The Twentieth International Symposium on Artificial Life and Robotics 2015 (AROB 20th 2015),
B-Con Plaza, Beppu, Japan, January 21-23, 2015

Modeling of Arbitrary-Shaped Objects by Evolving-Ellipsoids Approximation for Visual Servoing

Kenta Nishimura[1], Yusuke Sunami[1], Takayuki Matsuno[1], Akira Yanou[1] and Mamoru Minami[1]

[1] Graduate School of Natural Science and Technology, Okayama University Tsushimanaka 3-1-1 , Okayama, JAPAN.
(Tel: 81-86-251-8924)
[1]en20857@s.okayama-u.ac.jp

Abstract: Model-based visual servoing needs definitions of the objects and detected pose of the model is used for visual servo controller. When 3-D pose-based visual servoing is applied to 3-D unknown-shaped target, the model-based visual servoing method needs create the artificial 3-D model definitions for detecting the unknown 3-D target and tracking it. In order to have visual servoing system track the unknown-shaped target, automatic 3-D model generation mechanisms for arbitrary-shaped target objects should be created before detecting the designated target. In this paper, a 3-D modeling method of arbitrary shape objects is proposed and the performance is examined by modeling experiment of arbitrary shape objects and visual servoing experiments with twin hand-eye camera whose looking direction can rotate around pitching and yawing axes.

Keywords: Visual servoing, Eye-vergence

1 INTRODUCTION

Visual servoing is a control method of robot's motion through visual information in the feedback loop, which is obtained from visual cameras [1]-[4]. 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 done in the workspace [7]. These methods obtain different views to observe the object by increasing the number of cameras, leaving the system less adaptive for changing environment.

Image-based visual servoing has an advantage that it can be used under unknown environment, that is, it works without any definition or knowledge of environment. However, it has no concept of the shape of target object since Image-based visual servoing system could be constructed by just eliminating corresponding point-to-point distance errors between desired point position and actual point position in image plane. Therefore, the image-based visual servoing does not need to recognize target shape.

But there exists visual servoing application that the system be required to discern several objects' shape and to do visual servoing against one of the objects. In this case, target shape should be recognized and its pose should be detected as well. Therefore when shape recognition task and visual servoing task to the designated target are required to be simultaneously done, the model-based visual servoing method is superior than the image-based one since model-based method enables the system to understand the target shape based on the defined shape in computer and pose, enabling the robot system can handle the target object such as grasping and picking up. And that image-based visual servoing system are unable to do so.

The model-based visual servoing needs definition of shape of object target, which is called in this document as simply "model", which is thought to be demerit when comparing it with image-based method. However GA-based (Genetic-Algorithm-based) 3-D model-based pose detection method that we have proposed can obtain the pose of target object without extracting correspondent points between left and right images, where epipolar line is usually used to find corresponding points that represent the identical point on an object existing in 3-D real world.

To find corresponding two points in left and right camera using epipolar line sometimes becomes difficult due to dynamic luminous intensity changing or being hidden by something adrift in the foreground of the target. However 3-D model-based matching method dose not require to find corresponding points in the camera views since the 3-D model projected to left and right camera images has already corresponding points in the projected model shape.

This natural feature of model-based matching can made the 3-D pose detection robust against noises in images. The robustness of 3-D model based pose tracking has been conformed [8] and usefulness for pose tracking in dynamic/real-time video images, which is used for visual servoing against swimming fish [9]. However the mode-based method needs of course a model definition. This restricts an extent of applicable target of this method where model should be predetermined and known. In order to overcome this problem, automatic 3-D model generation mechanisms for arbitrary-shaped target objects is required created before detecting the designated target. In this paper to overcome the above obstacle to use model-based method against unknown object, we propose a modeling method of 3-D arbitrary shaped target

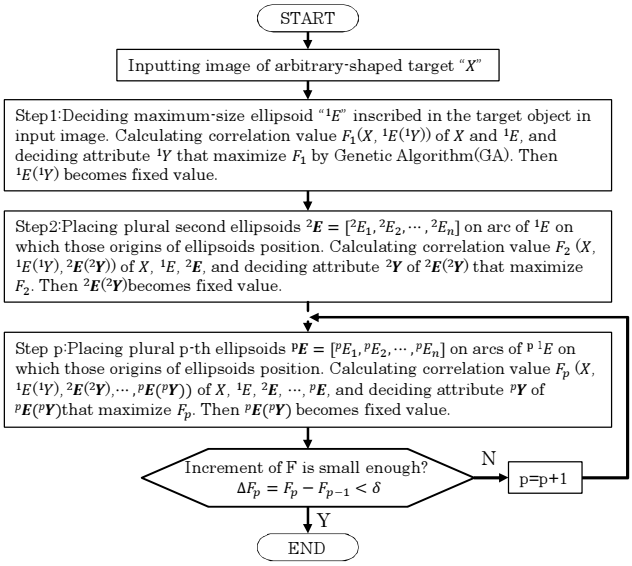

Fig. 1. Flowchart of approximation description of arbitrary-shaped target object by optimization through Genetic Algorithms by using progressive combination as a fitness-function.

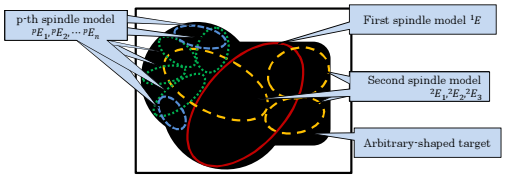

Fig. 2. Modeling method of arbitrary-shaped target object through progressive combinations.

object.

2 3-D ARBITRARY-SHAPE MODELLING

2.1 Needs for Arbitrary-shape Modelling

Robots used in factories need predetermined descriptions of target objects' shape, position and orientation, and that their information is crucial for adaptive behavior of robots. Thus the robots are hard to have leeways to deal some required tasks within changing environment since the changing could not be prescribed beforehand. To make a robot system that manages to operate autonomously in the unknown and unpredictable environment, it is required that the robot should perceive the changing in environment resulted from what the robots have done and should accomplish the task to be done adaptively and autonomously. Then this requirement indicates the abilities to perceive 3-D arbitrary-shaped target objects are essential for autonomous operations in changing environments. Therefore we propose in this paper to construct a recognition system of target object with arbitrary-

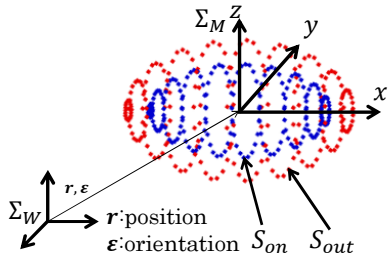

Fig. 3. 3-D model to approximate arbitrary-shape object for progressive combination.

shape by approximating the object with progressive combination of 3-D ellipsoid.

2.2 Progressive Ellipsoid Approximation

Here a 3-D ellipsoid object is considered as a fundamental element to approximate an arbitrary-shaped object to be perceived and to describe it in computer. The ellipsoid element description has an attribute including pose(position and orientation), radii of longer/shorter axes, color, etc.

Figure 1 shown a flowchart of description procedure for arbitrary-shaped object with progressive ellipsoid approximation. By optimizing a correlation function $F_1(X,{}^1E({}^1Y))$ between target object "X" in input image and attribute of first 3-D ellipsoid model "${}^1Y$" that has a peak when the first 3-D ellipsoid model is bounded internally in the target object and that the size and pose are chosen as large as the projected ellipsoid to be matched to the target object in the image, the first 3-D ellipsoid model could be determined, as a first step.

In step 2, second 3-D ellipsoid models that have their center positions placed on the surface of the first model with attributes ${}^2\boldsymbol{E} = [{}^2E_1,{}^2E_2,\cdots,{}^2E_n]$ are determined by the same procedures as the first one that maximizes the correlation value $F_2(X,{}^1E({}^1Y),{}^2\boldsymbol{E}({}^2\boldsymbol{Y}))$ by changing the attributions ${}^2\boldsymbol{Y} = [{}^2Y_1,{}^2Y_2,\cdots,{}^2Y_n]$ of the secondary models.

By repeating above procedures p times, the accumulated correlation function $F_p(X,{}^1E({}^1Y),{}^2\boldsymbol{E}({}^2\boldsymbol{Y}),\cdots,{}^p\boldsymbol{E}({}^p\boldsymbol{Y}))$ that includes first, second, $\cdots$, p-th can be calculated.

When the incremental changing of $\Delta F_p = F_p - F_{p-1}$ is less than predetermined δ, the approximation procedure steps. To make the approximation accuracy high, the δ should be set small, thus the requirement of how much precisely the 3-D ellipsoid series describe under unknown-shaped target can be adjusted through setting the value of δ.

Figure 2 shows the models determined by Genetic Algorithms, where one first model ${}^1E$, three second models ${}^2E_1,{}^2E_2,{}^2E_3,\cdots$ are depicted, where are determined by flow chart in Fig.1.

2.3 Definition of 3-D Ellipsoid Model

A 3-D ellipsoid model is shown in Fig.3. The model consists of points that exist on the surface of the ellipsoid "S^p_{on}", which includes m points and outside the one "S^p_{out}" that envelopes the S^p_{on}. "q" means the number of the 3-D model points used in 1-step GA process. A set of points on the surface S^p_{on} of p-th model is determined by

$$\begin{cases} x^p_{on,i} = a\sin\theta_i \\ y^p_{on,i} = b\cos\theta_i\sin\phi_i \quad (-\pi < \theta_i, \phi_i < +\pi \\ z^p_{on,i} = c\cos\theta_i\cos\phi_i \quad \text{and } i = 1, 2, \cdots, m) \end{cases} \tag{1}$$

and a set of points on the enveloping surface S^p_{out} with the radii being $a+\alpha$, $b+\alpha$, $c+\alpha$ along to each x,y,z axis is defined as

$$\begin{cases} x^p_{out,i} = (a+\alpha)\sin\theta_i \\ y^p_{out,i} = (b+\alpha)\cos\theta_i\sin\phi_i \ \ (-\pi < \theta_i, \phi_i < +\pi \\ z^p_{out,i} = (c+\alpha)\cos\theta_i\cos\phi_i \ \ \text{and } i = 1, 2, \cdots, m). \end{cases} \tag{2}$$

Then a set of the p-th 3-D ellipsoid model is defined based on Σ_M as

$$^M S^p_{on} = \{(x,y,z)|^M \boldsymbol{r}^p_{on} = [x^p_{on,i}, y^p_{on,i}, z^p_{on,i}], (i = 1, 2, \cdots, m)\} \tag{3}$$

$$^M S^p_{out} = \{(x,y,z)|^M \boldsymbol{r}^p_{out} = [x^p_{out,i}, y^p_{out,i}, z^p_{out,i}], (i = 1, 2, \cdots, m)\} \tag{4}$$

$$^M S^p = {}^M S^p_{on} \cap {}^M S^p_{out}. \tag{5}$$

Suppose that the p-th 3-D ellipsoid model positions based on the coordinate of Σ_M that is set at the center of 3-D model as shown in Fig.3 at the point of $^M\boldsymbol{r}_p$ and orientation $^M\boldsymbol{\epsilon}_p$ that is represented by unit quaternion, $\{\eta, \boldsymbol{\epsilon}\}$ and $\eta^2 + \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = 1$, then the each point in $^M S^P$ is translated into the pose based on Σ_W as

$$^W S^p = \{(x,y,z)|^W \boldsymbol{r}^p = \boldsymbol{r} + {}^W\boldsymbol{R}_M(\boldsymbol{\epsilon})^M\boldsymbol{r}^p, {}^M\boldsymbol{r}^p \in {}^M S^p\} \tag{6}$$

where $^W\boldsymbol{R}_M(\boldsymbol{\epsilon})$ is 3×3 orientation matrix determined by quaternion. The set of points in $^W S^P$ is projected to image plane of left and right camera, then the 3-D model-based correlation function is calculated based on points in $^W S^P$.

Then the best matched pose, that is $\boldsymbol{r}$ and $\boldsymbol{\epsilon}$, which is thought to represent pose of the real target object in 3-D space, can be identified.

3 OBJECT RECOGNITION METHOD

3.1 Model-based Matching Method

In this part, a model-based matching method was presented. The images input from a right-and-left video cameras are composed by hue value ranging from 0 to 360. $S_{R,in}$, and $S_{L,in}$ are the inside spaces of coordinate on the surface of the block model. Accordingly, $S_{R,out}$, and $S_{L,out}$ are the outside spaces. The H value of right image at the position $^{IR}\boldsymbol{r}_i$ is expressed as $p(^{IR}\boldsymbol{r}_i)$, and the H value of left image at the position $^{IL}\boldsymbol{r}_i$ is expressed as $p(^{IL}\boldsymbol{r}_i)$. $^{IR}\boldsymbol{r}_i$ and $^{IL}\boldsymbol{r}_i$ are the positions of the pixels. If it was defined that $\boldsymbol{m}(\boldsymbol{r}) = 1(\boldsymbol{r} \in \boldsymbol{S}_{R,in}$, or $\boldsymbol{r} \in \boldsymbol{S}_{L,in})$, $\boldsymbol{m}(\boldsymbol{r}) = -1(\boldsymbol{r} \in \boldsymbol{S}_{R,out}$, or $\boldsymbol{r} \in \boldsymbol{S}_{L,out})$, and $\boldsymbol{\phi}$ was the position/orientation of the model, then the fitting evaluation function was

$$\begin{aligned} F(\boldsymbol{\phi}) = \sum_{j=1}^{p} \Bigg\{ & \Big(\sum_{^{IR}\boldsymbol{r}_i \in S^j_{R,in}(\boldsymbol{\phi})} m * p(^{IR}\boldsymbol{r}_i) + \sum_{^{IR}\boldsymbol{r}_i \in S^j_{R,out}(\boldsymbol{\phi})} m * p(^{IR}\boldsymbol{r}_i) \Big) \\ & + \Big(\sum_{^{IL}\boldsymbol{r}_i \in S^j_{L,in}(\boldsymbol{\phi})} m * p(^{IL}\boldsymbol{r}_i) + \sum_{^{IL}\boldsymbol{r}_i \in S^j_{L,out}(\boldsymbol{\phi})} m * p(^{IL}\boldsymbol{r}_i) \Big) \Bigg\} /2 \\ = \sum_{j=1}^{p} & \Big\{ F^j_R(\boldsymbol{\phi}) + F^j_L(\boldsymbol{\phi}) \Big\} /2 \end{aligned} \tag{7}$$

Equation (7) was used as a fitness function in GA process. When the moving searching model fitted to the target object being imaged in the right and left images, the fitness function $\boldsymbol{F}(\boldsymbol{\phi})$ got maximum value.

Therefore, the problem of finding a target object and detecting its pose can be converted to searching $\boldsymbol{\phi}$ that maximize $\boldsymbol{F}(\boldsymbol{\phi})$. This problem will be solved with 1-step GA algorithm.

3.2 On-line Pose Tracking "1-step GA"

For real-time visual control purposes, we employ GA in a way that we denoted as "1-Step GA"[10] evolution in which the GA evolutional iteration is applied one time to the newly input image. While using the elitist model of the GA, the position/orientation of a target can be detect in every new image by that of the searching model given by the best individual in the population. This feature happens to be favorable for real-time visual recognition. We output the current best individual of the GA in every newly input image, and use it as real-time recognition result.

4 EYE-VERGENCE VISUAL SERVOING CONTROLLER

4.1 Desired-trajectory generation

In Fig.4, the world coordinate frame is denoted by Σ_W, the target coordinate frame is denoted by Σ_M, and the desired and actual end-effector coordinate frame is denoted by Σ_{Ed}, Σ_E respectively. The desired 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 pose Σ_{Ed} and the actual end-effector pose Σ_E is denoted as $^E\boldsymbol{T}_{Ed}$, which can be described by:

$$^E\boldsymbol{T}_{Ed}(t) = {}^E\boldsymbol{T}_M(t)^{Ed}\boldsymbol{T}^{-1}_M(t) \tag{8}$$

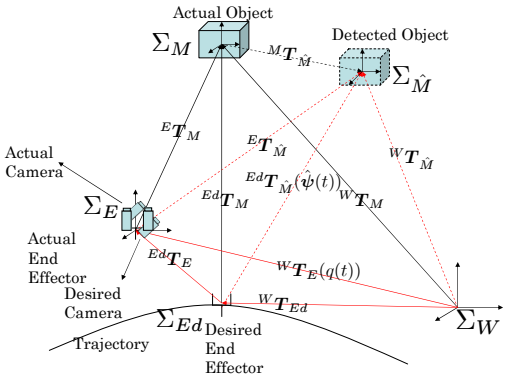

Fig. 4. Motion of the end-effector and object

(8) is a general representation of hand pose tracking error that satisfies arbitrary object motion ${}^{W}\boldsymbol{T}_{M}(t)$ and arbitrary visual servoing objective ${}^{Ed}\boldsymbol{T}_{M}(t)$. The relation ${}^{E}\boldsymbol{T}_{M}(t)$ can be estimated by 1-step GA [10], having been presented as an on-line model-based pose estimation method. Let $\Sigma_{\hat{M}}$ denote the detected object, It is natural there should always exist an error between the actual object Σ_{M} and the detected one $\Sigma_{\hat{M}}$. So in visual servoing, (8) will be rewritten based on $\Sigma_{\hat{M}}$ that includes the error ${}^{M}\boldsymbol{T}_{\hat{M}}$, as

$$ {}^{E}\boldsymbol{T}_{Ed}(t) \quad = \quad {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{Ed}\boldsymbol{T}_{\hat{M}}^{-1}(t). \tag{9} $$

Differentiating (9) with respect to time yields

$$ {}^{E}\dot{\boldsymbol{T}}_{Ed}(t) = {}^{E}\dot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\boldsymbol{T}_{Ed}(t) + {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{\hat{M}}\dot{\boldsymbol{T}}_{Ed}(t). \tag{10} $$

Differentiating (10) with respect to time again

$$ {}^{E}\ddot{\boldsymbol{T}}_{Ed}(t) = {}^{E}\ddot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\boldsymbol{T}_{Ed}(t) + 2\,{}^{E}\dot{\boldsymbol{T}}_{\hat{M}}(t)\,{}^{\hat{M}}\dot{\boldsymbol{T}}_{Ed}(t) + {}^{E}\boldsymbol{T}_{\hat{M}}(t)\,{}^{\hat{M}}\ddot{\boldsymbol{T}}_{Ed}(t), \tag{11} $$

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.4, there are two errors left to be decreased in the visual system. 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, ${}^{E}\boldsymbol{T}_{Ed}$. The error of ${}^{M}\boldsymbol{T}_{\hat{M}}$ is decreased by pose tracking method of the "1-step GA" [10], and the eye-vergence camera system, and the error of ${}^{E}\boldsymbol{T}_{Ed}$ depends on the performances of the hand visual servoing controller.

4.2 Hand Visual Servoing Controller

The block diagram of the eye-vergence visual servoing controller is shown as Fig.5. Based on the above analysis of the desired-trajectory generation, the desired hand velocity ${}^{W}\dot{\boldsymbol{r}}_{d}$ is calculated as,

$$ {}^{W}\dot{\boldsymbol{r}}_{d} = \boldsymbol{K}_{P_p}\,{}^{W}\boldsymbol{r}_{E,Ed} + \boldsymbol{K}_{V_p}\,{}^{W}\dot{\boldsymbol{r}}_{E,Ed}, \tag{12} $$

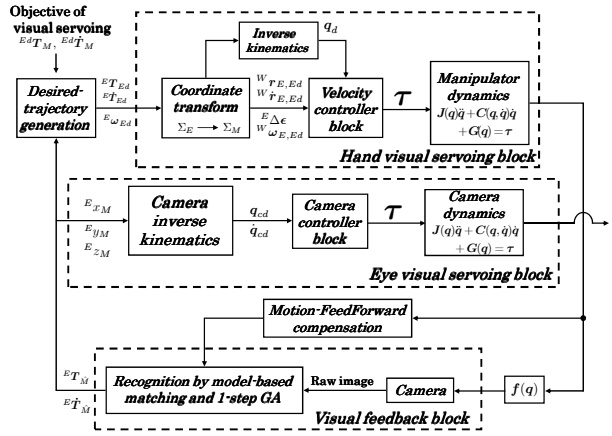

Fig. 5. Block diagram of Eye-Vergence visual servoing controller

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

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

$$ {}^{W}\boldsymbol{\omega}_{d} = \boldsymbol{K}_{P_o}\,{}^{W}\boldsymbol{R}_{E}\,{}^{E}\Delta\boldsymbol{\epsilon} + \boldsymbol{K}_{V_o}\,{}^{W}\boldsymbol{\omega}_{E,Ed}, \tag{13} $$

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, $\boldsymbol{K}_{P_o}$ and $\boldsymbol{K}_{V_o}$ are suitable feedback matrix gains. We define the desired hand pose as ${}^{W}\boldsymbol{\psi}_{d}^{T} = [{}^{W}\boldsymbol{r}_{d}^{T}, {}^{W}\boldsymbol{\epsilon}_{d}^{T}]^{T}$

The desired joint variable $\boldsymbol{q}_{Ed} = [q_{1d}, ..., q_{7d}]^{T}$ and $\dot{\boldsymbol{q}}_{Ed}$ is obtained by

$$ \boldsymbol{q}_{Ed} \quad = \quad \boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_{d}^{T}) \tag{14} $$

$$ \dot{\boldsymbol{q}}_{Ed} \quad = \quad \boldsymbol{J}_{E}^{+}(\boldsymbol{q}) \begin{bmatrix} {}^{W}\dot{\boldsymbol{r}}_{d} \\ {}^{W}\boldsymbol{\omega}_{d} \end{bmatrix} \tag{15} $$

where $\boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_{d}^{T})$ is the inverse kinematic function 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 robot arm is a 7 links manipulator, and the end-effector has 6-DoF, so it has a redundance. In the research before, we only calculated the position of the manipulator's end-effector, but not considering the joint angles through the position of the manipulator's end-effector. For one end-effector pose, there may exist infinite kinds of shapes, which will make the system dangerous. In this report, we made q_1 is 0, and used the inverse kinematics to calculate all joint angles. It can solve the redundancy problem. Meanwhile we took a controller to make the joint of angles approximately as the desired joint angles. So we defined the formula of the desired joint angles in the new controller as

$$ \dot{\boldsymbol{q}}_{Ed} = \boldsymbol{k}_{p}(\boldsymbol{q}_{Ed} - \boldsymbol{q}_{E}) + \boldsymbol{J}_{E}^{+}(\boldsymbol{q}) \begin{bmatrix} {}^{W}\dot{\boldsymbol{r}}_{d} \\ {}^{W}\boldsymbol{\omega}_{d} \end{bmatrix} \tag{16} $$

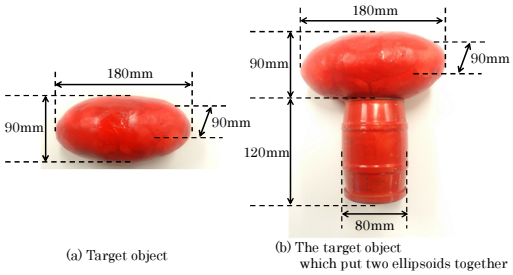

Fig. 6. target object.

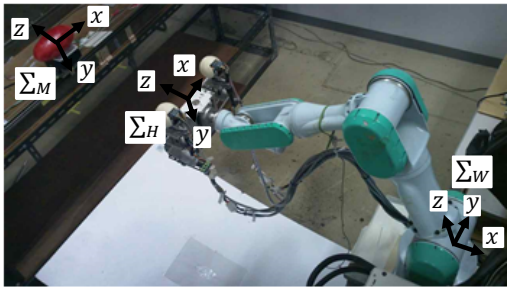

Fig. 7. Target object and Visual servoing system.

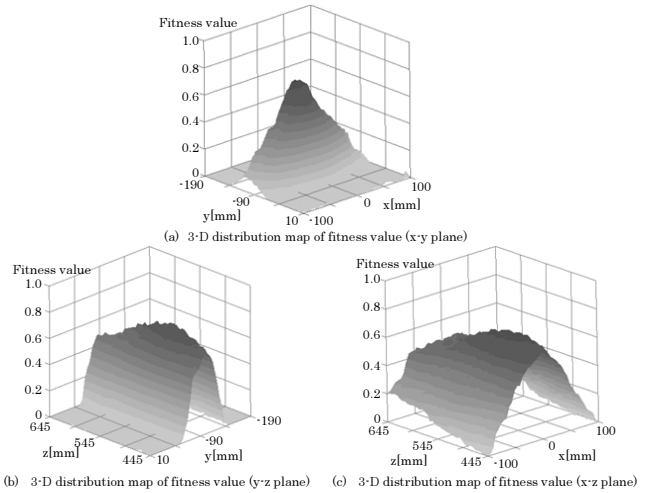

Fig. 8. 3-D distribution map of fitness value.

where $\boldsymbol{k}_p$ is P positive gain.

5 EXPERIMENT

5.1 The confirmation of the proposed method

To judge whether a target object can be searched, a distribution map of fitness value of the object in fig.6(a) is made. Manipulator coordinate system is shown in fig.7. The relation between the object and the end-effector Σ_H is set as $(x, y, z, \epsilon_1, \epsilon_2, \epsilon_3) = (0, -100, 530, 0, 0, 0)$. The conformity was calculated every 5[mm] in a range of the x-axis direction from -100[mm] to 100[mm], the y-axis direction from -190[mm] to 10[mm], the z-axis direction from 445[mm] to 645[mm] by an experiment. The result of the experiment is shown in Fig.8. The distribution maps of the fitness value

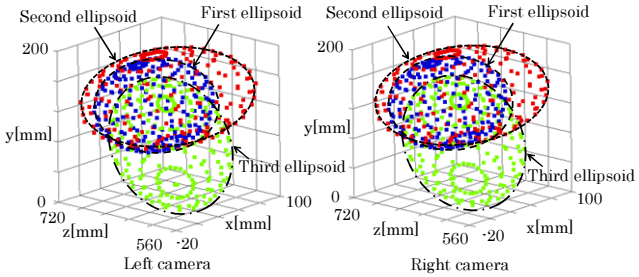

Fig. 9. Recognition result using three ellipsoids.

is drawn as mountain shapes, such as (a)x-y plane, (b)y-z plane, and (c)x-z plane. The fitness value described in color depth, maximize at the top of the mountain where the color is the darkest. In fig.8(a), the position expressed by the maximum of the fitness value is the same as the real position of the object. However the distribution of the maximum in the z direction is wider, which is shown as Fig.8(b) and (c).

5.2 Recognition experiment of arbitrary shape objects

The target object which put two ellipsoids together is shown in Fig.19(b). In a recognition experiment, 3 ellipsoids were used as a target object. The result of this experiment is shown in Fig.9. Two ellipsoids with different size, first and second ellipsoid, are used to be the model of the ellipsoid on the upper side shown in Fig.19(b). Then the third ellipsoid in Fig.9 model the ellipsoid on the underside shown in Fig.**??**. From the above, shape of the target object is matched with the modeling results. But, there is an position error in x and z axis.

5.3 Visual Servoing Experiment

5.3.1 Experiment condition

Visual servoing is performed by using the ellipsoid shown in fig.6(a). The initial hand pose is defined as Σ_{E_0}, and the initial object pose is defined as Σ_{M_0}. The homogeneous transformation matrix from Σ_W to Σ_{E_0} and from Σ_W to Σ_{M_0} are:

$$ {}^W\boldsymbol{T}_{E_0} = \begin{bmatrix} 0 & 0 & -1 & -890[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 500[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{17} $$

$$ {}^W\boldsymbol{T}_{M_0} = \begin{bmatrix} 0 & 0 & -1 & -1435[mm] \\ 1 & 0 & 0 & -100[mm] \\ 0 & -1 & 0 & 500[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{18} $$

The target object moves according to the following function as:

$$^{M_0}x_M(t) \quad = \quad 150 - 150\cos(\omega t)[mm]. \qquad (19)$$

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

$$^{Ed}\boldsymbol{\psi}_M = [0, -90[mm], 545[mm], 0, 0, 0]. \qquad (20)$$

the visual servoing experiments is conducted in two conditions $\omega = 0.209$ (Period: $T = 30[s]$) and $\omega = 0.628$ (Period: $T = 10[s]$). In this experiment, position (x, y, z coordinates) was unknown and a true value was given to posture. In this case, the variables in the genes of the 1-step-GA pose-tracking-system are only x, y, z, and the variables for orientation defined by quaternion are fixed as a true value and not updated by 1-step GA. After Modeling by the primary ellipsoid of a target object, visual servoing is performed based on the position and size which are given by the modeling.

5.3.2 Experiment result

The experiment result are shown in Fig.10 and Fig.11. A modeling result by the primary ellipsoid of a target object is an ellipsoid with the size of major axis 90 [mm] and minor axis 45 [mm], the position of target object is $(x, y, z) = (0, -90, 545)$. This modeling experiment result agrees with both the target object shown in Fig.6, and the initial position. In the visual servoing based on modeling result, End-effector has the phase delay, but GA is able to recognize the target object(Fig.10,Fig.11). However, in fig.11, recognition becomes unstable with the time passed. It is because that a range giving the biggest fitness value in the z direction is wide, unevenness occurs in the recognition of the z direction, and the result has a bad influence upon the x-direction.

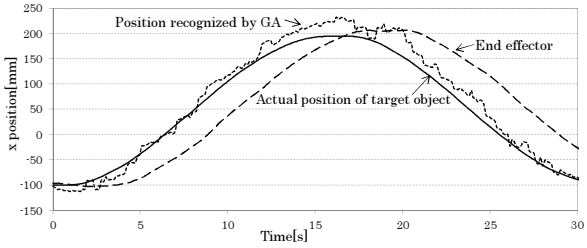

Fig. 10. $\omega = 0.209$

6 CONCLUSION

We modeled arbitrary-shaped object by performing the 3-D modeling which used the ellipsoid. It was confirmed that a target can be recognized from the modeling result. In the case of the one ellipsoid, we modeled the target object and applied it to a visual servoing. We will improve the recognition in z direction and perform a visual servo using plural ellipsoids in future.

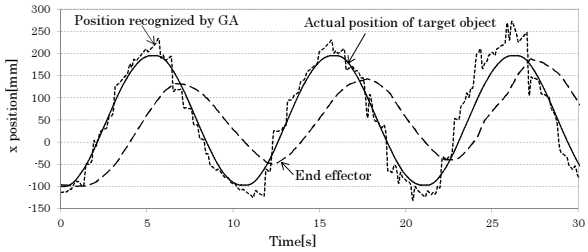

Fig. 11. $\omega = 0.628$

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] P.Y.Oh, and P.K.Allen, “Visual Servoing by Partitioning Degrees of Freedom”, IEEE Trans. on Robotics and Automation, vol. 17, no. 1, pp. 1-17, 2001.

[3] 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.

[4] P.K.Allen, A.Timchenko, B.Yoshimi, and P.Michelman, “Automated Tracking and Grasping of a Moving object with a Robotic Hand-Eye System”, IEEE Trans. on Robotics and Automation, vol. 9, no. 2, pp. 152-165, 1993.

[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] Wei. Song, M. Minami, Fujia Yu, Yanan Zhang and Akira Yanou “3-D Hand & Eye-Vergence Approaching Visual Servoing with Lyapunouv-Stable Pose Tracking”, IEEE Int. Conf. on Robotics and Automation (ICRA), pp.11, 2011.

[9] Keita Mori, Yuya Ito, Mamoru Minami and Akira Yanou, “Fish Catching Experiments to Overcome Fish’s Intelligence by Visual and Prediction Servoing in combination with Chaos and Random Motions”, Proceedings of SICE Annual Conference 2013, pp.1316-1321, 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).