# Quick Eye-Vergence system to Enhance visual-servoing Trackability

Fujia Yu, Wei Song, Mamoru Minami and Yanou Akira

***Abstract*— Visual servoing methods for hand-eye configuration are vulnerable for hand's dynamical oscillation, since nonlinear dynamical effects of whole manipulator stand against the stable tracking ability (trackability). Our proposal to solve this problem is that the controller for visual servoing of the hand and the one for eye-vergence should be separated independently based on decoupling each other, where the trackablity is verified by Lyapunov analysis. Then the effectiveness of the decoupled hand & eye-vergence visual servoing method is evaluated through simulations incorporated with actual dynamics of 7-DoF robot with additional 3-DoF for eye-vergence mechanism by amplitude and phase frequency analysis.**

## I. INTRODUCTION

Visual servoing can be classified into three major groups by its methods: position-based, image-based and hybrid visual servoing [1]. In most image-based researches they concentrated on a planar object [2], [3], while in the hybrid controller the object is always static [4]. A hand-eye configuration has a merit of the ability choosing the viewpoint adaptively instead of the tendency being unstable during servoing motion due to hand's dynamical oscillation. Thus enhancing both the camera's tracking ability and robot's servoing stability is inherent hazard for hand-eye configuration, since they deteriorate each other in visual servoing motion. Also keeping suitable viewpoint is important for pose estimation to track the target precisely. If a pose measurement system can provide servoing controller with correct pose without time-delay, it can improve the stability of servoing dynamics since it is common sense that the time-delay existing in feedback mechanism may mess up closed-loop stability. Some methods are proposed to improve observation, like using stereo camera [5], multiple cameras [6], and two cameras: with one fixed on the end-effector, and the other done in the workspace [7]. However, these methods to give different views to observe the object by only increasing the number of cameras remain in less adaptive for changing environment.

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. 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 [5], [8], [9]. 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.

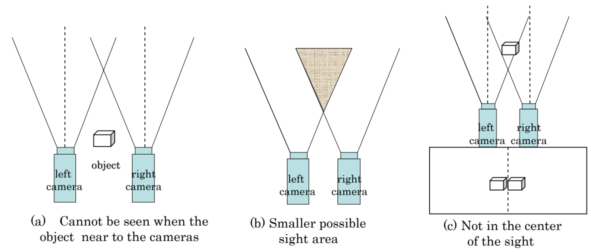


Fig. 1. Disadvantage of fix camera system

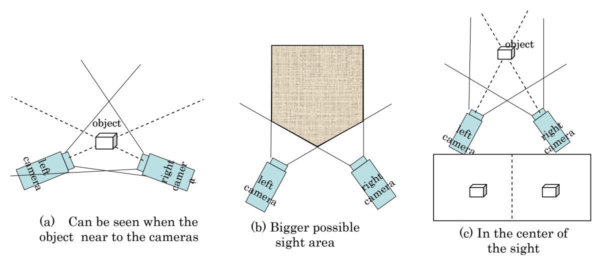


Fig. 2. Advantage of Eye-vergence system

In visual servoing application, it is important to keep the object in the visual eye sight to make the visual feed back not be severed 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

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

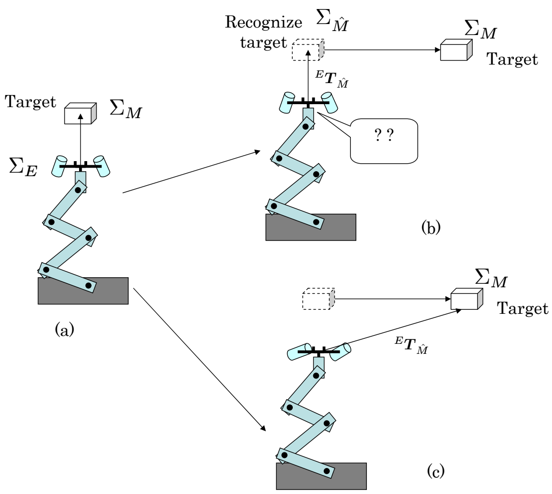


Fig. 3. Dynamics advantage of Eye-vergence system

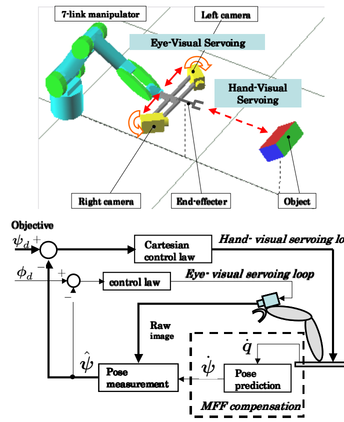


Fig. 4. Hand & Eye Visual servo system

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 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 [**?**], 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 appreciate our visual servoing proposal introduced in the next paragraph.

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. 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.

The effectiveness of the proposed hand & eye-vergence dual visual servoing will be evaluated through simulations incorporated with actual dynamics of 7-DoF robot with additional 3-DoF for eye-vergence mechanism of left and right camera's motion. We discuss the performance of the proposed system on the view points of how the new idea improve the stability and trackability against quick motion of the target by frequency analysis, and get the amplitude-frequency curve and phase-frequency curve between the motion of the target and the sight of the camera of fixed-camera system and eye-vergence system to confirm that the proposed system has better stability and trackability.

## II. Simulator and Robot Dynamics

The Mitsubishi PA-10 robot arm is a 7-DoF robot arm manufactured by Mitsubishi Heavy Industries. Our simulator is model PA-10 , by using the actual physical parameters of the PA-10.

The general equation of motion of manipulator is

$$\boldsymbol{M}(\boldsymbol{q}_E)\ddot{\boldsymbol{q}}_E + \boldsymbol{h}(\boldsymbol{q}_E, \dot{\boldsymbol{q}}_E) + \boldsymbol{g}(\boldsymbol{q}_E) + \boldsymbol{d}(\dot{\boldsymbol{q}}_E) = \boldsymbol{\tau}_E, \quad (1)$$

where, $\boldsymbol{q}_E$: the joint displacement and $\boldsymbol{q}_E = [q_1, q_2, \cdots, q_7]^T$, $\boldsymbol{\tau}_E$: the joint driving force and $\boldsymbol{\tau}_E = [\tau_1, \tau_2, \cdots, \tau_7]^T$, $\boldsymbol{M}(\boldsymbol{q}_E)$: the inertia matrix, $\boldsymbol{h}(\boldsymbol{q}_E, \dot{\boldsymbol{q}}_E)$: the vector representing the centrifugal and coriolis forces, $\boldsymbol{g}(\boldsymbol{q}_E)$: the vector representing the gravity load, $\boldsymbol{d}(\dot{\boldsymbol{q}}_E)$: the vector representing the frictional force. Here, we assume $\boldsymbol{d}(\dot{\boldsymbol{q}}_E) = \boldsymbol{0}$.

Two cameras are mounted on the end-effector, FCB-1X11A manufactured by Sony Industries. The frame frequency of stereo cameras is set as 33fps. The structure of the manipulator and the cameras are shown in the top of 4. To evaluate the trackability of a visual servoing system, we should have the whole model of the system first, in the research before we always ignore the mass and moment inertia of the camera system. In this paper we also consider the mass and moment inertia of the two cameras.

## III. Hand & Eye Visual Servoing

### A. Desired-trajectory generation

As shown in Fig. 5, the world coordinate frame is denoted by $\Sigma_W$, the target coordinate frame is denoted by $\Sigma_M$, and the desired and actual end-effector coordinate frame is denoted by $\Sigma_{Ed}$, $\Sigma_E$ separately. The desired relative relation between the target and the end-effector is given by Homogeneous Transformation as $^{Ed}\boldsymbol{T}_M$, the relation between the target and the actual end-effector is given by $^{E}\boldsymbol{T}_M$, then the difference between the desired end-effector

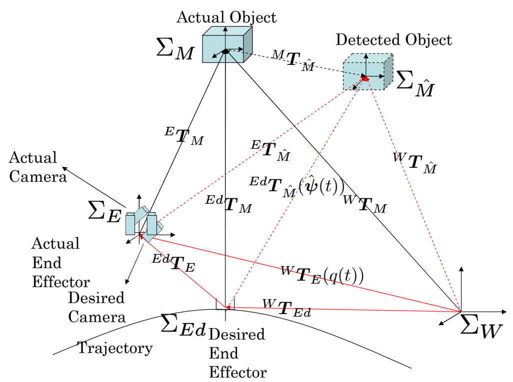


Fig. 5. Motion of the end-effector and object

pose $\Sigma_{Ed}$ and the actual end-effector pose $\Sigma_E$ is denoted as ${}^{E}\boldsymbol{T}_{Ed}$, ${}^{E}\boldsymbol{T}_{Ed}$ can be decribed by:

$$ {}^{E}\boldsymbol{T}_{Ed}(t) \;=\; {}^{E}\boldsymbol{T}_{M}(t)\,{}^{Ed}\boldsymbol{T}_{M}^{-1}(t) \tag{2} $$

(2) is a general deduction that satisfies arbitrary object motion ${}^{W}\boldsymbol{T}_{M}(t)$ and arbitrary visual servoing objective ${}^{Ed}\boldsymbol{T}_{M}(t)$. However, the relation ${}^{E}\boldsymbol{T}_{M}(t)$ is only observed by cameras using the on-line model-based recognition method and 1-step GA [5]. Let $\Sigma_{\hat{M}}$ denote the detected object, there always exist an error between the actual object $\Sigma_M$ and the detected one $\Sigma_{\hat{M}}$. So in visual servoing, (2) will be rewritten based on $\Sigma_{\hat{M}}$ that includes the error ${}^{M}\boldsymbol{T}_{\hat{M}}$, as

$$ {}^{E}\boldsymbol{T}_{Ed}(t) \;=\; {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{Ed}\boldsymbol{T}_{\hat{M}}^{-1}(t), \tag{3} $$

where ${}^{E}\boldsymbol{T}_{\hat{M}} = {}^{E}\boldsymbol{T}_{M}$ determined by the given visual servoing objective. Differentiating (3) with respect to time yields

$$ {}^{E}\dot{\boldsymbol{T}}_{Ed}(t) = {}^{E}\dot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\boldsymbol{T}_{Ed}(t) + {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{\hat{M}}\dot{\boldsymbol{T}}_{Ed}(t), \tag{4} $$

Differentiating Eq. (4) with respect to time again

$$ \begin{aligned} {}^{E}\ddot{\boldsymbol{T}}_{Ed}(t) = {}^{E}\ddot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\boldsymbol{T}_{Ed}(t) + 2\,{}^{E}\dot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\dot{\boldsymbol{T}}_{Ed}(t) + \\ {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{\hat{M}}\ddot{\boldsymbol{T}}_{Ed}(t), \end{aligned} \tag{5} $$

Where ${}^{\hat{M}}\boldsymbol{T}_{Ed}$, ${}^{\hat{M}}\dot{\boldsymbol{T}}_{Ed}$, ${}^{\hat{M}}\ddot{\boldsymbol{T}}_{Ed}$ are given as the desired visual servoing objective. ${}^{E}\boldsymbol{T}_{\hat{M}}$, ${}^{E}\dot{\boldsymbol{T}}_{\hat{M}}$, ${}^{E}\ddot{\boldsymbol{T}}_{\hat{M}}$ can be observed by cameras. As shown in Fig. 5, there are two errors that we have to decrease to $\boldsymbol{0}$ in the visual servoing process. First one is the error between the actual object and the detected one, ${}^{M}\boldsymbol{T}_{\hat{M}}$, and the other is the error between the desired end-effector and the actual one, ${}^{E}\boldsymbol{T}_{Ed}$. In our research, the error of ${}^{M}\boldsymbol{T}_{\hat{M}}$ is decreased by on-line recognition method of 1-step GA, MFF compensation method and the eye-vergence camera system, and the error of ${}^{E}\boldsymbol{T}_{Ed}$ can be decreased by the hand visual servoing controller.

### 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. The controller used for hand-visual servoing is proposed by B.Siciliano [10]. Here, we just show main equations of the hand visual servoing controller that are used to calculate input torque $\boldsymbol{\tau}$ as:

$$ \boldsymbol{a}_p = {}^{W}\ddot{\boldsymbol{r}}_{Ed} + \boldsymbol{K}_{D_p}\,{}^{W}\dot{\boldsymbol{r}}_{E,Ed} + \boldsymbol{K}_{P_p}\,{}^{W}\boldsymbol{r}_{E,Ed}, \tag{6} $$

$$ \boldsymbol{a}_o = {}^{W}\dot{\boldsymbol{\omega}}_{Ed} + \boldsymbol{K}_{D_o}\,{}^{W}\boldsymbol{\omega}_{E,Ed} + \boldsymbol{K}_{P_o}\,{}^{W}\boldsymbol{R}_{E}\,{}^{E}\Delta\boldsymbol{\epsilon}, \tag{7} $$

$$ \begin{aligned} \ddot{\boldsymbol{q}}_{Ed} = \boldsymbol{J}_E^{+}(\boldsymbol{q}_E)\Big(\begin{bmatrix}\boldsymbol{a}_p\\ \boldsymbol{a}_o\end{bmatrix} - \dot{\boldsymbol{J}}_E(\boldsymbol{q}_E,\dot{\boldsymbol{q}}_E)\dot{\boldsymbol{q}}_E\Big) + \big(\boldsymbol{I}- \\ \boldsymbol{J}_E^{+}(\boldsymbol{q}_E)\boldsymbol{J}_E(\boldsymbol{q}_E)\big)\big(\boldsymbol{E}_p(\boldsymbol{q}_{E_0}-\boldsymbol{q}_E)+\boldsymbol{E}_d(\boldsymbol{0}-\dot{\boldsymbol{q}}_E)\big), \end{aligned} \tag{8} $$

Here, $\ddot{\boldsymbol{q}}_E$ is a $7\times 1$ 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 (6), (7) are described in $\Sigma_W$, which can be calculated by the transformation ${}^{E}\boldsymbol{T}_{Ed}$, ${}^{E}\dot{\boldsymbol{T}}_{Ed}$, ${}^{E}\ddot{\boldsymbol{T}}_{Ed}$ in (3), (4), (5), using the rotational matrix ${}^{W}\boldsymbol{R}_E(\boldsymbol{q})$ through coordinate transformation.

And $\boldsymbol{J}_E^{+}(\boldsymbol{q}_E)$ in (8) 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. 6, ${}^{E}x_{\hat{M}}$, ${}^{E}y_{\hat{M}}$, ${}^{E}z_{\hat{M}}$ express position of the detected object in the end-effector coordinate. The desired angle of the camera joints are calculated by:

$$ \begin{aligned} q_{8d} &= atan2({}^{E}z_{\hat{M}}, {}^{E}y_{\hat{M}}) && (9)\\ q_{9d} &= atan2({}^{E}z_{\hat{M}}, -l_{8R} + {}^{E}x_{\hat{M}}) && (10)\\ q_{10d} &= atan2({}^{E}z_{\hat{M}}, l_{8L} + {}^{E}x_{\hat{M}}) && (11) \end{aligned} $$

where $l_{8L} = l_{8R} = 150[mm]$ that is the camera location. We set the center line of the camera as the $z$ axis of each camera coordinate, so the object will be in the center of the sight of the right camera when ${}^{R}x_{\hat{M}} = 0$ and ${}^{R}y_{\hat{M}} = 0$, ${}^{R}x_{\hat{M}}$, ${}^{R}y_{\hat{M}}$, ${}^{R}z_{\hat{M}}$ express the position of the detcted object in the right camera coordinate. While the object position relative to the cameras are:

$$ \begin{aligned} \frac{{}^{R}y_{\hat{M}}}{{}^{R}z_{\hat{M}}} &= tan(q_{8d} - q_8) && (12)\\ \frac{{}^{R}x_{\hat{M}}}{{}^{R}z_{\hat{M}}} &= tan(q_{9d} - q_9) && (13)\\ \frac{{}^{L}x_{\hat{M}}}{{}^{L}z_{\hat{M}}} &= tan(q_{10d} - q_{10}) && (14) \end{aligned} $$

Here we can use the relationship between the object and the right camera in Fig. 6 to define the trackability $c_R$ of the right camera on a object:

$$ c_R = \frac{1}{T}\int_0^T \frac{\sqrt{{}^{R}x_{\hat{M}}^2 + {}^{R}y_{\hat{M}}^2}}{{}^{R}z_{\hat{M}}}dt \tag{15} $$

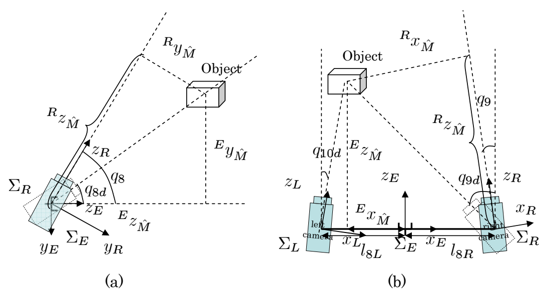


Fig. 6. Calculation of tilt and pan angles

here $T$ is the time used for tracking visual servoing experiment, and the trackability of the left camera and the end-effector can be calculated in the similar way, it is easy to see that when the object is always keeping in the center of the sight of the right camera, $c_R = 0$, in this case, we have the best trackability of the right camera.

In the same way as the hand visual servoing controller, the controller of the right camera is given by:

$$ {}^{E}\boldsymbol{\omega}_R = {}^{E}\boldsymbol{J}_R \dot{\boldsymbol{q}}_R \tag{16} $$

here ${}^{E}\boldsymbol{J}_R$ is the Jacobian matrix from the end-effector to the right camera, $\boldsymbol{q}_R = [q_8, q_9]^T$ and $\dot{\boldsymbol{q}}_R = [\dot{q}_8, \dot{q}_9]^T$, and

$$ {}^{E}\boldsymbol{\omega}_{Rd} = {}^{E}\boldsymbol{J}_{Rd} \dot{\boldsymbol{q}}_{Rd} \tag{17} $$

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

$$ {}^{E}\dot{\boldsymbol{\omega}}_{Rd} = {}^{E}\dot{\boldsymbol{J}}_{Rd} \dot{\boldsymbol{q}}_{Rd} + {}^{E}\boldsymbol{J}_{Rd} \ddot{\boldsymbol{q}}_{Rd}, \tag{18} $$

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:

$$ \ddot{\boldsymbol{q}}_{Rd} = {}^{E}\boldsymbol{J}_R^{+}(\boldsymbol{q}_R)(\dot{\boldsymbol{a}}_{oR} - {}^{E}\dot{\boldsymbol{J}}_R(\boldsymbol{q}_R)\dot{\boldsymbol{q}}_R) \tag{19} $$

$\ddot{\boldsymbol{q}}_{Rd}$ is a $2 \times 1$ vector that $\ddot{\boldsymbol{q}}_{Rd} = [\ddot{q}_{8d}, \ddot{q}_{9d}]^T$.

In the similar way we can calculate the desired angular acceleration of the left camera, for the two cameras use the same tilt link, $\ddot{q}_{Ld}$ is a scalar, where $\ddot{q}_{Ld} = \ddot{q}_{10d}$. 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 by the controller of the whole hand & eye-vergence dual visual servoing system is:

$$ \ddot{\boldsymbol{q}}_d = \begin{bmatrix} \ddot{\boldsymbol{q}}_{Ed} \\ \ddot{\boldsymbol{q}}_{Rd} \\ \ddot{q}_{Ld} \end{bmatrix} \tag{20} $$

$$ \boldsymbol{\tau} = \boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}}_d + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}). \tag{21} $$

Here, $\ddot{\boldsymbol{q}}_d$ and $\boldsymbol{\tau}$ are both $10 \times 1$ vectors, and $\boldsymbol{\tau}$ means the input torque of the 7-links manipulator and 3-links camera system.

## IV. Stability Analysis

The equation of motion of the whole system is

$$ \boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{\tau} \tag{22} $$

here we define

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

the compensation of robot's dynamics the outputs:

$$ \boldsymbol{\tau} = \boldsymbol{M}(\boldsymbol{q})\boldsymbol{\phi} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) \tag{23} $$

Taking the compensation (23) into (22), closed loop dynamics is:

$$ \ddot{\boldsymbol{q}} = \boldsymbol{\phi} \tag{24} $$

From the controllers of the manipulator and the cameras, we get closed loop hand camera motions are,

$$ \begin{aligned} \Delta\ddot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE} &= \boldsymbol{0} \quad (25) \\ \Delta\dot{\boldsymbol{\omega}}_{dE} + \boldsymbol{K}_{D_o}\Delta\boldsymbol{\omega}_{dE} + \boldsymbol{K}_{P_o}\boldsymbol{R}_E{}^{E}\boldsymbol{\epsilon}_{dE} &= \boldsymbol{0} \quad (26) \\ \Delta{}^{E}\dot{\boldsymbol{\omega}}_{dc} + \boldsymbol{K}_{D_o}\Delta{}^{E}\boldsymbol{\omega}_{dc} + \boldsymbol{K}_{P_o}{}^{E}\boldsymbol{R}_c{}^{c}\boldsymbol{\epsilon}_{dc} &= \boldsymbol{0} \quad (27) \end{aligned} $$

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.

$$ \begin{aligned} \nu = {} & \Delta\boldsymbol{p}_{dE}^T K_{P_p}\Delta\boldsymbol{p}_{dE} + \Delta\dot{\boldsymbol{p}}_{dE}^T\Delta\dot{\boldsymbol{p}}_{dE} \\ & + K_{P_o}((\eta_{dE} - 1)^2 + {}^{E}\boldsymbol{\epsilon}_{dE}^T{}^{E}\boldsymbol{\epsilon}_{dE}) + \frac{1}{2}\Delta\boldsymbol{\omega}_{dE}^T\Delta\boldsymbol{\omega}_{dE} \\ & + K_{P_o}(({}^{E}\eta_{dc} - 1)^2 + {}^{c}\boldsymbol{\epsilon}_{dc}^T{}^{c}\boldsymbol{\epsilon}_{dc}) + \frac{1}{2}\Delta{}^{E}\boldsymbol{\omega}_{dc}^T\Delta{}^{E}\boldsymbol{\omega}_{dc} \\ & \geq 0 \end{aligned} \tag{28} $$

so

$$ \begin{aligned} \dot{\nu} = {} & -2\Delta\dot{\boldsymbol{p}}_{dE}^T(\Delta\ddot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE}) \\ & + 2K_{P_o}((\eta_{dE} - 1)\dot{\eta}_{dE} + {}^{E}\boldsymbol{\epsilon}_{dE}^T{}^{E}\dot{\boldsymbol{\epsilon}}_{dE}) + \Delta\boldsymbol{\omega}_{dE}^T\Delta\dot{\boldsymbol{\omega}}_{dE} \\ & + 2K_{P_o}(({}^{E}\eta_{dc} - 1){}^{E}\dot{\eta}_{dc} + {}^{c}\boldsymbol{\epsilon}_{dc}^T{}^{c}\dot{\boldsymbol{\epsilon}}_{dc}) \\ & + \Delta{}^{E}\boldsymbol{\omega}_{dc}^T\Delta{}^{E}\dot{\boldsymbol{\omega}}_{dc} \end{aligned} \tag{29} $$

from (25) we can know that

$$ \Delta\ddot{\boldsymbol{p}}_{dE} + \boldsymbol{K}_{P_p}\Delta\boldsymbol{p}_{dE} = -\boldsymbol{K}_{D_p}\Delta\dot{\boldsymbol{p}}_{dE} \tag{30} $$

from the quaternion definition we can know that [10]

$$ \dot{\eta}_{dE} = -\frac{1}{2}{}^{E}\boldsymbol{\epsilon}_{dE}^T\Delta{}^{E}\boldsymbol{\omega}_{dE} \tag{31} $$

and

$$ {}^{E}\dot{\eta}_{dc} = -\frac{1}{2}{}^{c}\boldsymbol{\epsilon}_{dc}^T\Delta{}^{c}\boldsymbol{\omega}_{dc} \tag{32} $$

and

$$ {}^{E}\dot{\boldsymbol{\epsilon}}_{dE} = \frac{1}{2}\boldsymbol{E}(\eta_{dE}, {}^{E}\boldsymbol{\epsilon}_{dE})\Delta{}^{E}\boldsymbol{\omega}_{dE} \tag{33} $$

$$ {}^{c}\dot{\boldsymbol{\epsilon}}_{dc} = \frac{1}{2}\boldsymbol{E}({}^{c}\eta_{dc}, {}^{c}\boldsymbol{\epsilon}_{dc})\Delta{}^{c}\boldsymbol{\omega}_{dc} \tag{34} $$

where $\boldsymbol{E}(\eta, \boldsymbol{\epsilon}) = \eta \boldsymbol{I} - \boldsymbol{S}(\boldsymbol{\epsilon})$. Substitute (26), (27), (30), (31), (32), (33) and (34) into (29) we can get

$$\begin{aligned} \dot{\nu} &= -2K_{D_p}\Delta\dot{\boldsymbol{p}}_{dE}^T\Delta\dot{\boldsymbol{p}}_{dE} - K_{D_o}\Delta\boldsymbol{\omega}_{dE}^T\Delta\boldsymbol{\omega}_{dE} \\ &\quad -K_{D_o}\Delta^E\boldsymbol{\omega}_{dc}^T\Delta^E\boldsymbol{\omega}_{dc} \le 0 \end{aligned} \quad (35)$$

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 (25), 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 (26) and (27) we can know ${}^E\boldsymbol{\epsilon}_{dE} = \boldsymbol{0}$ and ${}^c\boldsymbol{\epsilon}_{dc} = \boldsymbol{0}$. The definition domain of $\theta$ is $(-\pi, \pi)$, so the manipulator and the cameras asymptotically converge to the invariant sets $s_p$, $s_o$ and $s_c$:

$$s_p = \{\Delta\boldsymbol{p}_{dE} = \boldsymbol{0},\ \Delta\dot{\boldsymbol{p}}_{dE} = \boldsymbol{0}\} \quad (36)$$

$$s_o = \{\eta_{dE} = 1,\ {}^E\boldsymbol{\epsilon}_{dE} = \boldsymbol{0},\ \Delta\boldsymbol{\omega}_{dE} = \boldsymbol{0}\} \quad (37)$$

$$s_c = \{\eta_{dc} = 1,\ {}^c\boldsymbol{\epsilon}_{dc} = \boldsymbol{0},\ \Delta^E\boldsymbol{\omega}_{dc} = \boldsymbol{0}\} \quad (38)$$

Thus, the hand & Eye-vergence visual servoing system will be converged to the sets $s_p$, $s_o$, $s_c$, as shown in (36), (37), (38). (36) and (37) 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.

$$\lim_{t\to\infty} {}^W\boldsymbol{r}_{E,Ed} = \boldsymbol{0} \qquad \lim_{t\to\infty} {}^W\dot{\boldsymbol{r}}_{E,Ed} = \boldsymbol{0} \quad (39)$$

$$\lim_{t\to\infty} {}^E\Delta\boldsymbol{\epsilon} = \boldsymbol{0} \qquad \lim_{t\to\infty} {}^W\boldsymbol{\omega}_{c,cd} = \boldsymbol{0}. \quad (40)$$

Then we have

$$\lim_{t\to\infty} {}^E\boldsymbol{T}_{Ed} = \boldsymbol{I} \qquad \lim_{t\to\infty} {}^E\dot{\boldsymbol{T}}_{Ed} = \boldsymbol{0} \quad (41)$$

Substituting Eq. (41) to Eq. (3), we have

$$\lim_{t\to\infty} {}^E\boldsymbol{T}_{\hat{M}} = \lim_{t\to\infty} {}^{Ed}\boldsymbol{T}_{\hat{M}} \quad (42)$$

Eq. (42) proves stable convergence of visual servoing.

(38) shows

$$\lim_{t\to\infty} {}^c\boldsymbol{\epsilon}_{dc} = \boldsymbol{0}, \qquad \lim_{t\to\infty} \eta_{dc} = 1 \quad (43)$$

so the rotation matrix from the actual orientation to the desired orientation of the camera ${}^c\boldsymbol{R}_{dc}$ will [10] :

$$\begin{aligned} \lim_{t\to\infty} {}^c\boldsymbol{R}_{dc} &= \lim_{t\to\infty} ({}^c\eta_{dc}^2 - {}^c\boldsymbol{\epsilon}_{dc}^T\,{}^c\boldsymbol{\epsilon}_{dc})\boldsymbol{I} + 2{}^c\boldsymbol{\epsilon}_{dc}{}^c\boldsymbol{\epsilon}_{dc}^T \\ &\quad +2{}^c\eta_{dc}\boldsymbol{S}({}^c\epsilon_{dc}) \\ &= \boldsymbol{I} \end{aligned} \quad (44)$$

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

$$\lim_{t\to\infty} \boldsymbol{q}_C = \boldsymbol{q}_{Cd} \quad (45)$$

here $\boldsymbol{q}_C = [q_8, q_9, q_{10}]$. From (12) to (14) we can get:

$$\lim_{t\to\infty} {}^Ry_M = 0, \quad \lim_{t\to\infty} {}^Rx_M = 0, \quad \lim_{t\to\infty} {}^Lx_M = 0 \quad (46)$$

so the object will become on the center line of the cameras, which means that the object will always keep in the center of the sight of the cameras.

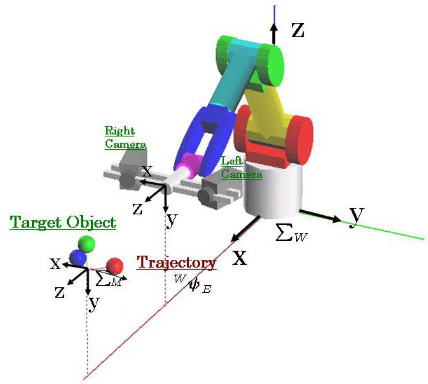


Fig. 7. Object and the visual-servoing system

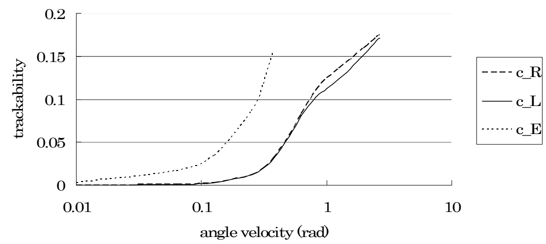


Fig. 8. trackability of the end effector and the cameras

## V. SIMULATION OF HAND & EYE-VERGENCE VISUAL SERVOING

To verify the effectiveness of the proposed hand & eye visual servoing system, we conduct the simulation of visual servoing to a 3D marker that is composed of a red ball, a green ball and a blue ball as Fig. 7.

### A. simulation condition

To cancel the error caused by the recognition of the object, we will give the position and orientation of the object to the robot directly in the simulation. The initial hand pose is defined as $\Sigma_{E_0}$, while the initial object pose is defined as $\Sigma_{M_0}$, and the homogeneous transformation matrix from $\Sigma_W$ to $\Sigma_{M_0}$ is:

$${}^W\boldsymbol{T}_{M_0} = \begin{bmatrix} 0 & 0 & 1 & 1388[mm] \\ -1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 455[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (47)$$

while the object will move according to the following time function,

$${}^{M_0}\boldsymbol{\psi}_M = [200\sin(\omega t)[mm], 0[mm], 0[mm], 0, 0, 0]^T \quad (48)$$

here, $\omega$ is the angular velocity of the motion of the object.

The relation between the object and the desired end-effector is.

$$\begin{cases} {}^{Ed}x_M(t) = 0 \\ {}^{Ed}y_M(t) = 0 \\ {}^{Ed}z_M(t) = 500[mm] \\ {}^{Ed}\epsilon_{1M}(t) = 0 \\ {}^{Ed}\epsilon_{2M}(t) = 0 \\ {}^{Ed}\epsilon_{3M}(t) = 0 \end{cases} \quad (49)$$

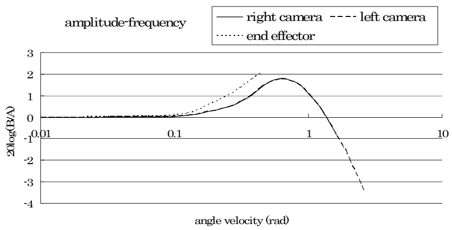


Fig. 9. amplitude-frequency curve of the end-effector and the cameras

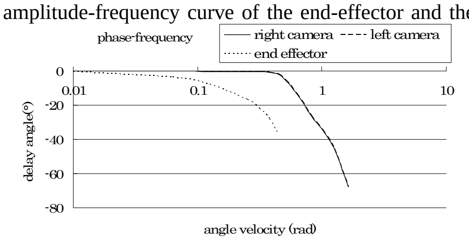


Fig. 10. phase-frequency curve of the end-effector and the cameras

### B. *simulation Results*

We can see the trackability under different angular velocitys in Fig. 8, because the position and orientation of the object are given directly to the robot, and we use the same manipulator controller in eye-vergence system and fixed camera system, the trackability of the end-effector of the eye-vergence system is same to the trackability of the fixed camera system. In Fig. 8 we can see that as the object angular velocity increase the trackability also increase, and when the $\omega$ in (48) get to 0.4396 the fixed camera system will lose the object in its sight while in eye-vergence system the higest angular velocity it can get and keeping the object always in the sight of the cameras is 2.6376, we can know that the eye-vergence system has the better stability than the fixed camera system. We can also see that the trackability of the hand eyes ($c_R$ and $c_L$) are smaller than the trackability of the end-effector ($c_E$) which is also the trackability of the fixed camera system under same $\omega$. From the definition we can know the eye-vergence system has better trackability.

Fig. 9 and Fig. 10 are the amplitude-frequency curve and phase-frequency curve of the fixed camera system and eye-vergence system. In both figures we use the logrithmic scalar of the $\omega$ in (48) as the x-axis. To check the observation ability of the camera we calculate out the point where the camera is gazing at. The gazing point of the right camera expressed in the world frame ${}^W\boldsymbol{p}_{GR}$ as it is shown in Fig. 6 can be calculated as follow:

$$ {}^W x_{GR} = {}^W x_E + {}^E z_{\hat{M}} \quad (50)$$

$$ {}^W y_{GR} = {}^W y_E - l_{8R} - {}^E z_{\hat{M}} \tan q_9 \quad (51)$$

$$ {}^W z_{GR} = {}^W z_E + {}^E z_{\hat{M}} \tan q_8 \quad (52)$$

here, ${}^W x_{GR}$, ${}^W y_{GR}$, ${}^W z_{GR}$ are the three elements of ${}^W\boldsymbol{p}_{GR}$. The gazing point of the left camera can be calculated in the similar way. The position of the end-effector and the gazing points of the cameras on y-axis of the world frame can be approximately expressed as a function $Bsin(\omega t+\phi)$, so the swing of the end-effector of the fixed camera system and the gazing point of the cameras is $B$, and the swing of motion of the object is $A$. We set $20\log\frac{B}{A}$ as the y-axis of the amplitude-frequency curve. In Fig. 9, for the smaller mass of the cameras the swing of the gazing point of the cameras are bigger than the end-effector, for the resonance reason the $20\log\frac{B}{A}$ may increase at first and decrease as the angular velocity of the object increase. In Fig. 10 for the motion function of the object in y-axis of the world frame is $0.2sin(\omega t)$, $\phi$ in the motion function of the end-effector and hand-eye cameras can be considered as the delay phase, we use the $\phi$ of the end-effector and the gazing points of the cameras as the y-axis of the phase-frequency figure, from this figure we can see that the eye-vergence system has smaller delay phase which means it will observe the object better.

## VI. 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. To check the trackability, we found a whole model of PA-10 including the cameras' mass and moment inertia. The controller of the system includes two loops: an outer 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. In the simulation we compare the trackability, amplitud-frequency and phase-frequency curves of the cameras of the eye-vergence system and the fixed camera system under moving object with different angular velocity, and get the conclusion that the trackability and stability of the eye-vergence system is better than that of the fixed-camera system.

## REFERENCES

[1] Amei massoud Farahmand, Azad Shademan, Martin Jägersand, Csaba szepesvári "Model-based and Model-free Reinforcement Learning for Visual Servoing" IEEE International Conference on Robotics and Automation Kobe, Japan, May 12-17, 2009

[2] Toshifumi Hiramatsu, Takanori Fukao, Keita Kurashiki, Koichi Osuka "Image-based Path Following Control of Mobile Robots with Central Catadioptric Cameras" IEEE International Conference on Robotics and Automation Kobe, Japan, May 12-17, 2009

[3] Omar Tahri, Youcef Mezouar "Generic Decoupled Image-Based Visual Servoing for Cameras Obeying the Unified Projection Model" IEEE International Conference on Robotics and Automation Kobe, Japan, May 12-17, 2009

[4] Dae-Jin Kim, Ryan Lovelett, and Aman Behal "Eye-in-Hand Stereo Visual Servoing of an Assistive Robot Arm in Unstructured Environments" IEEE International Conference on Robotics and Automation Kobe, Japan, May 12-17, 2009

[5] W. Song, M. Minami, Y. Mae and S. Aoyagi, "On-line Evolutionary Head Pose Measurement by Feedforward Stereo Model Matching", IEEE Int. Conf. on Robotics and Automation (ICRA), pp.4394-4400, 2007.

[6] J. Stavnitzky, D. Capson, "Mutiple Camera Model-Based 3-D Visual Servoing", IEEE Trans. on Robotics and Automation, vol. 16, no. 6, December 2000.

[7] C. Dune, E. Marchand, C. leroux, "One Click Focus with Eye-in-hand/Eye-to hand Cooperation", IEEE Int. Conf. on Robotics and Automation (ICRA), pp.2471-2476, 2007.

[8] 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.

[9] 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.

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