International Journal of Applied Electromagnetics and Mechanics 52 (2016) 867–873
DOI 10.3233/JAE-162205
IOS Press

6-dof Eye-vergence visual servoing by 1-step GA pose tracking

Yu Cui, Kenta Nishimura, Yusuke Sunami, Mamoru Minami* and Akira Yanou
Graduate School of Nature Science and Technology, Okayama University, Okayama, Japan

Abstract. Visual servoing to moving target with hand-eye cameras fixed at hand is inevitably affected by hand dynamical oscillations, then it is hard to keep target at the centre of camera's image, since nonlinear dynamical effects of whole manipulator stand against tracking abilities of robots. In order to solve this problem, we come up with a system where the visual servoing controller of the hand and eye-vergence is separated independently, so that the cameras can observe the target object at the center of the camera images through eye-vergence functions. The eye with light masses make the cameras' eye-sight direction rotate quickly than the hand's, so the track ability of the eye-vergence motion is superior to the one of hand's. In this report, merits of eye-vengence visual servoing for pose tracking have been confirmed through frequency response experiments on condition of full 6-dof pose being estimated in real time.

Keywords: Visual servoing, eye-vergence, GA

1. Introduction

Nowadays, in a field of robot vision, a control method called a visual servoing attracts attention [1–4]. Visual servoing, a method for controlling a robot using visual information in a feedback loop, is expected to be able to allow the robot to adapt to changing or unknown environments.

An eye-vergence visual servoing system, where the visual servoing controller of hand and eye-vergence is separated independently, is said to be superior to fixed camera visual servoing system [5], which have been reported in [1–4]. When humans want to keep tracking a fast-moving object, it is not easy to catch up with the object by head rotation motion, because it is hard for human to rotate the face to squarely position the object that moves quickly since the mass of head affects dynamically to the rotating motions. However human's eyes can keep staring at the object because of their small masses and inertial moments. On top of this, generally the final objective of visual servoing is deemed to lie in approaching the end-effector to a target object and then work on it, such as grasping. To achieve the above environment-adapting of visual servoing in which the desired distance between cameras and the target is time-varying, eye-vergence camera system is indispensable to keep suitable viewpoint all the time during the approaching visual servoing with camera rotation so as the cameras to observe the target at the center of camera images during the approaching motions.

To verify the superiority of eye-vergence system, frequency response experiments where target object moves with sinusoidal time profile have been conducted. In this report, the results of the experiments are presented. Based on them, it can be confirmed that eye-vergence have a better stability and traceability than fixed-eye visual servoing system.

*Corresponding author: Mamoru Minami, Graduate School of Nature Science and Technology, Okayama University, 3-1-1 Tsushimanaka, Okayama 700-8530, Japan. E-mail: minami-m@cc.okayama-u.ac.jp.

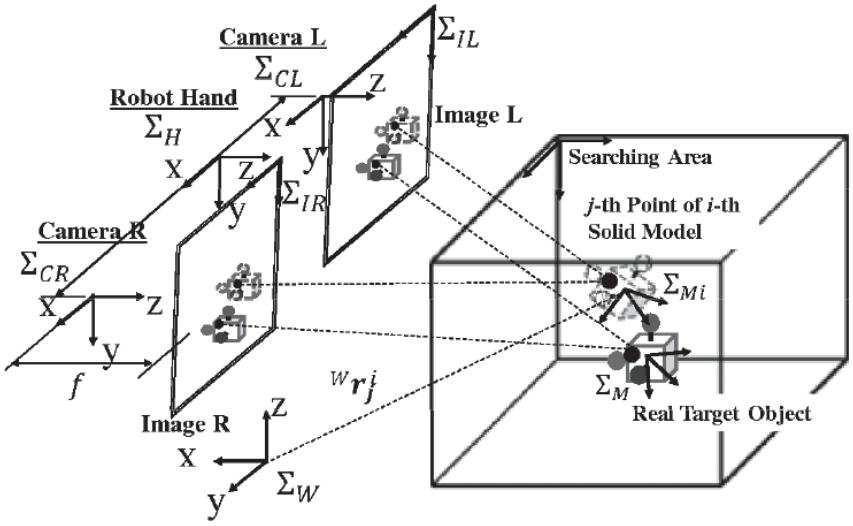


Fig. 1. Coordinate system of dual eyes – whose two cameras are set at the hand (represented by Σ_H) of robot – are defined as Σ_{CL} and Σ_{CR}. The coordinates of camera images are Σ_{IL} and Σ_{IR}, in which real target object defined by Σ_M and *i*-th solid model defined by Σ_{Mi} are projected from the 3D space depicted at right side).

2. 3D pose tracking method

2.1. Kinematics of stereo-vision

We utilize perspective projection as projection transformation. Figure 1 shows the coordinate system of the dual-eyes vision system. The target object's coordinate system is represented by Σ_M, the *i*-th solid model coordinate system is represented by Σ_{Mi} and the image coordinate systems of the left and right cameras are represented by Σ_{IL} and Σ_{IR}. Σ_W means world coordinate system and Σ_H dose hand coordinate system of the robot.

The position vectors projected in the Σ_{IR} and Σ_{IL} of arbitrary *j*-th point on target object can be described as ${}^{IR}\boldsymbol{r}_j^i$ and ${}^{IL}\boldsymbol{r}_j^i$. They can be calculated as below:

$$ {}^{IR}\boldsymbol{r}_j^i = f_R({}^W\phi^i, {}^{Mi}\boldsymbol{r}_j, \boldsymbol{q}), {}^{IL}\boldsymbol{r}_j^i = f_L({}^W\phi^i, {}^{Mi}\boldsymbol{r}_j, \boldsymbol{q}). \tag{1} $$

These relations connect the defined points, ${}^{Mi}\boldsymbol{r}_j$, on the 3D object and projected points on the left and right images, since the joint angle vector of robot, $\boldsymbol{q}$, can be detected by robot. ${}^{IR}\boldsymbol{r}_j^i$ and ${}^{IL}\boldsymbol{r}_j^i$ can be calculated when the value of ${}^W\phi^i$, which means pose (position and orientation) of the *i*-th model, is given by an assumption in Genetic Algorithm (GA) process.

2.2. Recognition method using genetic algorithm

The 3D marker as a target with color and shape information used in this paper is shown in Fig. 2. The pose of the target is represented by the coordinate systems, Σ_M, which is fixed at the target.

In order to measure the pose of the target object, many solid models having the same shape and color with the real target object but having different poses are prepared. The state that the *i*-th solid model and the real target object exist in 3D space is shown in Fig. 3 (the right side). The pose of the *i*-th solid model is represented as ${}^W\phi^i = [t_x^i, t_y^i, t_z^i, \epsilon_1^i, \epsilon_2^i, \epsilon_3^i]$ ($\epsilon = [\epsilon_1^i, \epsilon_2^i, \epsilon_3^i]$ is quaternion posture parameter) in Σ_W.

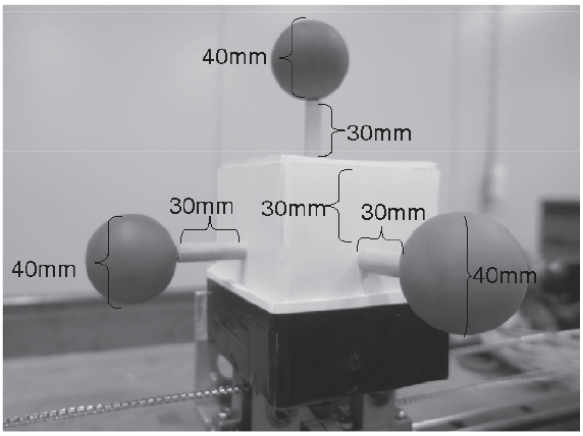


Fig. 2. 3D marker.

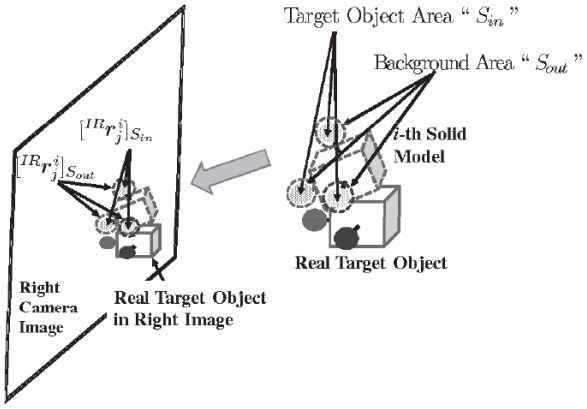


Fig. 3. i-th object model is projected to right image plane whose target object area denoted by "S_{in}" and background area denoted by "S_{out}." Projected points $^{IR}\boldsymbol{r}_j^i$ on "S_{in}" are represented as $[^{IR}\boldsymbol{r}_j^i]_{s_{in}}$ and on S_{out} as $[^{IR}\boldsymbol{r}_j^i]_{s_{out}}$.

The six elements of $^W\phi^i$ are expressed in binary with 12 bits. The shape and color of the solid model and the real target object projected to right camera image are shown in Fig. 3 (the left side).

Then a correlation function representing a matching degree of projected model against the real target object in the image is used as a fitness function in GA process. When the pose of the model and that of the real target object are identical, the fitness function value is to be maximized. We have confirmed that a correlation function having the highest peak with a condition of the pose of real target completely being identical to the model could be designed [6,7]. So the problem of recognition for target object can be changed into an optimization problem that explores the maximum of fitness function and finds the $^W\phi^i$ to give the maximum.

There are many methods for the exploration of maximum of functions. The easiest method is full exploration that calculates function value for all and find the maximum, which takes too much time for real-time recognition. Therefore, in this research GA, which can find the maximum more efficiently in shorter time, is used to solve it [8]. Furthermore, a 1-step GA is proposed for real-time visual control [9].

3. Hand & eye visual servoing controller

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

$$^W\dot{\boldsymbol{r}}_d = \boldsymbol{K}_{P_p}{}^W\boldsymbol{r}_{E,Ed} + \boldsymbol{K}_{V_p}{}^W\dot{\boldsymbol{r}}_{E,Ed}, \tag{2}$$

where $^W\boldsymbol{r}_{E,Ed}, ^W\dot{\boldsymbol{r}}_{E,Ed}$ can be calculated from homogeneous Matrix $^E\boldsymbol{T}_{Ed}$ and the time derivative, $^E\dot{\boldsymbol{T}}_{Ed}$ that connect end effector coordinate Σ_E and desired pose of end effector Σ_{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,

$$^W\boldsymbol{\omega}_d = \boldsymbol{K}_{P_o}{}^W\boldsymbol{R}_E{}^E\Delta\boldsymbol{\epsilon} + \boldsymbol{K}_{V_o}{}^W\boldsymbol{\omega}_{E,Ed}, \tag{3}$$

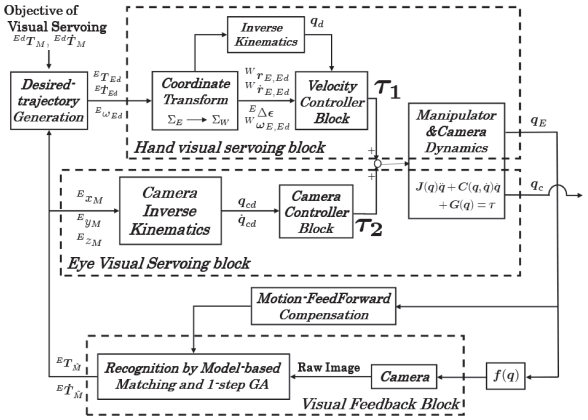


Fig. 4. Block diagram of the hand visual servoing system.

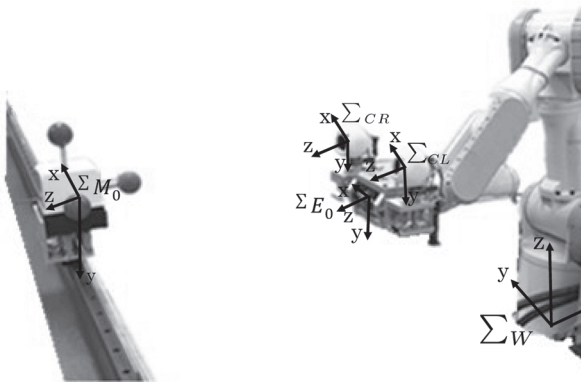


Fig. 5. Experimental visual servoing system.

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 into variables on base coordinates of Σ_W using ${}^{E}\boldsymbol{T}_{Ed}$ and ${}^{E}\dot{\boldsymbol{T}}_{Ed}$. Also, $\boldsymbol{K}_{P_o}$ and $\boldsymbol{K}_{V_o}$ are suitable feedback matrix gains. The desired hand pose is represented as ${}^{W}\boldsymbol{\psi}_d^T = [{}^{W}\boldsymbol{r}_d^T, {}^{W}\boldsymbol{\epsilon}_d^T]^T$.

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

$$\boldsymbol{q}_d = \boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_d) \tag{4}$$

$$\dot{\boldsymbol{q}}_d = \boldsymbol{K}_p(\boldsymbol{q}_{Ed} - \boldsymbol{q}_E) + \boldsymbol{K}_v \boldsymbol{J}_E^{+}(\boldsymbol{q}) \begin{bmatrix} {}^{W}\dot{\boldsymbol{r}}_d \\ {}^{W}\boldsymbol{\omega}_d \end{bmatrix} \tag{5}$$

where $\boldsymbol{K}_p$, $\boldsymbol{K}_v$ is positive definite matrices and $\boldsymbol{f}^{-1}({}^{W}\boldsymbol{\psi}_d^T)$ 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 Mitsubishi PA-10 robot arm is a 7 links manipulator, and the end-effector has 6-DoF, so it has a redundance.

The hardware control system of the PD servo system of the robot PA10 is expressed as

$$\boldsymbol{\tau} = \boldsymbol{K}_{SP}(\boldsymbol{q}_d - \boldsymbol{q}) + \boldsymbol{K}_{SD}(\dot{\boldsymbol{q}}_d - \dot{\boldsymbol{q}}) \tag{6}$$

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

Furthermore, two rotatable cameras whose controller is separated from the hand's are installed at the end-effector, which constitutes eye-vergence system.

4. Experiment of hand & eye-vergence visual servoing

4.1. Experimental conditions

The experimental system is shown in Fig. 5, and the target 3D marker moves in the landscape direction. The initial hand pose is defined as Σ_{E_0} and the initial object pose is defined as Σ_{M_0}. Both are stationary in space as shown in Fig. 5. The relation between the object and the desired end-effector is set as:

$${}^{Ed}\boldsymbol{\psi}_M = [0, -90[\mathrm{mm}], 545[\mathrm{mm}], 0, 0, 0]. \tag{7}$$

Target object moves in landscape direction according to the following time function as:

$${}^{M_0}x_M(t) = -150\cos(\omega t)[\mathrm{mm}], \tag{8}$$

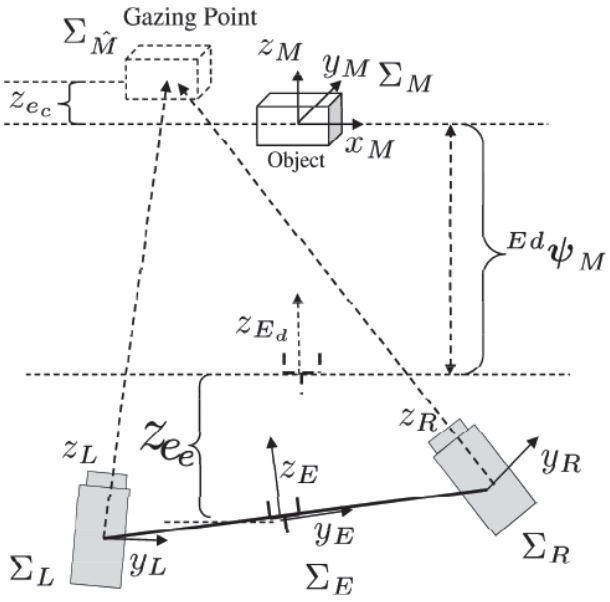


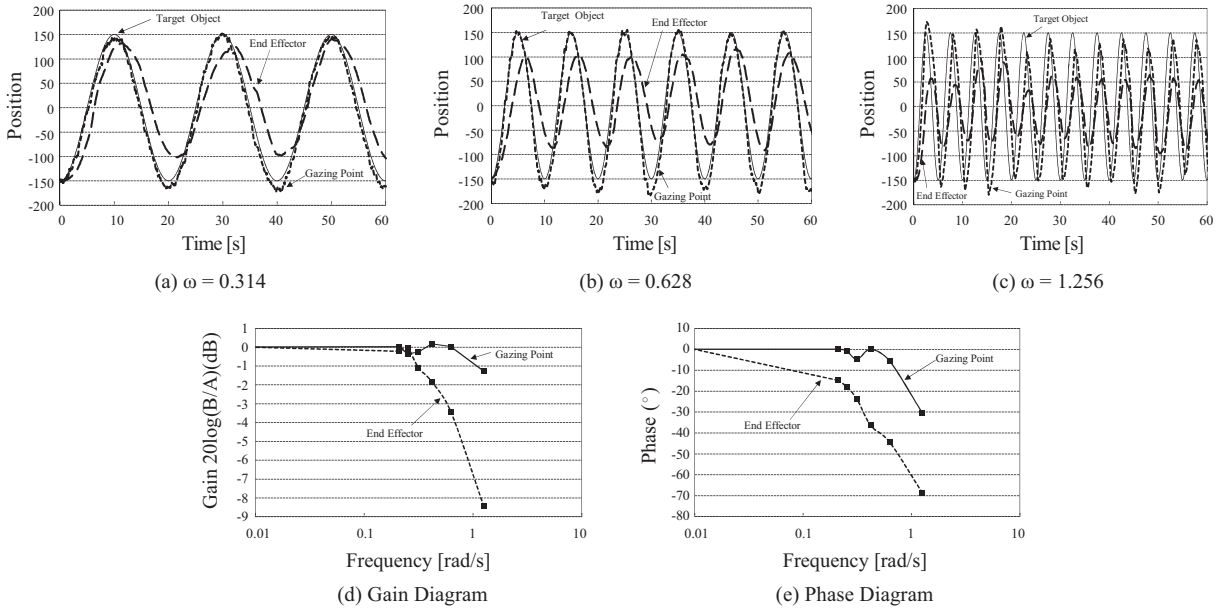
Fig. 6. Cameras' gazing point.

4.2. *Definition of trackability*

Here, to compare the trackability of the eye-vergence system and fixed camera system, we define a concept of gazing point. As it is shown in Fig. 6, the intersection of both cameras' gazing directions is defined as the gazing point of the cameras. Since the tilt angle of both left and right cameras is common by kinematical design, there always exists the intersection of gazing directions, i.e. gazing point in space.

4.3. *Experiment results*

The results of the position and the orientation of the target 3D marker recognized by two rotatable cameras are shown in Fig. 7 (6-dof), where the meaning of the Figs 7(a)–(b) are explained in the figure caption. In the fixed camera system, motion orbits of cameras and end effector are same, so the trackability of the fixed camera system can be reflected by the motion of the end effector. We evaluated eye-vergence system comparing to fixed camera system using gazing point performance that reflects the trackability of eye-vergence system with different angular velocity, ω ($\omega = 0.314$, $\omega = 0.628$, $\omega = 1.256$) as shown in Figs 7(a) (b) (c). We did each experiment for 60 s for each value of ω, and got the average delay time and the amplitude to draw the frequence response curve. When the target object moves slowly, both hand and cameras can keep tracking with it, as shown in Fig. 7(a). While, as the raise of the velocity of the target object, the hand can not keep tracking with it, but cameras can gaze at the object every time, as shown in Figs 7(b) and (c). The amplitude-frequence curve and the delay frequency curve are shown in Figs 7(d) and (e). Here, A means the amplitude of target object shown in Figs (a), (b) and (c), and B means the maximum amplitude of end effector or gazing point. In these two figures the abscissa axes are ω. From Fig. 7(d), the curve of the fixed camera system is always below the curves of the cameras, we can see that the amplitude of the eye-vergence system is more closed to the target object than the fixed camera system. And from Fig. 7(e), the curve of the fixed camera system is also below the curves of the cameras, which means that the delay time in the fixed camera system is longer than in the eye-vergence system.



(a) ω = 0.314 (b) ω = 0.628 (c) ω = 1.256

(d) Gain Diagram (e) Phase Diagram

Fig. 7. The object moves in lateral direction and it's position and orientation are estimated through "1-step GA", then they are feedbacked to the visual servoing controller as shown Visual Feedback Block in Fig. 4. Then the frequency response results in this figure include robot's dynamics and pose-tracking dynamics. In this figure, Target Object: The motion of target object to the referred by robot hand through visual servoing. End Effector: Trajectory tracking result by fixed camera visual servoing. Gazing Point: Trajectory tracking result of gazing point by eye-vergence visual servoing. (a) represents frequency response results of target object with $\omega = 0.314$, which is tracked by end effector and eye-vergence gazing point (defined by Fig. 6), (b) $\omega = 0.628$, (c) $\omega = 1.256$, (d) gain curve of Bode diagram, (e) phase curve of Bode diagram.

5. Conclusion

In this paper, we have carried out visual servoing experiments on frequency responses to evaluate the observability and trackability of hand & eye-vergence visual servoing with 1-step GA. Based on the experimental results, we have analyzed the trackability, amplitude-frequency and phase-frequency responses of the cameras of the eye-vergence system under moving object with different angular velocity. The trackability and observability of the eye-vergence system has been confirmed to function and the eye-vergence system's ability to track a target in the lateral direction has been clarified through the frequency response experiments.

References

[1] S. Hutchinson, G. Hager and P. Corke, A Tutorial on Visual Servo Control, *IEEE Trans. on Robotics and Automation* **12** (1996), 651–670.
[2] P.Y. Oh and P.K. Allen, Visual Servoing by Partitioning Degrees of Freedom, *IEEE Trans. on Robotics and Automation* **17** (2001), 1–17.
[3] E. Malis, F. Chaumentte and S. Boudet, 2-1/2-D Visual Servoing, *IEEE Trans. on Robotics and Automation* **15** (1999), 238–250.
[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* **9** (1993), 152–165.
[5] W. Song, M. Minami, F. Yu, Y. Zhang and A. Yanou, 3-D Hand & Eye-Vergence Approaching Visual Servoing with Lyapunouv-Stable Pose Tracking, *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Shanghai, (2011), pp. 11.

[6] K. Maeda, M. Minami, A. Yanou, H. Matsumoto, F. Yu and S. Hou, Frequency Response Experiments of 3-D Full Tracking Visual Servoing with Eye-Vergence Hand-Eye Robot System, *SICE Annual Conference*, Akita, (2012), pp. 101–107.
[7] W. Song, Y. Fujia and M. Minami, 3D Visual Servoing by Feedforward Evolutionary Recognition, *Journal of Advanced Mechanical Design, Systems, and Manufacturing (JSME)* **4** (2010), 739–755.
[8] T. Nagata, K. Konishi and H. Zha, Cooperative manipulations based on Genetic Algorithms using contact information, *Proceedings of the International Conference on Intelligent Robots and Systems*, Pittsburgh, (1995), pp. 400–5.
[9] M. Minami and 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* **21** (2009), 709–719.
[10] W. Song, M. Minami and 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* **2** (2007), 127–139.