(t))$, we have

$$\Delta F(\delta\psi(t)) \geq 0. \quad (8)$$

Based on the definition of $\Delta F(\delta\psi(t))$ in (7), in this research, we let GA work in the following way:

- (a) GA evolves to minimize $\Delta F(\delta\psi(t))$.
- (b) The elitist individual of GA is preserved at every generation (elitist gene preservation strategy).
- (c) $\psi_{max}^{GA}(t)$ does keep the same value in the evolving when the evolved new gene with different value gives the same value of ΔF .

Here, we present two assumptions.

[Assumption 1] $\Delta F(\delta\psi(t))$ is positive definite.

This means the distribution of $F(\psi(t))$ satisfies $\Delta F(\delta\psi(t)) = 0$ if and only if $\delta\psi(t) = 0$, which indicates $\Delta F(\delta\psi(t)) = 0$ has a sole minimum at $\delta\psi(t) = 0$ over the searching space \mathbf{L} , even though ΔF is multi-peak distribution having peaks and bottoms with limited number. When the model overlap the target object in the image, then the situation can make ΔF have the sole minimum in \mathbf{L} . $0 \leq F(\psi(t)) \leq 1$, since $F(\psi(t))$ is normalized to be less than 1 and negative value to be set as zero by a definition of correlation function $F(\psi(t))$ ². So the fitness function is always less than 1 except only one point which means the $\psi_{max}^{GA}(t)$ can express the target object's pose, as shown in Fig. 5(a). From (7), we can see when $\psi_{max}^{GA}(t) = 1$, $\Delta F(\delta\psi(t)) = 0$ (Fig. 5(b)), which means

that only in this case, $\psi_{max}^{GA}(t)$ can express the actual pose of the target object.

[Assumption 2] $\dot{F}(\psi_{max}^{GA}(t)) \geq 0$.

This means GA evolves itself to get a bigger fitness function value ($\dot{F}(\psi_{max}^{GA}(t)) > 0$) or keep a same value ($\dot{F}(\psi_{max}^{GA}(t)) = 0$). It is not only an assumption but also the character of GA if the target object is static, because the elitist individual is preserved in every generation of GA. However, when the target object is moving, $\dot{F}(\psi_{max}^{GA}(t)) \geq 0$ will indicate that the convergence speed to the target in the dynamical images should be faster than the moving speed of the target object. Furthermore, with the pose tracking in dynamic scene being input at a certain video rate, this assumption means that $F(\psi_{max}^{GA}(t))$ have the tendency of approaching to $F(\psi_{max}(t))$, and $\psi_{max}^{GA}(t)$ moves toward $\psi_{max}(t)$ in each period of the input image, or keeps a distance to $\psi_{max}(t)$. Since in this paper we think that the object's motion is enough slow comparing the calculation speed of GA's evolving to find $F(\psi_{max}^{GA}(t))$ from the view point that the one image be input every input video period and evolving iterations in input video period are enough to catch up with the $F(\psi_{max}^{GA}(t))$ being stationary during the input video period.

Differentiating (7) by time t, we have

$$\Delta \dot{F}(\delta\psi(t)) = \dot{F}(\psi_{max}(t)) - \dot{F}(\psi_{max}^{GA}(t)). \quad (9)$$

We defined $F(\psi_{max}(t)) = 1$ representing that the true pose of the target object gives the highest peak. Therefore, the time differentiation of $F(\psi_{max}(t))$ will be $\dot{F}(\psi_{max}(t)) = 0$. Thus, from (9) and [Assumption 2], we have

$$\Delta \dot{F}(\delta\psi(t)) = -\dot{F}(\psi_{max}^{GA}(t)) \leq 0. \quad (10)$$

$\psi_{max}^{GA}(t)$ represents current best GA solution. [Assumption 2] means GA can change its best gene $\psi_{max}^{GA}(t)$ to always reduce the value of ΔF regardless of dynamic image or static one, which indicates that the convergence speed to the target in the dynamically continuous images should be faster than the moving speed of the target object.

We cannot guarantee that the above two assumptions always hold, since they depend on some factors such as object's shape, object's speed, definition of $F(\psi(t))$, parameters of GA and viewpoint for observing, lightening environment, computer's performance et al. However, we can make efforts to improve the environment and correlation function and so on. Providing above two assumptions be satisfied, (8) and (10) hold, then $\Delta F(\delta\psi(t))$ is so-called Lyapunov function. The objective here is to verify that $\delta\psi(t)$ asymptotically stable, resulting in it converges to $\mathbf{0}$ by using the Lyapunov function of $\Delta F(\delta\psi(t))$, meaning $\psi_{max}^{GA}(t) \rightarrow \psi_{max}(t)$, ($t \rightarrow \infty$), and the following shows how to verify it.

Since $\Delta \dot{F}(\delta\psi(t))$ is only negative semi-definite, in the view of LaSalle theorem, $\delta\psi(t)$ asymptotically converges to the invariant set of the solutions $\delta\psi$ satisfying $\Delta \dot{F}(\delta\psi(t)) = 0$. Considering the following expression,

$$\Delta \dot{F}(\delta\psi(t)) = \frac{\partial \Delta F}{\partial \delta\psi} \cdot \delta\dot{\psi} = 0, \quad (11)$$

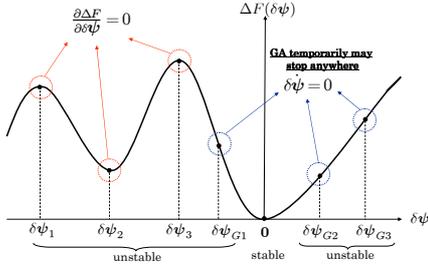


Fig. 6: The invariant set of the solutions of $\Delta\dot{F}(\delta\psi(t)) = 0$.

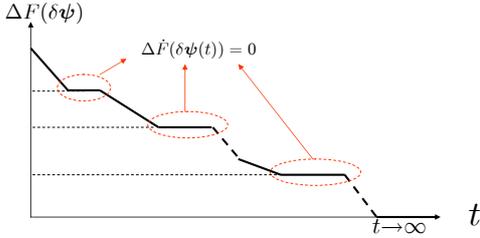


Fig. 7: The changing of $\Delta F(\delta\psi(t))$ with respect to time t in the whole GA's evolution.

the first part $\partial\Delta F/\partial\delta\psi$ describes partial differentiation of ΔF with respect to $\delta\psi$, implying steepest descending direction of ΔF in the space of $\delta\psi$; the second part $\delta\dot{\psi}$ describes the difference between the moving speed of the target object and the evolution speed of the best gene of GA, by the definition in (6).

Equation (11) shows the invariant set of the solutions of $\Delta\dot{F}(\delta\psi(t)) = 0$ includes (1): P_1 , the solution set of $\partial\Delta F/\partial\delta\psi = 0$; (2): P_2 , the solution set of $\delta\dot{\psi} = 0$; and (3): P_3 , the solution set satisfying $\partial\Delta F/\partial\delta\psi \neq \mathbf{0}$, $\delta\dot{\psi} \neq \mathbf{0}$, but their inner product is 0.

As shown in Fig. 6, P_1 includes the points of $\delta\psi$ that give the local maximum or minimum values of the function ΔF including $\mathbf{0}$. The number of these points is finite by [Assumption 1] denoted by p , that is

$$P_1 = \{\mathbf{0}, \delta\psi_1, \delta\psi_2, \dots, \delta\psi_{p-1}\}. \quad (12)$$

Concerning (1), the evolving process of GA may stay temporarily at the same ΔF value. If the target object is static, it means the best gene of GA stops at some moments for the reason that the limited individuals of GA could not improve a current solution that gives a smaller fitness function value ΔF during some generations. And when the target object is moving, $\delta\dot{\psi} = 0$ means at these moments that the evolution speed of the best gene of GA is equal to the moving speed of the target object, by (6). The number of these points is assumed to be possibly finite, denoted by q . Thus, we describe the set of P_2 as

$$P_2 = \{\mathbf{0}, \delta\psi_{G1}, \delta\psi_{G2}, \dots, \delta\psi_{G(q-1)}\}. \quad (13)$$

Notice that there is another solution set of $\delta\dot{\psi}$: P_3 . In this case, the vector of $\partial\Delta F/\partial\delta\psi$ is vertical to the vector of $\delta\dot{\psi}$ since the calculation $(\Delta F/\delta\psi) \cdot \delta\dot{\psi}$ in (11) means inner cross product, which means GA evolves in the direction that keeps a same fitness function value ΔF . This

GA's evolution way is forbidden in this research for the GA work rule (c) that we have stated above. Then, P_3 is null. So the invariant set that $\delta\psi(t)$ asymptotically converges to is

$$P = P_1 \cup P_2. \quad (14)$$

Here, $\delta\psi_1, \delta\psi_2, \dots, \delta\psi_{p-1}$ in P_1 are all unstable since $\delta F(\delta\psi_i) > 0$ ($i = 1, 2, \dots, p-1$), and only $\delta\psi = \mathbf{0}$ is stable from [Assumption 1], since when $t \rightarrow \infty$ there should always remain the possibility to get out of local maximum/minimum of $\delta\psi_1 \dots \delta\psi_{p-1}$. And all the points in P_2 except the point $\mathbf{0}$ are unstable because GA has possibility to get out of these points by its evolving nature. Therefore, $\mathbf{0}$ is the only stable point in the invariant set of P , that is, $\delta\psi(t)$ will finally converges to $\mathbf{0}$. The image of the changing of $\Delta F(\delta\psi(t))$ with respect to time t in the whole GA's evolution is shown in Fig.7.

The above verification shows $\delta\psi(t) \rightarrow \mathbf{0}$, which means

$$\psi_{max}^{GA}(t) \rightarrow \psi_{max}(t), \quad (t \rightarrow \infty) \quad (15)$$

Let t_ϵ denotes a convergence time, then

$$|\delta\psi(t)| = |\psi_{max}(t) - \psi_{max}^{GA}(t)| \leq \epsilon, \quad (\epsilon > 0, t \geq t_\epsilon) \quad (16)$$

In (16), ϵ is tolerable extent that can be considered as a observing error. Thus, it is possible to realize real-time optimization, because $\psi_{max}^{GA}(t)$ can be assumed to be in the vicinity of the theoretically optimal $\psi_{max}(t)$ after t_ϵ .

Above discussion is under the condition of continuous time. Here, when we consider evolution time of each generation of GA denoted by Δt . The GA's evolving process is described as

$$\psi_i^j(t) \xrightarrow{\text{evolve}} \psi_i^{j+1}(t + \Delta t). \quad (17)$$

Obviously, this time-discrete evolution with the interval of time Δt may enlarge the recognition error $\delta\psi(t)$. Should this undesirable influence of Δt be considered, the tolerable pose error ϵ will expand to ϵ' as,

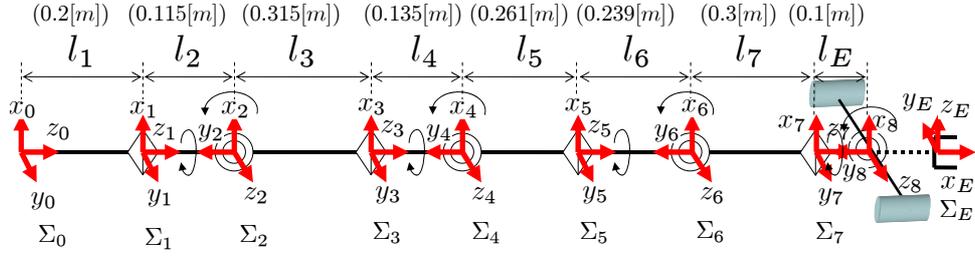
$$|\delta\psi(t)| \leq \epsilon', \quad (\epsilon' > \epsilon > 0). \quad (18)$$

Since the GA process to recognize the target's pose at the current time is executed at least one time with the period of Δt as the current quasi-optimal pose $\psi_{max}^{GA}(t)$ is output synchronously, we named this on-line recognition method as "1-step GA". We have confirmed that the above real-time optimization problem could be solved by "1-step GA" through several experiments to recognize swimming fish ⁷⁾, human face ³⁾ and rectangular solid block ⁸⁾.

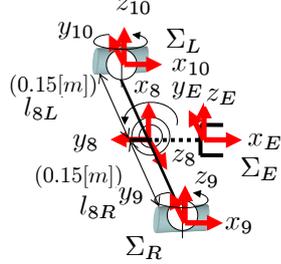
3 Hand & Eye Visual Servoing

3.1 Experiment circumstance

The Mitsubishi PA-10 robot arm is a 7-DoF robot arm manufactured by Mitsubishi Heavy Industries, as shown in Fig. 8 (a). Two rotatable cameras with two pan angles and one sharing tilt angle mounted on the end-effector are FCB-1X11A manufactured by Sony Industries (Fig.



(a) Sketch map of the manipulator



(b) Sketch map of camera system

Fig. 8: Sketch map of the eye-vergence system

8 (b)). The frame frequency of stereo cameras is set as 33fps. The image processing board, CT-3001, receiving the image from the CCD camera is connected to the DELL WORKSTATION PWS650 (CPU: Xeon, 2.00 GHz) host computer.

3.2 Desired-trajectory generation

As shown in Fig. 9, the world coordinate frame is denoted by Σ_W , the target coordinate frame is denoted by Σ_M , and the desired and actual end-effector coordinate frame is denoted by Σ_{Ed} , Σ_E separately. The desired relative relation between the target and the end-effector is given by Homogeneous Transformation as ${}^{Ed}T_M$, the relation between the target and the actual end-effector is given by ${}^E T_M$, then the difference between the desired end-effector pose Σ_{Ed} and the actual end-effector pose Σ_E is denoted as ${}^E T_{Ed}$, ${}^E T_{Ed}$ can be described by:

$${}^E T_{Ed}(t) = {}^E T_M(t) {}^{Ed} T_M^{-1}(t) \quad (19)$$

(19) is a general deduction that satisfies arbitrary object motion ${}^W T_M(t)$ and arbitrary visual servoing objective ${}^{Ed} T_M(t)$. However, the relation ${}^E T_M(t)$ is only observed by cameras using the on-line model-based recognition method and 1-step GA^{3), 9)}. Let $\Sigma_{\hat{M}}$ denote the detected object, there always exist an error between the actual object Σ_M and the detected one $\Sigma_{\hat{M}}$. So in visual servoing, (19) will be rewritten based on $\Sigma_{\hat{M}}$ that includes the error ${}^M T_{\hat{M}}$, as

$${}^E T_{Ed}(t) = {}^E T_{\hat{M}}(t) {}^{Ed} T_{\hat{M}}^{-1}(t), \quad (20)$$

where ${}^E T_{\hat{M}} = {}^E T_M$ determined by the given visual servoing objective. Differentiating (20) with respect to time yields

$${}^E \dot{T}_{Ed}(t) = {}^E \dot{T}_{\hat{M}}(t) {}^{\hat{M}} T_{Ed}(t) + {}^E T_{\hat{M}}(t) {}^{\hat{M}} \dot{T}_{Ed}(t), \quad (21)$$

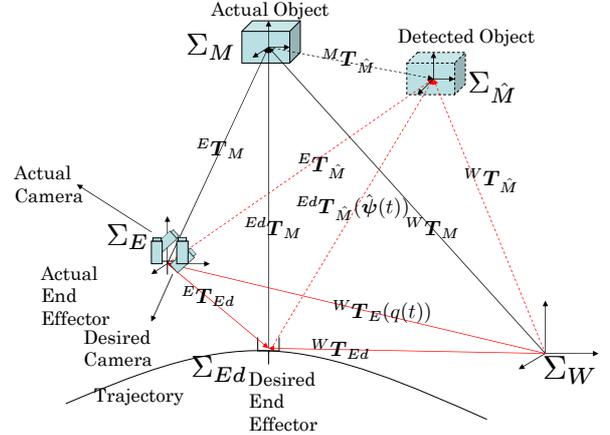


Fig. 9: Motion of the end-effector and object

Differentiating Eq. (21) with respect to time again

$${}^E \ddot{T}_{Ed}(t) = {}^E \ddot{T}_{\hat{M}}(t) {}^{\hat{M}} T_{Ed}(t) + 2 {}^E \dot{T}_{\hat{M}}(t) {}^{\hat{M}} \dot{T}_{Ed}(t) + {}^E T_{\hat{M}}(t) {}^{\hat{M}} \ddot{T}_{Ed}(t), \quad (22)$$

Where ${}^{\hat{M}} T_{Ed}$, ${}^{\hat{M}} \dot{T}_{Ed}$, ${}^{\hat{M}} \ddot{T}_{Ed}$ are given as the desired visual servoing objective. ${}^E T_{\hat{M}}$, ${}^{Ed} T_{\hat{M}}$, ${}^E \dot{T}_{\hat{M}}$ can be observed by cameras. As shown in Fig. 9, there are two errors that we have to decrease to $\mathbf{0}$ in the visual servoing process. First one is the error between the actual object and the detected one ${}^M T_{\hat{M}}$, and the other is the error between the desired end-effector and the actual one ${}^E T_{Ed}$. In our research, the error of ${}^M T_{\hat{M}}$ is decreased by on-line recognition method of 1-step GA, MFF compensation method and the eye-vergence camera system, and the error of ${}^E T_{Ed}$ can be decreased by the hand visual servoing controller.

3.3 Hand & Eye Visual Servoing Controller

The block diagram of our proposed hand & eye-vergence visual servoing controller is shown in Fig. 4. The hand-visual servoing is the outer loop. A detailed block diagram of hand visual servoing control is depicted in Fig.10. Based on the above analysis of the desired-trajectory generation, the desired hand velocity ${}^W \dot{r}_d$ is calculated as,

$${}^W \dot{r}_d = \mathbf{K}_{P_p} {}^W \mathbf{r}_{E,Ed} + \mathbf{K}_{V_p} {}^W \dot{r}_{E,Ed}, \quad (23)$$

where ${}^W \mathbf{r}_{E,Ed}$, ${}^W \dot{r}_{E,Ed}$ are given by transforming ${}^E T_{Ed}$ and ${}^E \dot{T}_{Ed}$ from Σ_E to Σ_W . \mathbf{K}_{P_p} and \mathbf{K}_{V_p} are positive definite matrix to determine PD gain.

The desired hand angular velocity ${}^W \boldsymbol{\omega}_d$ is calculated as,

$${}^W \boldsymbol{\omega}_d = \mathbf{K}_{P_o} {}^W \mathbf{R}_E^E \Delta \boldsymbol{\epsilon} + \mathbf{K}_{V_o} {}^W \boldsymbol{\omega}_{E,Ed}, \quad (24)$$

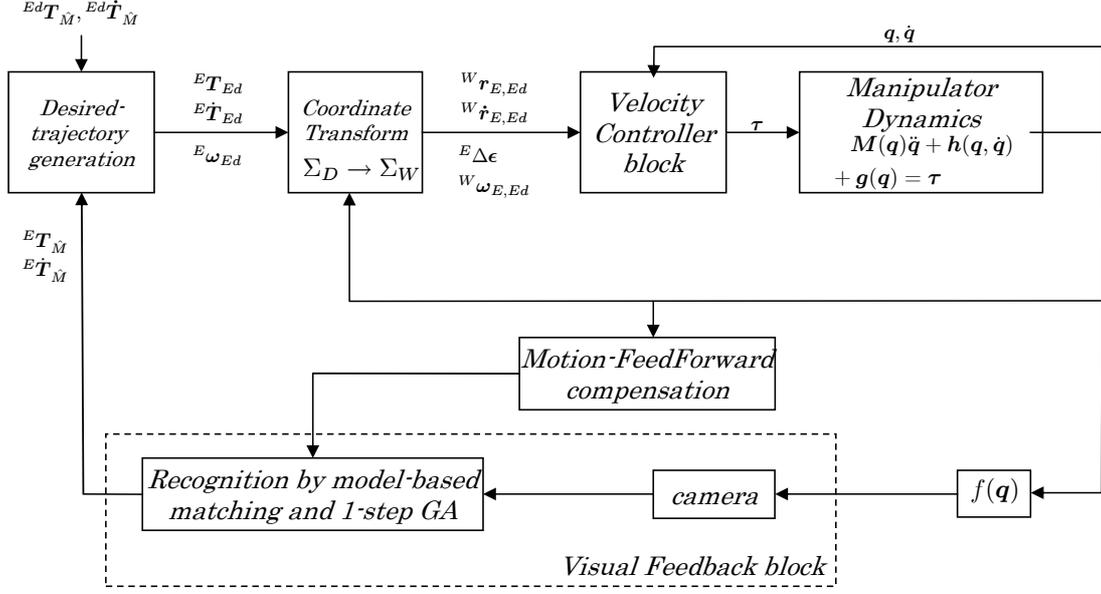


Fig. 10: Block diagram of the hand visual servoing system

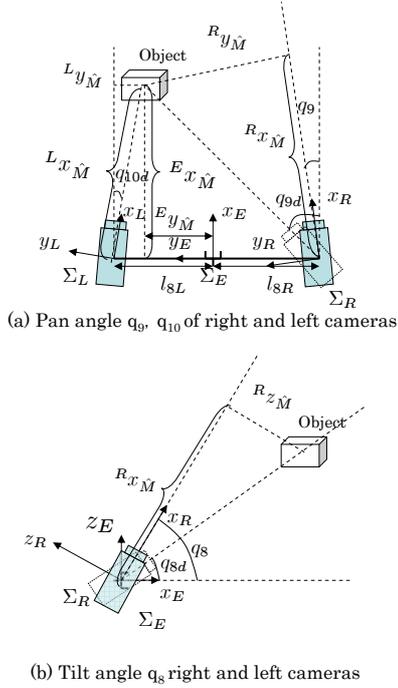


Fig. 11: Calculation of tilt and pan angles

where $E \Delta \epsilon$ is the quaternion error that from the recognition result directly, and $W \omega_{E,Ed}$ can be calculated by transforming $E T_{Ed}$ and $E \dot{T}_{Ed}$ from Σ_E to Σ_W . Also, K_{P_o} and K_{V_o} are suitable feedback matrix gains.

The desired joint variable \dot{q}_d is obtained by

$$\dot{q}_d = J^+(q) \begin{bmatrix} W \dot{\tau}_d \\ W \omega_d \end{bmatrix}. \quad (25)$$

where $J^+(q)$ is the pseudo inverse matrix of $J(q)$, and $J^+(q) = J^T(JJ^T)^{-1}$. The hardware control system of

the velocity-based servo system of PA10 is expressed as

$$\tau = K_{SP}(\dot{q}_d - \dot{q}) + K_{SI} \int_0^t (\dot{q}_d - \dot{q}) dt \quad (26)$$

where K_{SP} and K_{SI} are symmetric positive definite matrix to determine PI gain.

The eye-vergence visual servoing is the inner loop of the visual servoing system shown in Fig. 4. In this paper, we use two pan-tilt cameras for eye-vergence visual servoing. Here, the positions of cameras are supposed to be fixed on the end-effector. For camera system, q_8 is tilt angle, q_9 and q_{10} are pan angles, and q_8 is common for both cameras. As it is shown in Fig. 11, $E x_M, E y_M, E z_M$ express position of the detected object in the end-effector coordinate. The desired angle of the camera joints are calculated by:

$$q_{8d} = \text{atan2}(E z_M, E x_M) \quad (27)$$

$$q_{9d} = \text{atan2}(l_{8R} + E y_M, E x_M) \quad (28)$$

$$q_{10d} = \text{atan2}(-l_{8L} + E y_M, E x_M) \quad (29)$$

where $l_{8L} = l_{8R} = 120[\text{mm}]$ that is the camera location. We set the center line of the camera as the x axis of each camera coordinate. Then the controller of eye-visual servoing is given by

$$\dot{q}_8 = K_{P_T}(q_{8d} - q_8) + K_{D_T}(\dot{q}_{8d} - \dot{q}_8), \quad (30)$$

$$\dot{q}_9 = K_{P_C}(q_{9d} - q_9) + K_{D_C}(\dot{q}_{9d} - \dot{q}_9), \quad (31)$$

$$\dot{q}_{10} = K_{P_C}(q_{10d} - q_{10}) + K_{D_C}(\dot{q}_{10d} - \dot{q}_{10}). \quad (32)$$

where $K_{P_T}, K_{D_T}, K_{P_C}, K_{D_C}$ are positive control gain.

4 experiment of hand & eye-vergence visual servoing

To verify the effectiveness of the proposed hand & eye visual servoing system, we conduct the experiment of visual servoing to a 3D marker that is composed of a red

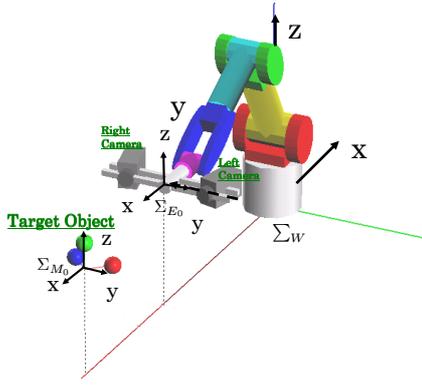


Fig. 12: Object and the visual-servoing system

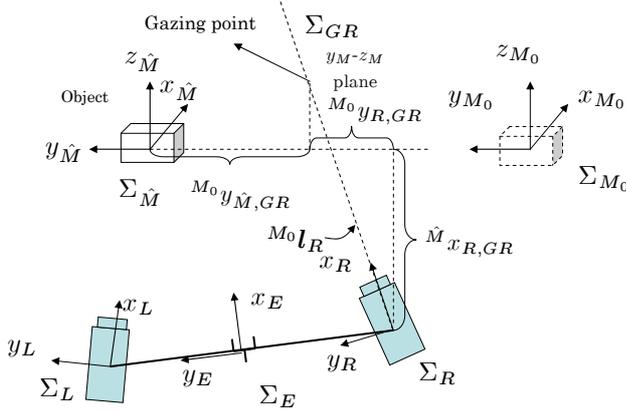


Fig. 13: Cameras' and End Effector's gazing point

ball, a green ball and a blue ball as Fig. 12. The radiuses of these three balls are set as 30[mm].

4.1 experiment condition

The initial hand pose is defined as Σ_{E_0} , while the initial object pose is defined as Σ_{M_0} , and the homogeneous transformation matrix from Σ_W to Σ_{M_0} is:

$${}^W\mathbf{T}_{M_0} = \begin{bmatrix} 0 & 0 & -1 & -1410[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 355[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (33)$$

The target object move according to the following time function

$${}^{M_0}\psi_M = [0, {}^{M_0}y_M(t), 0, 0, 0, 0]^T \quad (34)$$

$${}^{M_0}y_M(t) = -200 \sin(\omega t)[mm] \quad (35)$$

here, ω is the angular velocity of the motion of the object.

The relation between the object and the desired end-effector is set as:

$${}^{E_d}\psi_M = [800[mm], 0, 0, 0, 0, 0] \quad (36)$$

Here, to compare the trackability of the eye-vergence system and fixed camera system, we define a concept of gazing point. As it is shown in Fig. 13 the intersection of the gazing line of right camera and the $y_{M_0}-z_{M_0}$ plane is defined as the gazing point. The relative relation between

Σ_{M_0} and Σ_R is given by Homogeneous Transformation as ${}^{M_0}\mathbf{T}_R$, ${}^{M_0}\mathbf{T}_R$ conclude the rotation matrix ${}^{M_0}\mathbf{R}_R$ and the position vector ${}^{M_0}\mathbf{p}_R$, and the rotation matrix ${}^{M_0}\mathbf{R}_R$ can be written as $[{}^{M_0}\mathbf{x}_R, {}^{M_0}\mathbf{y}_R, {}^{M_0}\mathbf{z}_R]$. The direction of ${}^{M_0}\mathbf{l}_R$ in Fig. 13 is same to the direction of x_R , and ${}^{M_0}\mathbf{l}_R$ can be expressed as:

$${}^{M_0}\mathbf{l}_R = {}^{M_0}\mathbf{p}_R + k_R {}^{M_0}\mathbf{x}_R \quad (37)$$

here k_R is a scalar variable. The gazing point of the right camera expressed in Σ_{M_0} is ${}^{M_0}\mathbf{p}_{GR} = [0, {}^{M_0}y_{GR}, {}^{M_0}z_{GR}]^T$. For ${}^{M_0}\mathbf{l}_R = {}^{M_0}\mathbf{p}_{GR}$ in x direction, $({}^{M_0}\mathbf{p}_R)_x + k_R({}^{M_0}\mathbf{x}_R)_x = 0$. And usually $({}^{M_0}\mathbf{x}_R)_x \neq 0$, k_R can be calculated by $k_R = -({}^{M_0}\mathbf{p}_R)_x / ({}^{M_0}\mathbf{x}_R)_x$, and the y, z coordinate of the gazing point in Σ_{M_0} can be calculated by:

$${}^{M_0}y_{GR} = ({}^{M_0}\mathbf{p}_R)_y + k_R({}^{M_0}\mathbf{x}_R)_y \quad (38)$$

$${}^{M_0}z_{GR} = ({}^{M_0}\mathbf{p}_R)_z + k_R({}^{M_0}\mathbf{x}_R)_z \quad (39)$$

The target object's motion is given by (34), (35), because the motion of the target object M is parallel to the y_{M_0} , we take ${}^{M_0}y_M(t)$ as the input, and the gazing point of the right camera ${}^{M_0}y_{GR}(t)$ as the response. And define the concept of trackability by the frequency response of ${}^{M_0}y_{GR}(t)$, the trackability of the left camera can be defined in the same way.

4.2 Experiment Results

In Fig. 14, we show the result of our experiment, we change the ω in (34) from 0.01 to 1.256 and get the data of the gazing point of the cameras of eye-vergence system and the gazing point of the end-effector of the fixed camera system separately, we do the experiment 10 times at every ω we selected, and use the average delay time and the amplitude to draw the frequency response curve. The amplitude-frequency curve and the delay frequency curve are shown in Fig.14 (a) and Fig.14 (b). Here, for the fixed camera $A = {}^{M_0}y_M(t)$, $B = {}^{M_0}y_{GE}(t)$. For the right camera of Eye-Vergence system $A = {}^{M_0}y_M(t)$, $B = {}^{M_0}y_{GR}(t)$, for the left camera $A = {}^{M_0}y_M(t)$, $B = {}^{M_0}y_{GL}(t)$. In this two figures the abscissa axes are ω . In (a), (b), we sign the angular velocity when $\omega = 0.314, 0.628, 1.256$, and show the position of the gazing point of the cameras in eye-vergence experiment and the position of the gazing point of the end-effector in fixed camera experiment in (c), (d), (e). From (a), (b) we can see that the fixed-camera system cannot track the target object when ω is faster than 0.628 so in (e), there is only the data of the cameras and the target object. From Fig. 14 (a) we can see the data of the cameras and the end-effector all become smaller as ω increases but the curve of the fixed camera system is always below the curves of the cameras, which means that delay of the fixed camera system is bigger than the eye-vergence system, from (b) the the curve of the fixed camera system is also below the curves of the cameras, we can see that the amplitude of the eye-vergence system is more closed to the target object than the fixed camera system, so from (a) and (b) we can get the conclu-

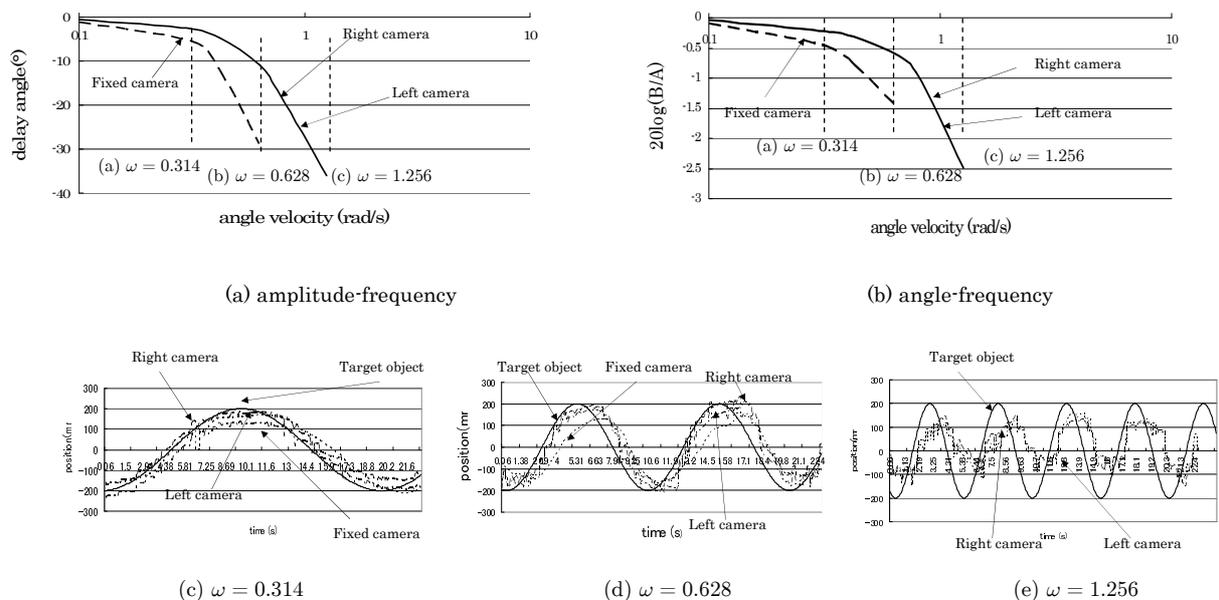


Fig. 14: Comparison of Cameras' and End-effector's Trackabilities by Frequency Response

sion that the eye-vergence system has the better trackability than the fixed-camera system. To be understood easily, we show the position of the gazing point of the cameras in eye-vergence experiment and the position of the gazing point of the end-effector in fixed camera experiment in (c), (d), (e). and $M_0 \dot{y}_M(0) = -200[mm/s]$, while the target object moved from static, so it cannot move stably at first, we use the data when the target object's motion became stable. From the figures it is also easily to see that comparing with the fixed camera system, the eye-vergence system can track the target object better.

5 Conclusion

In this paper, we put forward a new concept to evaluate the observation ability on a moving object of visual servoing system, and introduce the importance of it. Then we introduce the recognition method using "1-step GA" and our eye-vergence system. To check the trackability of eye-vergence visual servoing system, we did some experiments of eye-vergence system and fixed camera system separately. In the experiments we compared the amplitude-frequency and phase-frequency curves of the gazing point of the cameras of the eye-vergence system and the fixed camera system by moving object in different angular velocities, and get the conclusion that the trackability and stability of the eye-vergence system is better than that of the fixed-camera system by analyze the experiment data.

- 1) B.Siciliano and L.Villani: *Robot Force Control*, ISBN 0-7923-7733-8.
- 2) L. Reznik and V. Kreinovich Eds, *Soft Computing in Measurement and Information Acquisition*, ISBN 3-540-00246-4.
- 3) W. Song, M. Minami, Y. Mae and S. Aoyagi, "On-line Evolutionary Head Pose Measurement by Feed-

forward Stereo Model Matching", IEEE Int. Conf. on Robotics and Automation (ICRA) 4394/4400 (2007).

- 4) Omar Tahri and Francois Chaumette, "Point-Based and Region-Based Image Moments for Visual Servoing of Planar Objects", IEEE Tran. on Robotics, vol. 21, no. 6, Dec (2005).
- 5) Tarek Hamel and Robert Mahony, "Visual Servoing of an Under-Actuated Dynamic Rigid-Body System: An Image-Based Approach", IEEE Trans. on Robotics and Automation, VOL. 18, NO. 2, APRIL (2002).
- 6) Tomoyuki Shiozaki, Toshiyuki Murakami. "Trackability Based Motion Control in Mobile Hand-Eye System". SICE-ICASE International Joint Conference 5304/5309 (2006).
- 7) H. Suzuki, M. Minami, "Visual Servoing to catch fish Using Global/local GA Search", IEEE/ASME Transactions on Mechatronics, Vol.10, Issue 3, 352/357 (2005).
- 8) Wei Song, Mamoru Minami, and Seiji Aoyagi, "Evolutionary Pose Measurement by Stereo Model Matching" Journal of Advanced Computational Intelligence and Intelligent Informatics(JACIII), vol.9, No.2, 150/158 (2005)
- 9) W. Song, M. Minami, S. Aoyagi, "On-line Stable Evolutionary Recognition Based on Unit Quaternion Representation by Motion-Feedforward Compensation", International Journal of Intelligent Computing in Medical Sciences and Image Processing (IC-MED) Vol. 2, No. 2, 127/139 (2007).