Visual servoing by Lyapunov-guaranteed stable on-line 6-D pose tracking

Fujia Yu, Wei Song, Mamoru Minami, Akira Yanou and Mingcong Deng

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. In application field, we use this 1-step GA method into visual-servoing simulation in which by a 10-link manipulator hand and eye-vergence system, where Lyapunov method is utilized to guarantee the visual tracking convergence to a moving target and visual servoing to the target is also confirmed by the Lyapunov method.**

I. INTRODUCTION

Tasks in which visual information are used to direct a manipulator toward a target object are referred to visual servoing in [1], [2]. This field is the fusion of many areas, such as kinematics, dynamics, image recognition, and control theory. This paper deals with problems of the real-time 3-D pose (position and orientation) recognition of a target for visual servoing and the convergence proof of this recognition algorithm.

There is a variety of approaches for 3D target object pose estimation, and they can be classified into three general categories: (1) feature-based, (2) appearance-based, and (3) model-based. We use model-based method in our research. The matching degree of the model to the target can be estimated by a function, whose maximum value represents the best matching and can be solved by GA, using the matching function as a fitness function. An advantage of our method is that we use a 3D solid model which enables it to possess six degree of freedom (DOF), both the position and orientation, without following hindrances. In other methods like feature-based recognition, the pose of the target object should be determined by a set of image points, which makes it need a very strict camera calibration. Moreover, searching the corresponding points in Stereo-vision camera images is also complicated and time consuming, e.g., [3].

GA is well known as a method for solving parameter optimization problems [4]. The GA-based scene recognition method described here can be designated as "evolutionary recognition method", since for every step of the GA's evolution, it struggles to perform the recognition of a target in the input image. To recognize a target by CCD camera in real-time, and to avoid time lag waiting for the convergence to a target, we used GA in such manner that only one generation is processed over newly input image, which we called "1-Step GA". In this way, the GA searching process and the convergence to the target does not consist in one image but the convergence is achieved in the sequence of the input images to recognize it with real-time manner.

In this report, we present a hand & eye-vergence dual visual servoing system with a stability analysis of Lyapunov method, guaranteeing that both the tracking pose errors of hand and eye-vergence converge to zero. There is no research on eye-vergence visual servoing stability analysis, as far as we know.

Fujia Yu, Mamoru Minami, Akira Yanou, Mingcong Deng are with Graduate School of Natural Science and Technology, Okayama University Tsushimanaka3-1-1 , Okayama, JAPAN. { yufujia, minami, yanou, deng}@suri.sys.okayama-u.ac.jp Wei Song is with University of Shanghai songwei5726@hotmail.com

II. ON-LINE EVOLUTIONARY RECOGNITION

A. 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, $\boldsymbol{\psi} = [\boldsymbol{r}^T, \boldsymbol{\epsilon}^T]^T$, where $\boldsymbol{r} = [x, y, z]^T$, $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \epsilon_3]^T$. Here, the target's orientation is represented by unit quaternion [5], which has an advantage that can represent the orientation of a rigid body without singularities, when $-\pi < \theta < \pi$ (θ is defined below). The unit quaternion, is defined as

$$\boldsymbol{Q} = \{\eta, \boldsymbol{\epsilon}\}, \tag{1}$$

where

$$\eta = cos\frac{\theta}{2},\ \boldsymbol{\epsilon} = sin\frac{\theta}{2}\boldsymbol{k}, \tag{2}$$

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

$$\eta^2 + \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = 1. \tag{3}$$

In (3) η can be calculated by $\boldsymbol{\epsilon}$, so we just use three parameters $\boldsymbol{\epsilon}$ 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 $F(\boldsymbol{\psi})$ by using the color information of the target. Please refer to [6] for a detailed definition of $F(\boldsymbol{\psi})$. When the searching models fit to the target objects being imaged in the right and left images, $F(\boldsymbol{\psi})$ gives the maximum value. Therefore the 3-D object's position/orientation measurement problem can be converted to a searching problem of $\boldsymbol{\psi}$ that maximizes

 SY0009/10/0000-0154

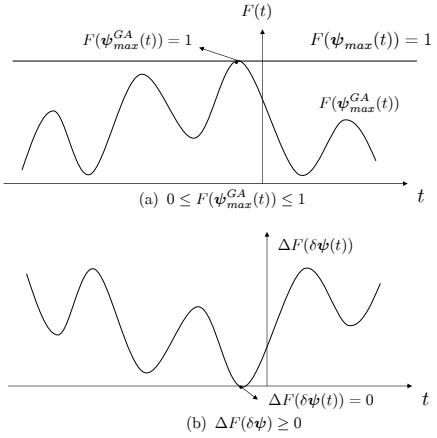

Fig. 1. $\Delta F(\delta\boldsymbol{\psi}(t))$ is positive definite

$F(\boldsymbol{\psi})$. We solve this optimization problem by 1-step GA method that will be explained in the next section.

B. GA-based On-line Recognition "1-step GA"

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

$$\boldsymbol{\psi}_{max}(t) = \left\{\boldsymbol{\psi}(t) \mid \max_{\boldsymbol{\psi}\in\boldsymbol{L}} F(\boldsymbol{\psi}(t))\right\} \tag{4}$$

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

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

$$\boldsymbol{\psi}_{max}^{GA}(t) = \left\{\boldsymbol{\psi}_i^j(t) \mid \max_{\boldsymbol{\psi}_i^j\in\boldsymbol{L}} F(\boldsymbol{\psi}_i^j(t))\right\}. \tag{5}$$

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

$$\delta\boldsymbol{\psi}(t) = \boldsymbol{\psi}_{max}(t) - \boldsymbol{\psi}_{max}^{GA}(t). \tag{6}$$

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

$$\Delta F(\delta\boldsymbol{\psi}(t)) = F(\boldsymbol{\psi}_{max}(t)) - F(\boldsymbol{\psi}_{max}^{GA}(t)), \tag{7}$$

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

$$\Delta F(\delta\boldsymbol{\psi}(t)) \geq 0. \tag{8}$$

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

- **(a)** GA evolves to minimize $\Delta F(\delta\boldsymbol{\psi}(t))$.
- **(b)** The elitist individual of GA is preserved at every generation (elitist gene preservation strategy).
- **(c)** $\boldsymbol{\psi}_{max}^{GA}(t)$ 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] $\Delta F(\delta\boldsymbol{\psi}(t))$ is positive definite.

This means the distribution of $F(\boldsymbol{\psi}(t))$ satisfies $\Delta F(\delta\boldsymbol{\psi}(t)) = 0$ if and only if $\delta\boldsymbol{\psi}(t) = 0$, which indicates $\Delta F(\delta\boldsymbol{\psi}(t)) = 0$ has a sole minimum at $\delta\boldsymbol{\psi}(t) = 0$ over the searching space $\boldsymbol{L}$, 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 $\boldsymbol{L}$. $0 \leq F(\boldsymbol{\psi}(t)) \leq 1$, since $F(\boldsymbol{\psi}(t))$ is normalized to be less than 1 and negative value to be set as zero by a definition of correlation function $F(\boldsymbol{\psi}(t))$ ([7]). So the fitness function is always less than 1 except only one point which means the $\boldsymbol{\psi}_{max}^{GA}(t)$ can express the target object's pose, as shown in Fig. 1(a). From (7), we can see when $\boldsymbol{\psi}_{max}^{GA}(t) = 1$, $\Delta F(\delta\boldsymbol{\psi}(t)) = 0$ (Fig. 1(b)), which means that only in this case, $\boldsymbol{\psi}_{max}^{GA}(t)$ can express the actual pose of the target object.

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

This means GA evolves itself to get a bigger fitness function value ($\dot{F}(\boldsymbol{\psi}_{max}^{GA}(t)) > 0$) or keep a same value ($\dot{F}(\boldsymbol{\psi}_{max}^{GA}(t)) = 0$). 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, $\dot{F}(\boldsymbol{\psi}_{max}^{GA}(t)) \geq 0$ 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 $F(\boldsymbol{\psi}_{max}^{GA}(t))$ have the tendency of approaching to $F(\boldsymbol{\psi}_{max}(t))$, and $\boldsymbol{\psi}_{max}^{GA}(t)$ moves toward $\boldsymbol{\psi}_{max}(t)$ in each period of the input image, or keeps a distance to $\boldsymbol{\psi}_{max}(t)$. Since in this paper we think that the object's motion is enough slow comparing the calculation speed of GA's evolving to find $F(\boldsymbol{\psi}_{max}^{GA}(t))$ 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 $F(\boldsymbol{\psi}_{max}^{GA}(t))$ being stationary during the input video period.

Differentiating (7) by time t, we have

$$\Delta\dot{F}(\delta\boldsymbol{\psi}(t)) = \dot{F}(\boldsymbol{\psi}_{max}(t)) - \dot{F}(\boldsymbol{\psi}_{max}^{GA}(t)). \tag{9}$$

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

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

$\boldsymbol{\psi}_{max}^{GA}(t)$ represents current best GA solution. [Assumption 2] means GA can change its best gene $\boldsymbol{\psi}_{max}^{GA}(t)$ 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

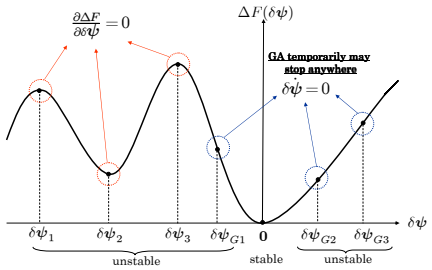

Fig. 2. The invariant set of the solutions of $\Delta\dot{F}(\delta\boldsymbol{\psi}(t)) = 0$.

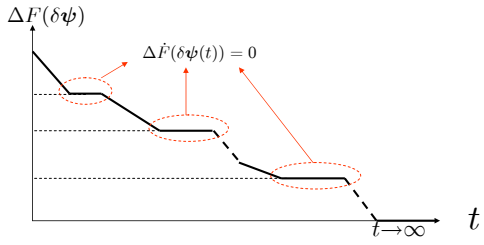

Fig. 3. The changing of $\Delta F(\delta\boldsymbol{\psi}(t)$ with respect to time t in the whole GA's evolution.

object's shape, object's speed, definition of $F(\boldsymbol{\psi}(t))$, 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 $\Delta F(\delta\boldsymbol{\psi}(t))$ is so-called Lyapunov function. The objective here is to verify that $\delta\boldsymbol{\psi}(t)$ asymptotically stable, resulting in it converges to $\mathbf{0}$ by using the Lyapunov function of $\Delta F(\delta\boldsymbol{\psi}(t))$, meaning $\boldsymbol{\psi}_{max}^{GA}(t)\longrightarrow\boldsymbol{\psi}_{max}(t), \quad (t\to\infty)$, and the following shows the way.

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

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

the first part $\partial\Delta F/\partial\delta\boldsymbol{\psi}$ describes partial differentiation of ΔF with respect to $\delta\boldsymbol{\psi}$, implying steepest descending direction of ΔF in the space of $\delta\boldsymbol{\psi}$; the second part $\delta\dot{\boldsymbol{\psi}}$ 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 $\Delta\dot{F}(\delta\boldsymbol{\psi}(t)) = 0$ includes (1): P_1, the solution set of $\partial\Delta F/\partial\delta\boldsymbol{\psi} = 0$; (2): P_2, the solution set of $\delta\dot{\boldsymbol{\psi}} = 0$; and (3): P_3, the solution set satisfying $\partial\Delta F/\partial\delta\boldsymbol{\psi} \neq \mathbf{0}$, $\delta\dot{\boldsymbol{\psi}} \neq \mathbf{0}$, but their inner product is 0.

As shown in Fig. 2, P_1 includes the points of $\delta\boldsymbol{\psi}$ 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

$$P_1 = \{\mathbf{0}, \delta\boldsymbol{\psi}_1, \delta\boldsymbol{\psi}_2, \cdots, \delta\boldsymbol{\psi}_{p-1}\}. \tag{12}$$

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, $\delta\dot{\boldsymbol{\psi}} = 0$ 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

$$P_2 = \{\mathbf{0}, \delta\boldsymbol{\psi}_{G1}, \delta\boldsymbol{\psi}_{G2}, \cdots \delta\boldsymbol{\psi}_{G(q-1)}\}. \tag{13}$$

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

$$P = P_1 \bigcup P_2. \tag{14}$$

Here, $\delta\boldsymbol{\psi}_1, \delta\boldsymbol{\psi}_2, \cdots, \delta\boldsymbol{\psi}_{p-1}$ in P_1 are all unstable since $\delta F(\delta\boldsymbol{\psi}_i) > 0$ $(i = 1, 2, \cdots, p-1)$, and only $\delta\boldsymbol{\psi} = \mathbf{0}$ is stable from [Assumption 1], since when $t \to \infty$, there should always remain the possibility to get out of local maximum/minimum of $\delta\boldsymbol{\psi}_1 \cdots \delta\boldsymbol{\psi}_{p-1}$. 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, $\delta\boldsymbol{\psi}(t)$ will finally converges to $\mathbf{0}$. The image of the changing of $\Delta F(\delta\boldsymbol{\psi}(t))$ with respect to time t in the whole GA's evolution is shown in Fig.3.

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

$$\boldsymbol{\psi}_{max}^{GA}(t)\longrightarrow\boldsymbol{\psi}_{max}(t), \quad (t\to\infty) \tag{15}$$

Let t_ϵ denotes a convergence time, then

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

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

$$\boldsymbol{\psi}_i^j(t) \stackrel{\text{evolve}}{\longrightarrow} \boldsymbol{\psi}_i^{j+1}(t + \Delta t). \tag{17}$$

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

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

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 [8] and human face [9].

III. Dynamics of Hand and Eye

The equation of motion of a robot is

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{\tau} \tag{19}$$

here we define

$$\boldsymbol{q}_E = \begin{bmatrix} q_1 \\ \vdots \\ q_7 \end{bmatrix}, \boldsymbol{q}_R = \begin{bmatrix} q_8 \\ q_9 \end{bmatrix}, \boldsymbol{q}_L = \begin{bmatrix} q_8 \\ q_{10} \end{bmatrix}, \boldsymbol{v}_E = \begin{bmatrix} \dot{\boldsymbol{r}}_E \\ \boldsymbol{\omega}_E \end{bmatrix}$$

the compensation of robot's dynamics the outputs:

$$\boldsymbol{\tau} = \boldsymbol{M}(\boldsymbol{q})\boldsymbol{\phi} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) \tag{20}$$

Taking the compensation (20) into (19), closed loop dynamics is:

$$\ddot{\boldsymbol{q}} = \boldsymbol{\phi} \tag{21}$$

$$\boldsymbol{\phi}_E = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_7 \end{bmatrix}, \boldsymbol{\phi}_R = \begin{bmatrix} \phi_8 \\ \phi_9 \end{bmatrix}, \boldsymbol{\phi}_L = \begin{bmatrix} \phi_8 \\ \phi_{10} \end{bmatrix},$$

On the other hand the position compensation of the end-effector is:

$$\boldsymbol{a}_p = \ddot{\boldsymbol{p}}_d + \boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE} \tag{22}$$

here

$$\Delta\boldsymbol{p}_{dE} = \boldsymbol{p}_d - \boldsymbol{p}_E \tag{23}$$

here, $\boldsymbol{p}_d$ and $\boldsymbol{p}_E$ means the desired position and actual position of the end-effector.

The orientation compensation of the end-effector is:

$$\boldsymbol{a}_o = \dot{\boldsymbol{\omega}}_d + \boldsymbol{K}_{D_o}\Delta\boldsymbol{\omega}_{dE} + \boldsymbol{K}_{P_o}\boldsymbol{R}_E{}^E\boldsymbol{\epsilon}_{dE} \tag{24}$$

and

$$\Delta\boldsymbol{\omega}_{dE} = \boldsymbol{\omega}_d - \boldsymbol{\omega}_E \tag{25}$$

here, $\boldsymbol{\omega}_d$ and $\boldsymbol{\omega}_E$ means the desired angular velocity and actual angular velocity of the end-effector, where $\boldsymbol{K}_{D_o}$ and $\boldsymbol{K}_{P_o}$ are suitable positive definitive feedback matrix gains. ${}^E\boldsymbol{\epsilon}_{dE}$ means the orientation error between the desired and the actual end-effector orientation, while the letter in the top left corner express the coordinate where the vector or the rotation matrix is expressed in. When there is no letter on the top left corner it means the vector or the matrix is expressed in the world frame. ${}^E\boldsymbol{\epsilon}_{dE}$ can be calculated by (2).

$\boldsymbol{\phi}_E$ can be calculated by:

$$\boldsymbol{\phi}_E = \boldsymbol{J}_E^+(\boldsymbol{q}_E)\Big(\begin{bmatrix} \boldsymbol{a}_p \\ \boldsymbol{a}_o \end{bmatrix} - \dot{\boldsymbol{J}}_E(\boldsymbol{q}_E, \dot{\boldsymbol{q}}_E)\dot{\boldsymbol{q}}_E\Big) \tag{26}$$

here we define $\boldsymbol{a}^T = [\boldsymbol{a}_p^T, \boldsymbol{a}_o^T]^T$

The compensation of the camera is:

$${}^E\boldsymbol{a}_{co} = {}^E\dot{\boldsymbol{\omega}}_d + \boldsymbol{K}_{D_o}\Delta^E\boldsymbol{\omega}_{dc} + \boldsymbol{K}_{P_o}{}^E\boldsymbol{R}_c{}^c\boldsymbol{\epsilon}_{dc} \tag{27}$$

${}^c\boldsymbol{\epsilon}_{dc}$ which means the orientation error between the desired and the actual camera orientation expressed in the camera coordinate, can be also calculated by (2).

Take the right camera as an example, $\boldsymbol{\phi}_R$ can be calculated as:

$$\boldsymbol{\phi}_R = {}^E\boldsymbol{J}_R^+(\boldsymbol{q}_R)({}^E\boldsymbol{a}_{Ro} - {}^E\dot{\boldsymbol{J}}_R(\boldsymbol{q}_R)\dot{\boldsymbol{q}}_R) \tag{28}$$

here, $\boldsymbol{J}_E$ is the Jacobian matrix from the world coordinate to the end-effector, ${}^E\boldsymbol{J}_R$ is the Jacobian matrix from the end-effector to the right camera, the equations will be introduced more detailed in the section 5.

Take (26) and (28) into (21) we have

$$\boldsymbol{a} = \boldsymbol{J}_E(\boldsymbol{q}_E)\ddot{\boldsymbol{q}}_E + \dot{\boldsymbol{J}}_E(\boldsymbol{q}_E)\dot{\boldsymbol{q}}_E \tag{29}$$

$${}^E\boldsymbol{a}_{Ro} = {}^E\boldsymbol{J}_R(\boldsymbol{q}_R)\ddot{\boldsymbol{q}}_R + {}^E\dot{\boldsymbol{J}}_R(\boldsymbol{q}_R)\dot{\boldsymbol{q}}_R \tag{30}$$

so

$$\begin{bmatrix} \ddot{\boldsymbol{p}}_E \\ \dot{\boldsymbol{\omega}}_E \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_p \\ \boldsymbol{a}_o \end{bmatrix} \tag{31}$$

$${}^E\dot{\boldsymbol{\omega}}_R = {}^E\boldsymbol{a}_{Ro} \tag{32}$$

${}^E\dot{\boldsymbol{\omega}}_L = {}^E\boldsymbol{a}_{Lo}$ can be deduced in the similar way.

Submit (22), (24) into (31), then we have

$$\ddot{\boldsymbol{p}}_E = \ddot{\boldsymbol{p}}_d + \boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE} \tag{33}$$

$$\dot{\boldsymbol{\omega}}_E = \dot{\boldsymbol{\omega}}_d + \boldsymbol{K}_{D_o}\Delta\boldsymbol{\omega}_{dE} + \boldsymbol{K}_{P_o}\boldsymbol{R}_E{}^E\boldsymbol{\epsilon}_{dE} \tag{34}$$

submit (27) into (32), we have

$${}^E\dot{\boldsymbol{\omega}}_c = {}^E\dot{\boldsymbol{\omega}}_d + \boldsymbol{K}_{D_o}\Delta^E\boldsymbol{\omega}_{dc} + \boldsymbol{K}_{P_o}{}^E\boldsymbol{R}_c{}^c\boldsymbol{\epsilon}_{dc} \tag{35}$$

Finally we get closed loop hand camera motions are,

$$\Delta\ddot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE} = \boldsymbol{0} \tag{36}$$

$$\Delta\dot{\boldsymbol{\omega}}_{dE} + \boldsymbol{K}_{D_o}\Delta\boldsymbol{\omega}_{dE} + \boldsymbol{K}_{P_o}\boldsymbol{R}_E{}^E\boldsymbol{\epsilon}_{dE} = \boldsymbol{0} \tag{37}$$

$$\Delta^E\dot{\boldsymbol{\omega}}_{dc} + \boldsymbol{K}_{D_o}\Delta^E\boldsymbol{\omega}_{dc} + \boldsymbol{K}_{P_o}{}^E\boldsymbol{R}_c{}^c\boldsymbol{\epsilon}_{dc} = \boldsymbol{0} \tag{38}$$

IV. Prove of the convergence of the system by Lyapunov method

Here, we discuss about the convergence of our proposed Hand & eye vergence dual visual servoing system. We invoke a Lyapunov argument, the feed back gains are taken as scalar matrices, i.e. $\boldsymbol{K}_{D_p} = K_{D_p}\boldsymbol{I}$, $\boldsymbol{K}_{P_p} = K_{P_p}\boldsymbol{I}$, $\boldsymbol{K}_{D_o} = K_{D_o}\boldsymbol{I}$ and $\boldsymbol{K}_{P_o} = K_{P_o}\boldsymbol{I}$. Here we assume that the feedback gains of the links are the same.

$$\begin{aligned} \nu = {} & \Delta\boldsymbol{p}_{dE}^T K_{P_p}\Delta\boldsymbol{p}_{dE} + \Delta\dot{\boldsymbol{p}}_{dE}^T\Delta\dot{\boldsymbol{p}}_{dE} \\ & + K_{P_o}((\eta_{dE} - 1)^2 + {}^E\boldsymbol{\epsilon}_{dE}^T{}^E\boldsymbol{\epsilon}_{dE}) + \frac{1}{2}\Delta\boldsymbol{\omega}_{dE}^T\Delta\boldsymbol{\omega}_{dE} \\ & + K_{P_o}(({}^E\eta_{dc} - 1)^2 + {}^c\boldsymbol{\epsilon}_{dc}^T{}^c\boldsymbol{\epsilon}_{dc}) + \frac{1}{2}\Delta^E\boldsymbol{\omega}_{dc}^T\Delta^E\boldsymbol{\omega}_{dc} \\ & \geq 0 \end{aligned} \tag{39}$$

so

$$\begin{aligned}\dot{\nu} &= -2\Delta\dot{\boldsymbol{p}}_{dE}^T(\Delta\ddot{\boldsymbol{p}}_{dE}+\boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE})\\ &\quad +2K_{P_o}((\eta_{dE}-1)\dot{\eta}_{dE}+{}^E\boldsymbol{\epsilon}_{dE}^T{}^E\dot{\boldsymbol{\epsilon}}_{dE})+\Delta\boldsymbol{\omega}_{dE}^T\Delta\dot{\boldsymbol{\omega}}_{dE}\\ &\quad +2K_{P_o}(({}^E\eta_{dc}-1)^E\dot{\eta}_{dc}+{}^c\boldsymbol{\epsilon}_{dc}^T{}^c\dot{\boldsymbol{\epsilon}}_{dc})\\ &\quad +\Delta^E\boldsymbol{\omega}_{dc}^T\Delta^E\dot{\boldsymbol{\omega}}_{dc} \qquad (40)\end{aligned}$$

from (36) we can know that

$$\Delta\ddot{\boldsymbol{p}}_{dE}+\boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE}=-\boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} \qquad (41)$$

from the quaternion definition we can know that [5]

$$\dot{\eta}_{dE}=-\frac{1}{2}{}^E\boldsymbol{\epsilon}_{dE}^T\Delta^E\boldsymbol{\omega}_{dE} \qquad (42)$$

and

$${}^E\dot{\eta}_{dc}=-\frac{1}{2}{}^c\boldsymbol{\epsilon}_{dc}^T\Delta^c\boldsymbol{\omega}_{dc} \qquad (43)$$

and

$${}^E\dot{\boldsymbol{\epsilon}}_{dE} = \frac{1}{2}\boldsymbol{E}(\eta_{dE},{}^E\boldsymbol{\epsilon}_{dE})\Delta^E\boldsymbol{\omega}_{dE} \qquad (44)$$

$${}^c\dot{\boldsymbol{\epsilon}}_{dc} = \frac{1}{2}\boldsymbol{E}({}^c\eta_{dc},{}^c\boldsymbol{\epsilon}_{dc})\Delta^c\boldsymbol{\omega}_{dc} \qquad (45)$$

where $\boldsymbol{E}(\eta,\boldsymbol{\epsilon})=\eta\boldsymbol{I}-\boldsymbol{S}(\boldsymbol{\epsilon})$. Substitute (37), (38), (41), (42), (43), (44) and (45) into (40) we can get

$$\begin{aligned}\dot{\nu} &= -2K_{D_p}\Delta\dot{\boldsymbol{p}}_{dE}^T\Delta\dot{\boldsymbol{p}}_{dE}-K_{D_o}\Delta\boldsymbol{\omega}_{dE}^T\Delta\boldsymbol{\omega}_{dE}\\ &\quad -K_{D_o}\Delta^E\boldsymbol{\omega}_{dc}^T\Delta^E\boldsymbol{\omega}_{dc}\le 0 \qquad (46)\end{aligned}$$

For $\boldsymbol{K}_{P_p}$ and $\boldsymbol{K}_{P_o}$ are positive-definite, only if when $\Delta\dot{\boldsymbol{p}}_{dE}=\boldsymbol{0}$, $\Delta\boldsymbol{\omega}_{dE}=\boldsymbol{0}$ and $\Delta\boldsymbol{\omega}_{dc}=\boldsymbol{0}$, $\dot{\nu}=0$, For $\Delta\dot{\boldsymbol{p}}_{dE}=\boldsymbol{0}$ then $\Delta\ddot{\boldsymbol{p}}_{dE}=\boldsymbol{0}$, from (36), we can know that $\Delta\boldsymbol{p}_{dE}=\boldsymbol{0}$, When $\Delta\boldsymbol{\omega}_{dE}=\boldsymbol{0}$ and $\Delta\boldsymbol{\omega}_{dc}=\boldsymbol{0}$, $\Delta\dot{\boldsymbol{\omega}}_{dE}=\boldsymbol{0}$ and $\Delta\dot{\boldsymbol{\omega}}_{dc}=\boldsymbol{0}$, from (37) and (38) we can know ${}^E\boldsymbol{\epsilon}_{dE}=\boldsymbol{0}$ and ${}^c\boldsymbol{\epsilon}_{dc}=\boldsymbol{0}$. The definition domain of θ is $(-\pi,\pi)$, so the manipulator and the cameras asymptotically converge to the invariant sets s_p, s_o and s_c:

$$s_p = \{\Delta\boldsymbol{p}_{dE}=\boldsymbol{0},\ \Delta\dot{\boldsymbol{p}}_{dE}=\boldsymbol{0}\} \qquad (47)$$

$$s_o = \{\eta_{dE}=1,\ {}^E\boldsymbol{\epsilon}_{dE}=\boldsymbol{0},\ \Delta\boldsymbol{\omega}_{dE}=\boldsymbol{0}\} \qquad (48)$$

$$s_c = \{\eta_{dc}=1,\ {}^c\boldsymbol{\epsilon}_{dc}=\boldsymbol{0},\ \Delta^E\boldsymbol{\omega}_{dc}=\boldsymbol{0}\} \qquad (49)$$

Thus, the hand & Eye-vergence visual servoing system will be converged to the sets s_p, s_o, s_c, as shown in (47), (48), (49). (47) and (48) shows the hand is exponentially stable for any choice of positive definnitive $\boldsymbol{K}_{D_p}$, $\boldsymbol{K}_{P_p}$, $\boldsymbol{K}_{D_o}$, $\boldsymbol{K}_{P_o}$, thus.

$$\lim_{t\to\infty}{}^W\boldsymbol{r}_{E,Ed}=\boldsymbol{0} \qquad \lim_{t\to\infty}{}^W\dot{\boldsymbol{r}}_{E,Ed}=\boldsymbol{0} \qquad (50)$$

$$\lim_{t\to\infty}{}^E\Delta\boldsymbol{\epsilon}=\boldsymbol{0} \qquad \lim_{t\to\infty}{}^W\boldsymbol{\omega}_{c,cd}=\boldsymbol{0}. \qquad (51)$$

Then we have

$$\lim_{t\to\infty}{}^E\boldsymbol{T}_{Ed}=\boldsymbol{I} \qquad \lim_{t\to\infty}{}^E\dot{\boldsymbol{T}}_{Ed}=\boldsymbol{0} \qquad (52)$$

Substituting Eq. (52) to Eq. (56), we have

$$\lim_{t\to\infty}{}^E\boldsymbol{T}_{\hat{M}}=\lim_{t\to\infty}{}^{Ed}\boldsymbol{T}_{\hat{M}} \qquad (53)$$

Eq. (53) proves stable convergence of visual servoing.

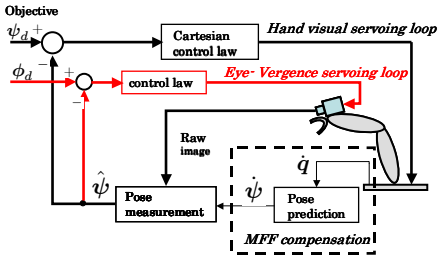

Fig. 4. Visual servoing system of PA-10

(49) shows

$$\lim_{t\to\infty}{}^c\boldsymbol{\epsilon}_{dc}=\boldsymbol{0}, \qquad \lim_{t\to\infty}\eta_{dc}=1 \qquad (54)$$

so the rotation matrix from the actual orientation to the desired orientation of the camera ${}^c\boldsymbol{R}_{dc}$ will [5] :

$$\begin{aligned}\lim_{t\to\infty}{}^c\boldsymbol{R}_{dc} &= \lim_{t\to\infty}({}^c\eta_{dc}^2-{}^c\boldsymbol{\epsilon}_{dc}^T{}^c\boldsymbol{\epsilon}_{dc})\boldsymbol{I}+2^c\boldsymbol{\epsilon}_{dc}{}^c\boldsymbol{\epsilon}_{dc}^T\\ &\quad +2^c\eta_{dc}\boldsymbol{S}({}^c\epsilon_{dc})\\ &= \boldsymbol{I} \qquad (55)\end{aligned}$$

The orientation error can exponentially converge to $\boldsymbol{0}$.

V. Hand & Eye Visual Servoing

A. Desired-trajectory generation

The desired relative relationship of Σ_M and Σ_E is given by Homogeneous Transformation as ${}^{Ed}\boldsymbol{T}_M(t)$, the difference of the desired camera pose Σ_{Ed} and the actual camera pose Σ_E is denoted as ${}^E\boldsymbol{T}_{Ed}$. ${}^E\boldsymbol{T}_{Ed}$ can be described by

$${}^E\boldsymbol{T}_{Ed} = {}^E\hat{\boldsymbol{T}}_M(\hat{\psi}(t))\ {}^M\boldsymbol{T}_{Ed}(t), \qquad (56)$$

Notice that Eq. (56) is a general deduction that satisfies arbitrary object motion ${}^W\boldsymbol{T}_M(t)$ and arbitrary objective of visual servoing ${}^{Ed}\boldsymbol{T}_M(t)$.

Differentiating Eq. (56) with respect to time yields

$${}^E\dot{\boldsymbol{T}}_{Ed}(t)={}^E\dot{\hat{T}}_M(t){}^M\boldsymbol{T}_{Ed}(t)+{}^E\hat{\boldsymbol{T}}_M(t){}^M\dot{\boldsymbol{T}}_{Ed}(t), \qquad (57)$$

Differentiating Eq. (57) with respect to time again

$$\begin{aligned}{}^E\ddot{\boldsymbol{T}}_{Ed}(t)={}^E\ddot{\hat{T}}_M(t){}^M\boldsymbol{T}_{Ed}(t)+2{}^E\dot{\hat{T}}_M(t){}^M\dot{\boldsymbol{T}}_{Ed}(t)+\\ {}^E\hat{\boldsymbol{T}}_M(t){}^M\ddot{\boldsymbol{T}}_{Ed}(t), \qquad (58)\end{aligned}$$

Where, ${}^M\boldsymbol{T}_{Ed}$, ${}^M\dot{\boldsymbol{T}}_{Ed}$, ${}^M\ddot{\boldsymbol{T}}_{Ed}$ are given as the desired visual servoing objective. ${}^E\hat{\boldsymbol{T}}_M$, ${}^E\dot{\hat{T}}_M$, ${}^E\ddot{\hat{T}}_M$ can be observed by cameras using the 1-step GA method.

B. 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.

Here, we just show main equations of the hand visual servoing controller that are used to calculate input torque $\boldsymbol{\tau}$ as:

$$\begin{aligned}\ddot{\boldsymbol{q}}_{Ed}=\boldsymbol{J}_E^+(\boldsymbol{q}_E)\Big(\begin{bmatrix}\boldsymbol{a}_p\\ \boldsymbol{a}_o\end{bmatrix}-\dot{\boldsymbol{J}}_E(\boldsymbol{q}_E,\dot{\boldsymbol{q}}_E)\dot{\boldsymbol{q}}_E\Big)+\big(\boldsymbol{I}-\\ \boldsymbol{J}_E^+(\boldsymbol{q}_E)\boldsymbol{J}_E(\boldsymbol{q}_E)\big)\big(\boldsymbol{E}_p(\boldsymbol{q}_{E_0}-\boldsymbol{q}_E)+\boldsymbol{E}_d(\boldsymbol{0}-\dot{\boldsymbol{q}}_E)\big), \qquad (59)\end{aligned}$$

Here, $\ddot{\boldsymbol{q}}_E$ is a 7×1 vector representing the angles of the first 7 links of the PA-10 manipulator. The quaternion error from the actual orientation to the desired orientation of the end effector ${}^E\Delta\boldsymbol{\epsilon}$ can be extracted from the transformation ${}^E\boldsymbol{T}_{Ed}$, and the other error variables in (22), (24) are described in Σ_W, which can be calculated by the transformation ${}^E\boldsymbol{T}_{Ed}$, ${}^E\dot{\boldsymbol{T}}_{Ed}$, ${}^E\ddot{\boldsymbol{T}}_{Ed}$ in (56), (57), (58), using the rotational matrix ${}^W\boldsymbol{R}_E(\boldsymbol{q})$ through coordinate transformation.

And $\boldsymbol{J}_E^+(\boldsymbol{q}_E)$ in (26) is the pseudo-inverse of $\boldsymbol{J}_E(\boldsymbol{q}_E)$ given by $\boldsymbol{J}_E^+(\boldsymbol{q}_E) = \boldsymbol{J}_E^T(\boldsymbol{J}_E\boldsymbol{J}_E^T)^{-1}$. $\boldsymbol{K}_{D_p}$, $\boldsymbol{K}_{P_p}$, $\boldsymbol{K}_{D_o}$, $\boldsymbol{K}_{P_o}$ are positive control gains.

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. 5, ${}^Ex_{\hat{M}}$, ${}^Ey_{\hat{M}}$, ${}^Ez_{\hat{M}}$ express position of the detected object in the end-effector coordinate, Σ_E. The desired angle of the camera joints are calculated by:

$$q_{8d} = atan2({}^Ez_{\hat{M}}, {}^Ey_{\hat{M}}) \tag{60}$$

$$q_{9d} = atan2({}^Ez_{\hat{M}}, -l_{8R} + {}^Ex_{\hat{M}}) \tag{61}$$

$$q_{10d} = atan2({}^Ez_{\hat{M}}, l_{8L} + {}^Ex_{\hat{M}}) \tag{62}$$

where $l_{8L} = l_{8R} = 150[mm]$ that is the camera location.

$${}^E\boldsymbol{\omega}_R = \begin{bmatrix} 1 & \cos q_8 \\ 0 & 0 \\ 0 & \sin q_8 \end{bmatrix} \begin{bmatrix} \dot{q}_8 \\ \dot{q}_9 \end{bmatrix} \tag{63}$$

define

$${}^E\boldsymbol{J}_R = \begin{bmatrix} 1 & \cos q_8 \\ 0 & 0 \\ 0 & \sin q_8 \end{bmatrix} \tag{64}$$

here ${}^E\boldsymbol{J}_R$ is the Jacobian matrix from the end-effector to the right camera, $\dot{\boldsymbol{q}}_R = [\dot{q}_8, \dot{q}_9]^T$, (63) also can be written as

$${}^E\boldsymbol{\omega}_R = {}^E\boldsymbol{J}_R\dot{\boldsymbol{q}}_R \tag{65}$$

and

$${}^E\boldsymbol{\omega}_{Rd} = {}^E\boldsymbol{J}_{Rd}\dot{\boldsymbol{q}}_{Rd} \tag{66}$$

and ${}^E\dot{\boldsymbol{\omega}}_{Rd}$ can be calculated by:

$${}^E\dot{\boldsymbol{\omega}}_{Rd} = {}^E\dot{\boldsymbol{J}}_{Rd}\dot{\boldsymbol{q}}_{Rd} + {}^E\boldsymbol{J}_{Rd}\ddot{\boldsymbol{q}}_{Rd}, \tag{67}$$

and the quarternion error from the actual orientation to the desired orientation of the right camera ${}^R\Delta\epsilon$ can be calculated by $[q_8, q_9]$ and $[q_{8d}, q_{9d}]$, so the compensation of the joint of the right camera can be calculated by:

$$\boldsymbol{\phi}_R = {}^E\boldsymbol{J}_R^+(\boldsymbol{q}_R)({}^E\boldsymbol{a}_{Ro} - {}^E\dot{\boldsymbol{J}}_R(\boldsymbol{q}_R)\dot{\boldsymbol{q}}_R) \tag{68}$$

In the similar way we can calculate the desired angular acceleration of the left camera. By controlling the cameras we can get better observation effect to decrease ${}^M\boldsymbol{T}_{\hat{M}}$ and to move the end-effector to the desired position and orientation

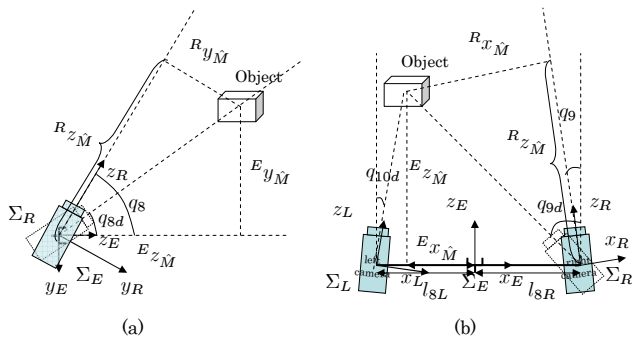

Fig. 5. Calculation of tilt and pan angles

By the controller of the whole hand & eye-vergence dual visual servoing system is:

$$\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_E \\ \boldsymbol{\phi}_R \\ \phi_{10} \end{bmatrix} \tag{69}$$

$$\boldsymbol{\tau} = \boldsymbol{M}(\boldsymbol{q})\boldsymbol{\phi} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}). \tag{70}$$

Here, $\boldsymbol{\tau}$ is a 10×1 vector, and $\boldsymbol{\tau}$ means the input torque of the 7-links manipulator and 3-links camera system.

VI. Conclusion

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, and apply it into simulation, and prove the convergence of motion of the system by Lyapunov theorem. In the future we will improve the stability and recognition speed of this method, and apply it into more fields.

References

[1] S.Hutchinson, G.Hager, and P.Corke, "A Tutorial on Visual Servo Control", IEEE Trans. on Robotics and Automation, vol. 12, no. 5, pp. 651-670, 1996.

[2] E.Malis, F.Chaumentte and S.Boudet, "2-1/2-D Visual Servoing", IEEE Trans. on Robotics and Automation, vol. 15, no. 2, pp. 238-250, 1999.

[3] Y.Maeda, G.Xu, "Smooth Matching of Feature and Recovery of Epipolar Equation by Tabu Search", IEICE, Vol.J83-D-2, No.3, pp.440-448, 1999.

[4] D. E.Goldberg, *Genetic algorithm in Search, Optimization and Machine Learning. Reading*, Addison-Wealey, 1989.

[5] B.Siciliano and L.Villani: *Robot Force Control*, ISBN 0-7923-7733-8.

[6] 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, Page 127-139 (2007).

[7] L. Reznik and V. Kreinovich Eds, *Soft Computing in Measurement and Information Acquisition*, ISBN 3-540-00246-4.

[8] 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 (2005.6).

[9] M. MINAMI, J. ZHU and M. MIURA, *Real-time Evolutionary Recognition of Human with Adaptation to Environmental Condition*, The Second Int. Conf. on Computational Intelligence, Robotics and Autonomous Systems (CIRAS 2003).