Proceedings of the 2016 IEEE/SICE International Symposium on
System Integration, Sapporo Convention Center, Sapporo, Japan,
December 13-15, 2016

Tracking Performances of Eye-Vergence Visual Servoing System under Different Light Condition with Evolutionary Pose Tracking

Hongzhi Tian[1], Yejun Kou, Ryuki Funakubo and Mamoru Minami

***Abstract*— People have a strong ability to adapt to the changes of light. And light changes in the natural environment is inevitable. In the past we have proposed eye-vergence visual servoing system. However we did not know how it performs in dark or bright environment. If a visual servoing system has a higher light environmental robustness it will have a wider range of application. Therefore we design tracking experiments under different illumination and partial occlusion condition. Though the experiment it is verified that the system is not easily affected by ambient brightness and can continually track the object. Compared with the partial occlusion condition, the influence of changing of light condition is smaller.**

I. INTRODUCTION

Nowadays, in a field of robot vision, a control method called a visual servoing attracts attentions [1], which can be classified into three major groups: position-based [2], image-based [3] and hybrid visual servoing [4].

The visual servoing, a method for controlling a robot using visual information in the feedback loop, is expected to be able to allow the robot to adapt to changing or unknown environments. Some methods have already been proposed to improve observation abilities, by using stereo cameras [5], multiple cameras [6], and two cameras; with one fixed on the end-effector, and the other one fixed in the workspace [7]. These methods obtain different views to observe an object by increasing the number of cameras.

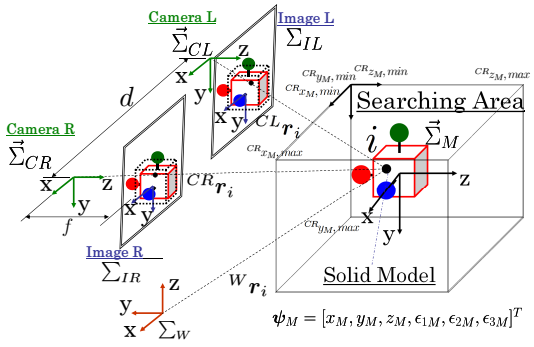


Fig. 1. Coordinate systems of dual eyes

When the end-effector is close to a target object, eye-vergence camera system can look at it in the center of camera images all the time by utilizing the changeable cameras' eye directions [9] so as to look the target the center of camera images. This advantage of eye-vergence can be called "kinematical merit."

[1]Hongzhi Tian is with Division of Mechanical and Systems Engineering, Graduate school of Natural Science and Technology, Okayama University, 1-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan psnc8ytd@s.okayama-u.ac.jp

It is easy to catch up to the object in case of the object moving slowly, but when the object become moving faster and faster, human's face be hardly kept positioned squarely to the object, while human's eye can still keep staring at the object because of its small mass and inertial moment. This another advantage of eye vergence can be called "dynamical merit."

Even though the eye-vergence servoing system has merits of kinematical and dynamical performances, it is not useful if it cannot work at normal conditions that could be affected by lighting condition varieties caused by climate changes. These merits of eye-vergence concerning kinematical and dynamical effects to keep tracking a moving target in the camera's view is deemed to be important and useful. Needless to say in visual servoing application, keeping closed loop of visual feedback is vital from a view point of closed loop control stability.

For practical application the light condition is an important effect element for visual servoing system to directly recognize the target object. In this report, some visual servoing experiments are conducted about object's lateral direction movement with light changing and even with partial occlusion, and confirmed that the visual servoing can be achieved in different light condition.

II. 3D POSE TRACKING METHOD

As shown in Fig.4 (d), in this paper, a 3D-ball-object whose size and color are known is taken as an example of the target object. Of couse, other objects having different shapes can also be measured by model-based matching strategy if their character is given, for example, in [8] a model of fish is used to track fish in real time, and in [5] a model of human face is used for face detection.

A. Kinematics of Stereo-Vision

Perspective projection is utilized as projection transformation. Fig.1 shows the coordinate system of the dual-eyes vision system. The target object's coordinate system is represented by $\vec{\Sigma}_M$ and image coordinate systems of the left and right cameras are represented by Σ_{IL} and Σ_{IR}. The difference between $\vec{\Sigma}$ and Σ is explained in section IV C. A point i on the target can be described using these coordinates and homogeneous transformation matrices. At first, a homogeneous transformation matrix from right camera coordinates, $\vec{\Sigma}_{CR}$ to $\vec{\Sigma}_M$ is defined as ${}^{CR}\boldsymbol{T}_M$. And an arbitrary point i on the target object in $\vec{\Sigma}_{CR}$ and $\vec{\Sigma}_M$ is defined as ${}^{CR}\boldsymbol{r}_i$ and ${}^{M}\boldsymbol{r}_i$. Then ${}^{CR}\boldsymbol{r}_i$ is,

$$ {}^{CR}\boldsymbol{r}_i = {}^{CR}\boldsymbol{T}_M \, {}^{M}\boldsymbol{r}_i. \tag{1}$$

978-1-5090-3328-7/16/$31.00 ©2016 IEEE

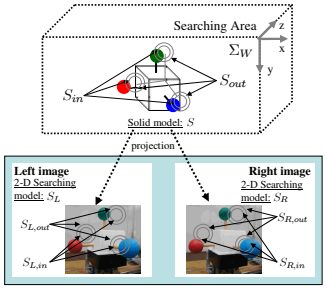


Fig. 2. Definition of a solid model and left/right searching models

Where ${}^{M}\boldsymbol{r}_i$ is predetermined fixed vectors. Using a homogeneous transformation matrix from Σ_W to the right camera coordinate system $\vec{\Sigma}_{CR}$, i.e., ${}^{W}\boldsymbol{T}_{CR}$, then ${}^{W}\boldsymbol{r}_i$ is got as,

$$ {}^{W}\boldsymbol{r}_i = {}^{W}\boldsymbol{T}_{CR}\ {}^{CR}\boldsymbol{r}_i. \tag{2}$$

The position vector of i point in right image coordinates, ${}^{IR}\boldsymbol{r}_i$ is described by using projection matrix $\boldsymbol{P}$ of camera as,

$$ {}^{IR}\boldsymbol{r}_i = \boldsymbol{P}\ {}^{CR}\boldsymbol{r}_i. \tag{3}$$

By the same way as above, using a homogeneous transformation matrix of fixed values defining the kinematical relation from the left camera coordinate system $\vec{\Sigma}_{CL}$ to the right camera coordinate system $\vec{\Sigma}_{CR}$, ${}^{CL}\boldsymbol{T}_{CR}$, ${}^{CL}\boldsymbol{r}_i$ is,

$$ {}^{CL}\boldsymbol{r}_i = {}^{CL}\boldsymbol{T}_{CR}\ {}^{CR}\boldsymbol{r}_i. \tag{4}$$

${}^{IR}\boldsymbol{r}_i$, ${}^{IL}\boldsymbol{r}_i$ is described by the following (5) through projection matrix $\boldsymbol{P}$.

$$ {}^{IL}\boldsymbol{r}_i = \boldsymbol{P}\ {}^{CL}\boldsymbol{r}_i = \boldsymbol{P}{}^{CL}\boldsymbol{T}_{CR}\ .{}^{CR}\boldsymbol{r}_i \tag{5}$$

Then position vectors projected in the Σ_{IR} and Σ_{IL} of arbitrary point i on target object can be described ${}^{IR}\boldsymbol{r}_i$ and ${}^{IL}\boldsymbol{r}_i$. Here, position and orientation of $\vec{\Sigma}_M$ based on $\vec{\Sigma}_{CR}$ has been defined as $\boldsymbol{\psi}_M$, which means ${}^{CR}\boldsymbol{T}_M$ in (1) is determined by $\boldsymbol{\psi}_M$. Then (3), (5) are rewritten as,

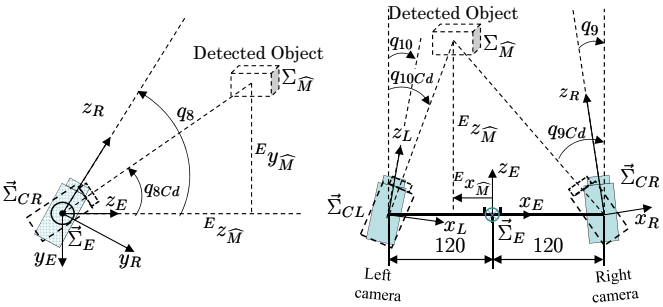
$$ \begin{cases} {}^{IR}\boldsymbol{r}_i = \boldsymbol{f}_R(\boldsymbol{\psi}_M,\ {}^{M}\boldsymbol{r}_i) \\ {}^{IL}\boldsymbol{r}_i = \boldsymbol{f}_L(\boldsymbol{\psi}_M,\ {}^{M}\boldsymbol{r}_i). \end{cases} \tag{6}$$

This relation connects the arbitrary points on the object and projected points on the left and right images corresponding to a 3-D pose $\boldsymbol{\psi}_M$ of the object. The measurement of $\boldsymbol{\psi}_M(t)$ in real time will be solved by consistent convergence of a matching model to the target object by a "Real-Time Multi-Step GA (RT-MS GA)" which will be explained in next subsection.

B. Model-Based Matching

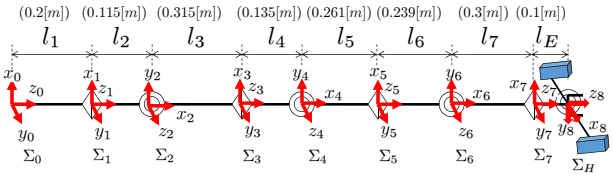
As shown in Fig.2 (on the top) we design a 3D solid model named S in a rectangular block to track the 3D-ball-object. The set of coordinates inside of the model is depicted as S_{in} and the outside space enveloping S_{in} is denoted as S_{out}. Projecting $S_{L,in}(\boldsymbol{\psi}_M)$ and $S_{out}(\boldsymbol{\psi}_M)$ onto the 2-D coordinates of left camera Σ_{IL}, points' set have been projected points' set to Σ_{IL}.

The projection for the right camera is in the same way. The left and right 2-D searching models, named S_L and S_R, are shown in Fig.2 (on the bottom).

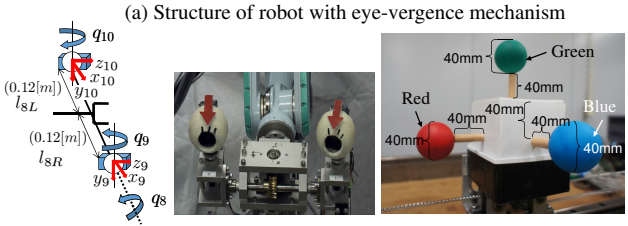


(a) Tilt angle q_8 (b) Pan angle q_9 and q_{10}

Fig. 3. Definition of tilt and pan angles with relation of detected object



(a) Structure of robot with eye-vergence mechanism



(b) Eye-vergence mechanism (c) Photo of Eye-vergence (d) 3D target marker

Fig. 4. Frame structure of eye-vergence manipulator and 3D target marker

Supposing there are distributed solid models in the searching space in Σ_W, each has its own pose $\boldsymbol{\psi}_M$. To determine which solid model is most close to the real target, a correlation function used fitness function in GA is defined for evaluation.

The input images will be directly matched by the projected moving models S_L and S_R, which are located only by $\boldsymbol{\psi}_M$ that includes the kinematical relations of the left and right camera coordinates. Therefore, if the camera parameters and kinematical relations are completely accurate, and the solid searching model describes precisely the target object shape, then the $S_{L,in}$ and $S_{R,in}$ will completely lie on the target reflected on the left and right images, provided that true value of $\boldsymbol{\psi}_M$ is found by GA's evolving calculations.

Color information is used to search for the target object in the images. We use the HSV space [11] to recognized the color. Because the hue of HSV space is hard be influenced by the light change. Let $h({}^{IL}\boldsymbol{r}_i)$ denote the hue value of i-th point in the searching model $S_{L,in}$ and $h({}^{IL}\boldsymbol{r}_j)$ denote the one of j-th point which is in $S_{L,out}$ in Fig.2. And the h is predetermined according to the known color of model.

Let b_i denote the hue value of point at the real image

which is overlap with i-th model point.

Then the evaluation function of the left moving surface-strips model is given as,

$$F_L(\boldsymbol{\psi}_M) = \frac{1}{n_{in}+n_{out}}\left(\sum_{{}^{IL}\boldsymbol{r}_i \in S_{L,in}(\boldsymbol{\psi}_M)} \delta\Big(h({}^{IL}\boldsymbol{r}_i) - b_{ki}\Big) - \sum_{{}^{IL}\boldsymbol{r}_j \in S_{L,out}(\boldsymbol{\psi}_M)} \delta\Big(h({}^{IL}\boldsymbol{r}_j) - b_{kj}\Big)\right) \tag{7}$$

where n_{in} represents the sum of the sampling points in $S_{L,in}$, n_{out} is the one in $S_{L,out}$ and δ is defined as

$$\delta(n) \quad = \quad \begin{cases} 1 & (n < 20) \\ 0 & (n \geq 20). \end{cases}$$

In the case of $F_L(\boldsymbol{\psi}_M) < 0$, $F_L(\boldsymbol{\psi}_M)$ is given to zero. The first part of this function expresses how much each color area of $S_{L,in}$ defined by $\boldsymbol{\psi}_M$ lies on the target being imaged on the left and right cameras. And the second part of (7) means the matching degree of its contour-strips. The difference between the internal surface and the contour-strips of the surface-strips model can make the estimation more sensible, especially in distance recognition between the target to the cameras which determine the size of the projected model onto the image plane. The right one is defined in the same way. Then the whole evaluation function is given as

$$F(\boldsymbol{\psi}_M) = [F_L(\boldsymbol{\psi}_M) + F_R(\boldsymbol{\psi}_M)]/2. \tag{8}$$

Equation (8) is used as a fitness function in GA process. When the searching model fits to the target object being imaged in the right and left images, then the fitness function $F(\boldsymbol{\psi}_M)$ gives maximum value, i.e., $F = 1$.

Therefore the problem of finding a target object and detecting its position/orientation can be converted to searching $\boldsymbol{\psi}_M$ that maximizes $F(\boldsymbol{\psi}_M)$. This optimization problem is solved by GA, which will be explained in the next section. The genes of GA representing possible pose solution $\boldsymbol{\psi}_{GA}$ is defined as,

$$\underbrace{\overbrace{01\cdots01}^{t_x}\overbrace{00\cdots01}^{t_y}\overbrace{11\cdots01}^{t_z}\overbrace{01\cdots01}^{\epsilon_1}\overbrace{01\cdots11}^{\epsilon_2}\overbrace{01\cdots10}^{\epsilon_3}}_{12bit\quad 12bit\quad 12bit\quad 12bit\quad 12bit\quad 12bit}.$$

C. On-line Pose Tracking "Real-Time Multi-Step GA"

For real-time visual control purpose, GA has been employed in a way denoted as "1-Step GA" evolution [12]. The used cameras' frame rate is 30fps. That means every 33ms cameras output a new image to a computer. In the past, subject to computing speed of the computer, GA explore process per frame can be done only once, so it was called as "1-Step GA." With advances in computing power of computers the system can now explore multiple GA explore processes in each frame (actually 9 times), and accuracy has also been improved. Now it is renamed as "Real-Time Multi-Step GA [13]."

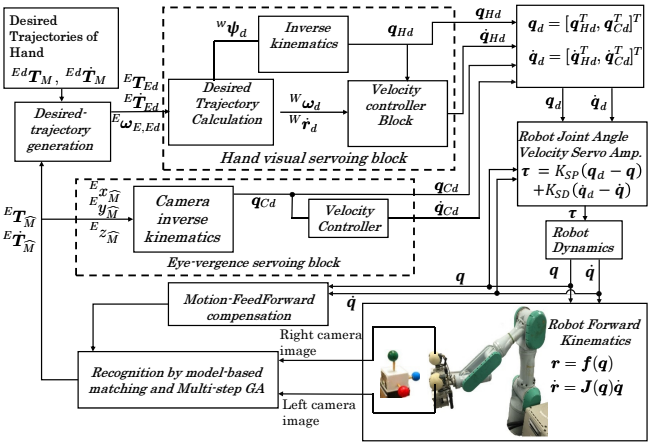


Fig. 5. Block diagram of the hand visual servoing system

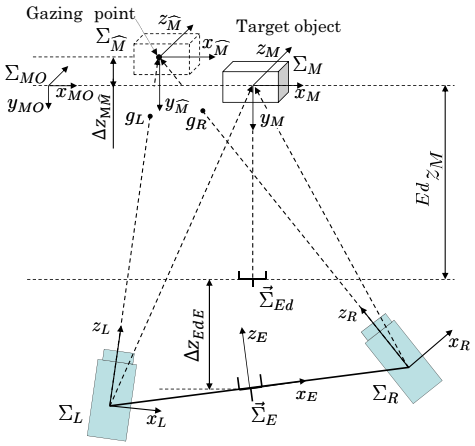


Fig. 6. Enlarged drawing of Fig.8 with gazing point. Pay attention to gazing point G, detected object $\vec{\Sigma}_{\widehat{M}}$ and target object $\vec{\Sigma}_M$. They are different from each other.

III. HAND & EYE VISUAL SERVOING CONTROLLER

A. Hand Visual Servoing Controller

The block diagram of our proposed hand & eye-vergence visual servoing controller is shown in Fig.5. The hand-visual servoing is the outer loop.

Based on the above analysis of the desired-trajectory generation, the desired hand velocity ${}^W\dot{\boldsymbol{r}}_d$ is calculated as,

$${}^W\dot{\boldsymbol{r}}_d = \boldsymbol{K}_{PP}{}^W\boldsymbol{r}_{E,Ed} + \boldsymbol{K}_{VP}{}^W\dot{\boldsymbol{r}}_{E,Ed}, \tag{9}$$

where ${}^W\boldsymbol{r}_{E,Ed}, {}^W\dot{\boldsymbol{r}}_{E,Ed}$ can be calculated from ${}^E\boldsymbol{T}_{Ed}$ and ${}^E\dot{\boldsymbol{T}}_{Ed}$. $\boldsymbol{K}_{PP}$ and $\boldsymbol{K}_{VP}$ are positive definite matrix to determine PD gain.

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

$${}^W\boldsymbol{\omega}_d = \boldsymbol{K}_{PO}{}^W\boldsymbol{R}_E{}^E\Delta\boldsymbol{\epsilon} + \boldsymbol{K}_{VO}{}^W\boldsymbol{\omega}_{E,Ed}, \tag{10}$$

where ${}^E\Delta\boldsymbol{\epsilon}$ is a quaternion error [10] calculated from the pose tracking result, and ${}^W\boldsymbol{\omega}_{E,Ed}$ can be computed by transforming the base coordinates of ${}^E\boldsymbol{T}_{Ed}$ and ${}^E\dot{\boldsymbol{T}}_{Ed}$ from Σ_E to Σ_W. Also, $\boldsymbol{K}_{PO}$ and $\boldsymbol{K}_{VO}$ are suitable feedback

matrix gains. The desired hand pose is defined as ${}^{W}\boldsymbol{\psi}_d^T = [{}^{W}\boldsymbol{r}_d^T, {}^{W}\boldsymbol{\epsilon}_d^T]^T$

The manipulator is 7 links, and the end-effector has 6-DoF, so q_1 is made 0 to solve the redundancy problem. The desired joint variable $\boldsymbol{q}_{Ed} = [0, q_{2d}, \ldots, q_{7d}]^T$ and $\dot{\boldsymbol{q}}_{Ed}$ is obtained by

$$\boldsymbol{q}_{Ed} = \boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_d^T) \tag{11}$$

where $\boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_d^T)$ is the inverse kinematic function.

Using the inverse kinematics it can make the joint of angles approximately as the desired joint angles. The formula of the desired joint angles was defined in the controller as

$$\dot{\boldsymbol{q}}_{Ed} = \boldsymbol{K}_P(\boldsymbol{q}_{Ed} - \boldsymbol{q}_E) + \boldsymbol{J}_E^{+}(\boldsymbol{q}) \begin{bmatrix} {}^{W}\dot{\boldsymbol{r}}_d \\ {}^{W}\boldsymbol{\omega}_d \end{bmatrix} \tag{12}$$

where $\boldsymbol{K}_P$ is P positive gain and $\boldsymbol{J}_E^{+}(\boldsymbol{q})$ is the pseudo-inverse matrix of $\boldsymbol{J}_E(\boldsymbol{q})$, and $\boldsymbol{J}_E^{+}(\boldsymbol{q}) = \boldsymbol{J}_E^T(\boldsymbol{J}_E\boldsymbol{J}_E^T)^{-1}$.

The hardware control system of the velocity-based servo system of PA10 is expressed as

$$\boldsymbol{\tau} = \boldsymbol{K}_{SP}(\boldsymbol{q}_d - \boldsymbol{q}) + \boldsymbol{K}_{SD}(\dot{\boldsymbol{q}}_d - \dot{\boldsymbol{q}}) \tag{13}$$

where $\boldsymbol{K}_{SP}$ and $\boldsymbol{K}_{SD}$ are symmetric positive definite matrices to determine PD gain.

B. Eye-vergence Visual Servoing Controller

The eye-vergence visual servoing is the inner loop of the visual servoing system shown in Fig.5. In this paper, two pan-tilt cameras are used 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.3 (a) and (b), ${}^{E}x_{\widehat{M}}$, ${}^{E}y_{\widehat{M}}$, ${}^{E}z_{\widehat{M}}$ express position of the detected object in the end-effector coordinate. The desired angle of the camera joints are calculated by:

$$q_{8Cd} = atan2({}^{E}y_{\widehat{M}}, {}^{E}z_{\widehat{M}}) \tag{14}$$

$$q_{9Cd} = atan2(l_{8R} - {}^{E}x_{\widehat{M}}, {}^{E}z_{\widehat{M}}) \tag{15}$$

$$q_{10Cd} = atan2(l_{8L} + {}^{E}x_{\widehat{M}}, {}^{E}z_{\widehat{M}}) \tag{16}$$

where $l_{8L} = l_{8R} = 120[mm]$ that is the camera location in Σ_E. The center line of the left and right camera viewing directions are defined as the z-axis of camera coordinate systems respectively. The target joint angular velocity $\dot{\boldsymbol{q}}_{Cd} = [\dot{q}_{8Cd}, \dot{q}_{9Cd}, \dot{q}_{10Cd}]$ of Eye-Vergence is calculated by:

$$\dot{q}_{iCd} = K_P(q_{iCd} - q_i) \qquad (i = 8, 9, 10) \tag{17}$$

$\dot{q}_{iCd}$ is input to the pulse motor for a camera angle control as a pulse train. Where $K_P = 1$ is the spring constant.

Because the motion of camera motor is an open loop, it is controlled to rotate a certain degree without getting the actual angle during the rotation, which make the accurate camera angle cannot be got. So the desired camera angles are input in every 33ms, and the input is limited to a certain value.

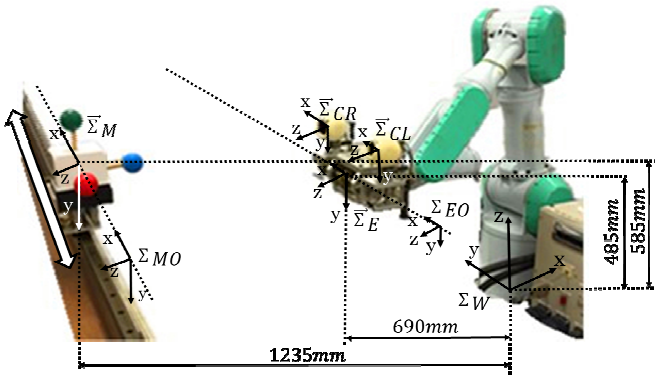


Fig. 7. Layout of the visual-servoing experiment

IV. EXPERIMENT OF HAND & EYE-VERGENCE VISUAL SERVOING

A. Experimental System

To verify the effectiveness of the hand & eye visual servoing system through real robot-PA-10 robot arm-manufactured by Mitsubishi Heavy Industries. Two rotatable cameras mounted on the end-effector are FCB-1X11A manufactured by Sony Industries. The frame frequency of stereo cameras is set as 30fps. The image processing board, CT-3001, receiving the image from the CCD camera is connected to the host computer(CPU: Intel Core i7-3770 , 3.40 GHz).

The structure of the manipulator and the cameras are shown in Fig.4 (a) and (b). And the 3D marker is shown in Fig.4 (d). The coordinate of the target object and the manipulator in experiment are shown in Fig.7.

B. Experiment Condition

EO, $MO$and EC represent the initial hand pose, the initial object pose and the midpoint of round-trip tracking movements of hand respectively. Therefore their coordinate systems are defined as Σ_{EO}, Σ_{EC} and Σ_{MO} separately. The homogeneous transformation matrix from Σ_W to Σ_{EC} and Σ_{MO} are:

$${}^{W}\boldsymbol{T}_{EC} = \begin{bmatrix} 0 & 0 & -1 & -690[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 485[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{18}$$

$${}^{W}\boldsymbol{T}_{MO} = \begin{bmatrix} 0 & 0 & -1 & -1235[mm] \\ 1 & 0 & 0 & -150[mm] \\ 0 & -1 & 0 & 585[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{19}$$

The target object move according to the following time function

$${}^{MO}z_M(t) = 150 - 150\cos(\omega t)[mm] \tag{20}$$

Target position and orientation relationship between the object and the end-effector is set as:

$${}^{Ed}\boldsymbol{\psi}_M = [0, -100[mm], 545[mm], 0, 0, 0] \tag{21}$$

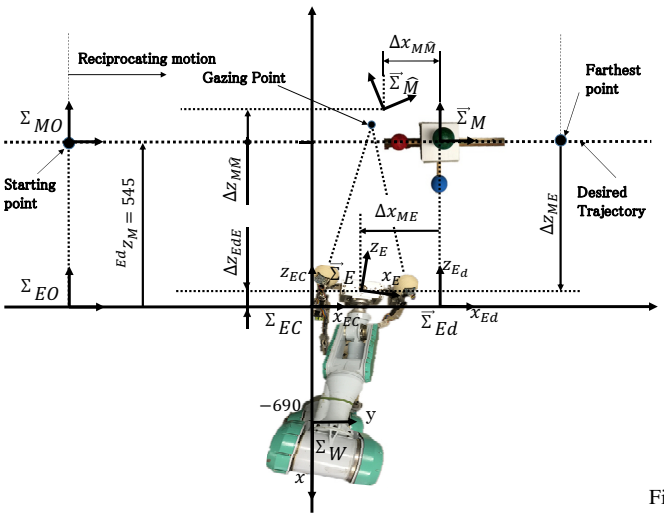


Fig. 8. Target object and definition of coordinates depicted in the x-z plane of Σ_{EC}. Initial position of the object is represented by Σ_{MO}; actual object $\vec{\Sigma}_M$; detected object $\vec{\Sigma}_{\widehat{M}}$; initial position of the hand Σ_{EO}; actual end effector $\vec{\Sigma}_E$; and desired end-effector $\vec{\Sigma}_{Ed}$.

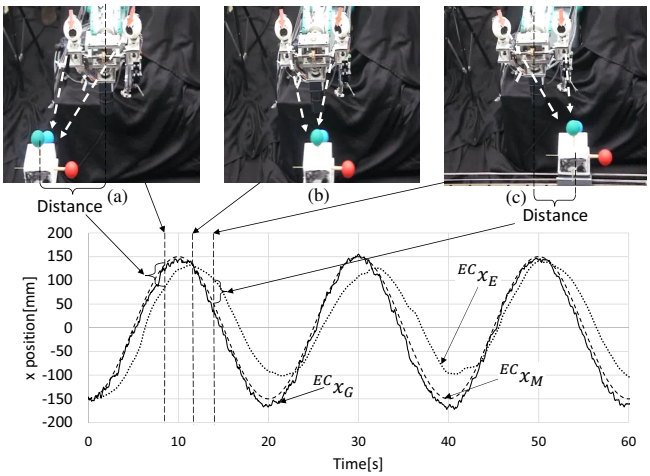


Fig. 9. Eye-vergence system and object position under 500lx light condition

The object is subjected to reciprocating motion of the sine wave in orbit. Pose relationship of the coordinate system of the object and the visual servoing system is shown in Fig.7.

C. Symbol Meaning

M represents the object and $\widehat{M}$ represents the estimated object. Then $\vec{\Sigma}_M$ denotes the coordinate system that moves along with the object. The relationship between coordinate systems such as the actual pose of the hand $\vec{\Sigma}_E$ or the recognized pose of the object $\vec{\Sigma}_{\widehat{M}}$ which is viewed from the x-z plane of the center coordinate system Σ_{EC} is shown in Fig.8. $\vec{\Sigma}$ represents that a coordinate system is running when it is viewed from the world coordinate system Σ_W. The coordinate system represented by Σ keeps fixed in the world coordinate system Σ_W. In other words $\vec{\Sigma}_E$, $\vec{\Sigma}_{Ed}$, $\vec{\Sigma}_M$ and $\vec{\Sigma}_{\widehat{M}}$are all moving in the world coordinate system Σ_W. On the other hand Σ_{EO}, Σ_{EC} and Σ_{MO} keeps fixed in the world coordinate system Σ_W. The motion of object$M$, hand$E$ and gazing point G in the x-axis direction of Σ_{EC} are represented by $^{EC}x_M$, $^{EC}x_E$ and $^{EC}x_G$. And as shown in Fig.10.

$$\Delta x_{EdE} = {}^{EC}x_{Ed} - {}^{EC}x_E \tag{22}$$

$$\Delta x_{MG} = {}^{EC}x_M - {}^{EC}x_G \tag{23}$$

represents a follow-up error of the hand and the gazing point respectively. According to the tracking relationship (21) $^{EC}x_{Ed} = {}^{EC}x_M$.

As it is shown in Fig.6, the intersection of both cameras' gazing directions is defined as the gazing point of cameras to examine trackability of the eye-vergence system. Because the gazing point has been calculated on the basis of the recognition result of the object by the Multi-Step GA, recognition error is included in the Gazing point.

D. Experiment Results

1) Relation between position diagram and the real machine: Fig.9 shows the positional relationship between the hand and the object in the condition that the tracking all the six position and orientation variables are recognized. And the motion period of the object is $T = 20$[s]. Movement trajectory of the object M, hand E and gazing point G are represented by dashed line, dotted line and solid line respectively. At the time of Fig.9 (b) the hand is just in front of the object. At the time of (a) and (c), since the moving velocity of the object is fast, hand is not able to track the object. Since the tracking state of hand is same as fixed camera system, the dotted line indicated by $^{EC}x_E$ in the figure also represents the movement of fixed camera system. At this time, it is clear that the distance between the hand $^{EC}x_E$ and the target object $^{EC}x_M$ on the x-axis direction is farther than that between the gazing point $^{EC}x_G$ and $^{EC}x_M$ of the camera. From the error between the gazing point $^{EC}x_G$ and the object $^{EC}x_M$ it can be seen that it is easier for eye-vergence system to track the object than the fixed camera system.

2) Position tracking result and analysis of the tracking experiment: In Fig.10 the pictures (1)∼(4) shows the images attached by left and right cameras under different experimental conditions. In the experiment we change the illumination of environment. Around object the illumination changes in the following order 500[lx], 1000[lx] and 2000[lx], followed by a disturbance experiment in which the object is partially covered for contrast.

Fig.10 (a) shows the fitness change during the experiment. It is clear that due to increased illumination the fitness value gradually reduce. However the influence of light changing is not so strong compared with partial cover. Because the fitness in the filed of partial occlusion is significantly lower than other stages.

In Fig.10 (b), with the same definition in Fig.9, movement trajectory of the object M, hand E and gazing point G are represented by dashed line, dotted line and solid line

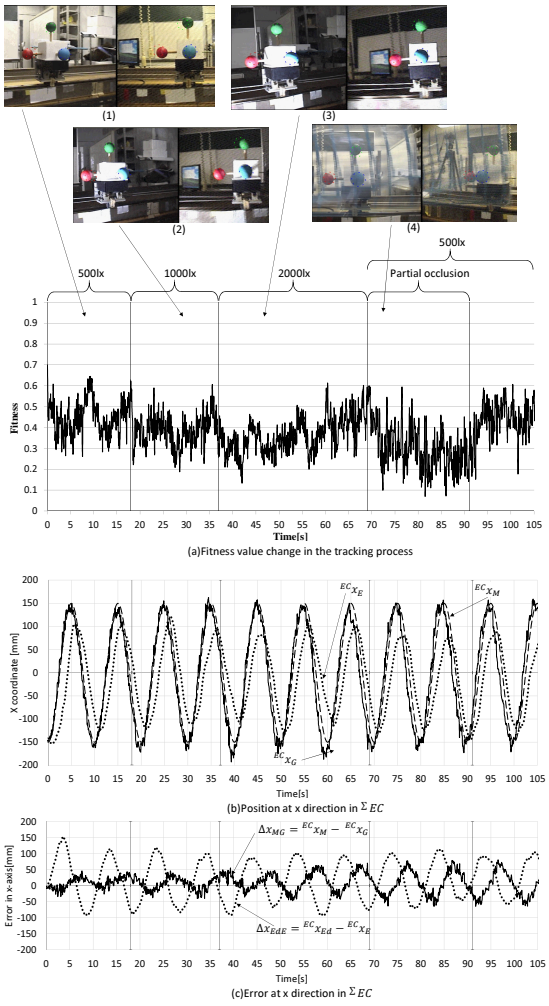


Fig. 10. Tracking result under different illumination and interference conditions

respectively. At this time, it is clear that the distance between the hand ${}^{EC}x_E$ and the target object ${}^{EC}x_M$ on the x-axis direction is farther than that between the gazing point ${}^{EC}x_G$ and ${}^{EC}x_M$ of the camera. From the error between ${}^{EC}x_G$ and ${}^{EC}x_M$, it can be seen that it is easier for eye-vergence system to track the object than the fixed camera system. It is obvious that after 50[s] compared with ${}^{EC}x_M$, ${}^{EC}x_G$ has a obvious phase delay. It is so bright that a part of the three balls is seen as white. Therefore they can not be correctly identified. It results in a reduction of fitness value in 2000[lx].

Fig.10 (c) shows the recognition error Δx_{MG} and tracking error Δx_{EdE}. With the gradual deterioration of the experimental conditions, the amplitude of the error of the gazing point Δx_{MG} gradually becomes larger. And the Δx_{EdE} change is not obvious.

V. CONCLUSION

In this paper the eye-vergence visual servoing controller have been described in detail. In order to evaluate the adaptability to light environment we designed lateral tracking experiment under different illumination conditions. Through the experiment it is verified that the system is not easily affected by ambient brightness and can continually track the object. Compared with the partial occlusion condition, the influence of changing of light condition is smaller.

In this experiment, we just considered about the illumination of light. And another important element that we need to consider is brightness of object. The brightness is the result of reflection. For the future we will think about the reflection and discuss the impact of brightness on trackability.

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] Wolfgang Sepp, Stefan Fuchs and Gerd Hirzinger, "Hierarchical Featureless Tracking for Position-Based 6-DoF Visual Servoing", Proceedings of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robotics and Systems (IROS), pp.4310-4315, 2006.

[3] 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

[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-inhand/Eye-to hand Cooperation, IEEE Int. Conf. on Robotics and Automation (ICRA), pp.2471-2476, 2007.

[8] Minami, Mamoru, Hidekazu Suzuki, Julien Agbanhan, and Toshiyuki Asakura. "Visual servoing to fish and catching using global/local GA search." In Advanced Intelligent Mechatronics, 2001. Proceedings. 2001 IEEE/ASME International Conference on, vol. 1, pp. 183-188. IEEE, 2001.

[9] Nishimura, Kenta, Sen Hou, Koichi Maeda, Mamoru Minami, and Akira Yanou. "Analyses on on-line evolutionary optimization performance for pose tracking while eye-vergence visual servoing." In 2013 IEEE International Conference on Mechatronics and Automation, pp. 698-703. IEEE, 2013.

[10] 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).

[11] Smith, Alvy Ray. "Color gamut transform pairs." ACM Siggraph Computer Graphics 12, no. 3 (1978): 12-19.

[12] M.Minami, W.Song, "Hand-eye-motion Invariant Pose Estimation with On-line 1-step GA -3D Pose Tracking Accuracy Evaluation in Dynamic Hand-eye Oscillation", Journal of Robotics and Mechatronics, Vol.21, No.6, pp.709-719 (2009.12)

[13] Myint, Myo, Kenta Yonemori, Akira Yanou, Khin Nwe Lwin, Mamoru Minami, and Shintaro Ishiyama. "Visual Servoing for Underwater Vehicle Using Dual-Eyes Evolutionary Real-Time Pose Tracking." Journal of Robotics and Mechatronics 28, no. 4 (2016): 543.

[14] Koichi Maeda, Mamoru Minami, Akira Yanou, Hiroaki Matsumoto, Fujia Yu, Sen Hou Frequency Response Experiments of 3-D Full Tracking Visual Servoing with Eye-Vergence Hand-Eye Robot SystemSICE Annual Conference pp.101-107 2012.

[15] Yoon, Youngrock, Guilherme N. DeSouza, and Avinash C. Kak. "Real-time tracking and pose estimation for industrial objects using geometric features." Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on. Vol. 3. IEEE, 2003.