The Twenty-Third International Symposium on Artificial Life and Robotics 2018 (AROB 23rd 2018),
The Third International Symposium on BioComplexity 2018 (ISBC 3rd 2018),
B-Con Plaza, Beppu, Japan, January 18-20, 2018

Verification of Illumination Tolerance for Human Recognition Moblie Robot based on Visual Servoing

Zeyi Zhang[1], Xiang Li[1], Mamoru Minami[1], Takayuki Matsuno[1]

[1]Okayama University, Japan
(Tel: 81-86-251-8233, Fax: 81-86-251-8233)
[1]peiw4znm@s.okayama-u.ac.jp

Abstract: Nowadays, automated guided vehicle (AGV) has been widely used in production field for handling work. In some cases, such as equipment maintenance, AGV needs to track the staff in real-time. To meet the above demand, we propose a real-time human recognition and tracking system for AGV based on visual servoing. Using dual-eyes camera, the proposed system can distinguish human from background by color, estimate relative position and control AGV to achieve human tracking in real-time. We employ Genetic Algorithm (GA) and a RGB image termed here as raw-image to execute recognition process by a model-based matching method. As the main contribution for this paper, we have developed visual servoing using adaptive system for various natural light conditions including backlight based on HSV model. Experiment using a mobile robot with the proposed system were condected in a complex light environment. We regard a worker in blue uniform as the recognition target. The experimental results confirmed that the proposed system is able to provide high homing accuracy and robustness against disturbances that the influence from not only the captured camera images under different lighting conditions but also the movement of the robot.

Keywords: Human-Recognition, Genetic Algorithm, Model-Based-Matching

1 INTRODUCTION

Nowadays, automated guided vehicle (AGV) has been widely used in storehouse or facility manufacturing for moving materials or goods. With the development of technology, the navigation of AGV has been changing from using guide tape to laser sensor. Nevertheless, the pre-laying of guide tape or expensive sensors limits the further spread of AGV. Moreover, in some cases, such as equipment maintenance and indoor sales, AGV needs to track human in real-time. With this motivation, we propose a real-time human recognition and tracking system for AGV based on visual servoing using dual-camera.

Many researchers [1]-[4] applied a binary image and some edge extraction methods that need several preprocessing steps to recognize the images. Those various filtering stages seem to be a time-consuming process, and to us, these are not convenient for real-time recognition. The tradition image recognition of human [5] often uses extraction technology of many characteristics for shape or position relation of eye, mouth and nose. However, there are some problems to extract a lot of characteristics because of non-clear shadow, which makes image recognition of human difficult. To solve these problems, we use directly the unprocessed RGB image termed here as raw-image. Basically, this research is based on a model-based matching method. We employ a model for the recognition purposes of a human considered here as the target. An fitness evaluate function, whose computation is based on the configuration of the human model, is used to evaluate the extent to which the human recognition model matches with the object being imaged, by changing the recognition problem into an optimization problem. Therefore, We use a Genetic Algorithm (GA) in the image recognition, because of its high performance of optimization.

GA is well known as a method for solving search and parameter optimization problems [6]. Moreover, to use the GA process in real time, i.e., to extract its position from the consecutively input images, in the past research we used the GA such a way that every input image is evaluated only one time by target-model-based fitness function, which we named Step GA [7]. To perform in real time, we used GA with long history and modified as real-time multi-step GA for searching the best model.

As the main contribution for this paper, we have constructed a human recogniton mobile robot system and enhance the system's ability to adapt to a changing light environment. In order to verify the effectiveness of the proposed system, human recognition and tracking experiments using a mobile robot with dual-eye camera were condected in a complex light environment.

2 HUMAN RECOGNITION MOBILE ROBOT

The human recognition mobile robot consists of 2 parts as shown in Fig.1. Two cameras (FCB-IX-11-A) installed at the top of robot are used as visual sensor for human recognition. In Movement Section, a two-wheel drive bogie is controlled to move the robot along desired path. The bogie can move

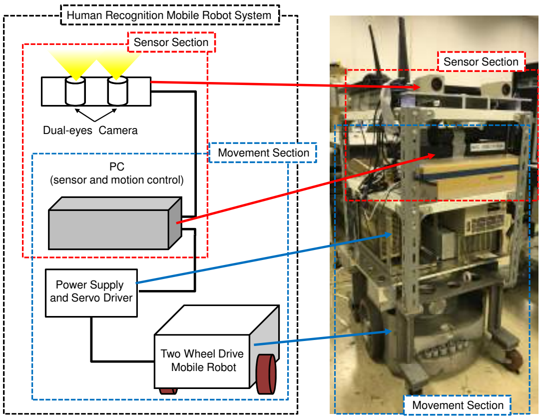

Fig. 1. Human Recognition Mobile Robot System

up to 0.6 [m/s].

3 HUMAN RECOGNITION ALGORITHM

In order to recognize human accurately, it is an effective method that distinguishing the color from the human to background. Regarding human as the objective of recognition, the effectiveness of recognition can be improved while adding extraction of color of the cloth and hair which are human characteristics to the evaluation function. In this paper, we assume that the recognition target is wearing dark blue clothes whose hair color is black.

3.1 Human Recognition Model

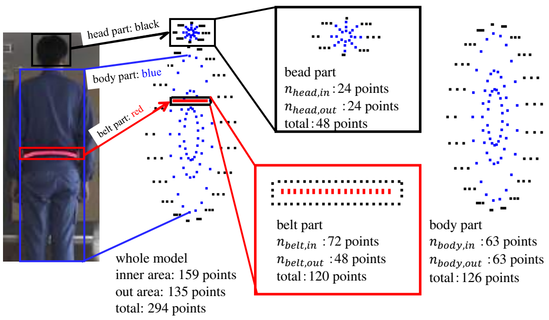

Fig. 2. Human recognition model

As it is shown by Fig.2, the recognition model can be divided into head part, body part and belt part. Each of them is consists of inner area and outer area. We added the bright belt to the model because in close range the camera may cannot capture the whole model, moreover, a bright belt can improve system accuracy in the dark. In this paper we applied 3D model-based recognition approach and multi-step GA to get human position.

3.2 HSV Color Depending Recognition System

The sample image and RGB value of blue cloth are shown as Fig.4. The 1-6 are sample about the blue cloth under different situations such as outdoors, indoor, illumination is

Fig. 3. The camera may cannot capture the whole model in close range

high or low etc. From the value of RGB, it is understood that there is great difference among the RGB value at the various situation even the same cloth.

Sample	1	2	3	4	5	6
R	61	113	48	30	41	74
G	81	127	49	28	52	80
B	161	168	69	39	115	115
H	228	225	237	251	231	231
S	62	33	31	28	64	36
V	63	66	27	15	45	45

Fig. 4. RGB and HSV

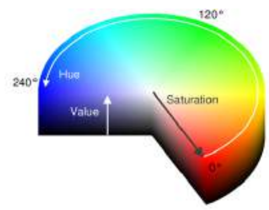

Fig. 5. Hue circle

So that, the HSV color system is considered (Here, the conversion expressions from RGB value to HSV value are omitted). H[0-359] of HSV represents the hue, which defines the direction of 5:00 at the hue circle shown in Fig.5 as 0, and the value becomes big counterclockwise. S[0-100] represents color, which is expressed by the radius of hue circle. V[0-255] is luminosity, and is the length of axis vertical to the hue circle. Color becomes dark with the value of V decreasing. Here, the value of H shown in Fig.3 should be noticed. Though the range of H is 0-359, the color of blue clothes can be specified by limiting the value of H to 220-260 without relation to the change of illumination and race.

3.3 Fitness Evaluate Function

In order to distinguish human from background by comparing hue value, it is necessary to construct a set of evaluation system about the matching degree between the human recognition model and the actual image. In this paper, we evaluate the combination of the raw-image and the recognition model by fitness function $F_{all}(\boldsymbol{\psi})(\boldsymbol{\psi} = x, z)$ which is given as Eq.(1).

$$
\begin{aligned}
F_{all}(\boldsymbol{\psi}) &= 0.5\Big(\big(\sum_{i=1}^{n_{body,in}} p_{body,in}({}^{IR}\boldsymbol{r}_i) + \sum_{i=1}^{n_{body,out}} p_{body,out}({}^{IR}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{body,in} + n_{body,out}) \\
&+ \quad \big(\sum_{i=1}^{n_{body,in}} p_{body,in}({}^{IL}\boldsymbol{r}_i) + \sum_{i=1}^{n_{body,out}} p_{body,out}({}^{IL}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{body,in} + n_{body,out})\Big) \\
&+ \quad 0.5\Big(\big(\sum_{i=1}^{n_{head,in}} p_{head,in}({}^{IR}\boldsymbol{r}_i) + \sum_{i=1}^{n_{head,out}} p_{head,out}({}^{IR}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{head,in} + n_{head,out}) \\
&+ \quad \big(\sum_{i=1}^{n_{head,in}} p_{head,in}({}^{IL}\boldsymbol{r}_i) + \sum_{i=1}^{n_{head,out}} p_{head,out}({}^{IL}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{head,in} + n_{head,out})\Big) \\
&+ \quad 1.0\Big(\big(\sum_{i=1}^{n_{belt,in}} p_{belt,in}({}^{IR}\boldsymbol{r}_i) + \sum_{i=1}^{n_{belt,out}} p_{belt,out}({}^{IR}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{belt,in} + n_{belt,out}) \\
&+ \quad \big(\sum_{i=1}^{n_{belt,in}} p_{belt,in}({}^{IL}\boldsymbol{r}_i) + \sum_{i=1}^{n_{belt,out}} p_{belt,out}({}^{IL}\boldsymbol{r}_i)\big) \\
&/ \quad (n_{belt,in} + n_{belt,out})\Big) \\
&= \quad 0.5(F_{body,R}(\boldsymbol{\psi}) + F_{body,L}(\boldsymbol{\psi})) \\
&+ \quad 0.5(F_{head,R}(\boldsymbol{\psi}) + F_{head,L}(\boldsymbol{\psi})) \\
&+ \quad 1.0(F_{belt,R}(\boldsymbol{\psi}) + F_{belt,L}(\boldsymbol{\psi}))
\end{aligned} \tag{1}
$$

In Eq.(1),$p_{body,in}({}^{IL}\boldsymbol{r}_i)$,$p_{body,in}({}^{IR}\boldsymbol{r}_i)$,$p_{body,out}({}^{IL}\boldsymbol{r}_i)$, $p_{body,out}({}^{IR}\boldsymbol{r}_i)$ means the matching degree of the raw-image and each point of the body part of the model. The value of $p_{body,in}(\boldsymbol{r}_i)$ and $p_{body,out}(\boldsymbol{r}_i)$ are calculated by Eq.(2)and Eq.(3).

$$
p_{body,in}(\boldsymbol{r}_i) = \begin{cases} 1 & (220 \leq hue_u \leq 260), \\ -1 & (220 \geq hue_u, or\ hue \geq 260). \end{cases} \tag{2}
$$

$$
p_{body,out}(\boldsymbol{r}_i) = \begin{cases} -1 & (220 \leq hue_u \leq 260), \\ 1 & (220 \geq hue_u, or\ hue \geq 260). \end{cases} \tag{3}
$$

In Eq.(2) and Eq.(3), $p_{body,in}(\boldsymbol{r}_i)$**,** $p_{body,out}(\boldsymbol{r}_i)$ mean the fitness of each surface points and belt points in body part. hue_u means the hue value of the point represented by HSV model. It can be known that the fitness function will takes the maximum value at the exact position of the model. The fitness of the points in head part and belt part are calculated in similar method.

3.4 Model-based matching using GA

In this reaearch, we used Model-based Matching to calculate the human 3D -Position in searching area from the 2D-image taken by dual-eyes cameras. Apart from other recognition methods based on 2D to 3D reconstruction, the proposed 3D model-based recognition is based on 3D to 2D projection. Moreover, the concept based on the group of pixels rather than individual pixels highlights merits of model-based method over feature-based ones. Fig.6 showed the

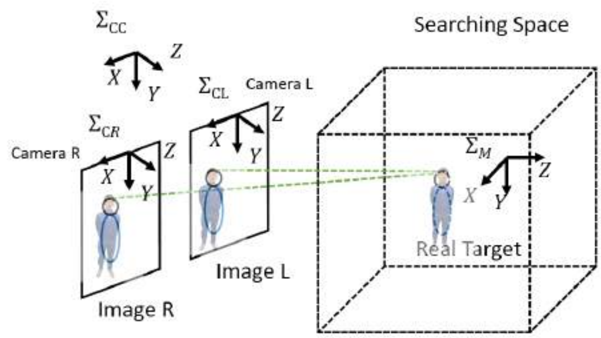

Fig. 6. Model-based Matching System using Dual-eyes Vision System

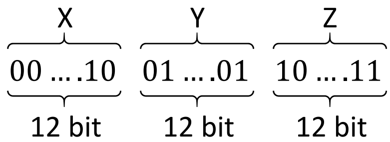

Fig. 7. Gene Representation for Position

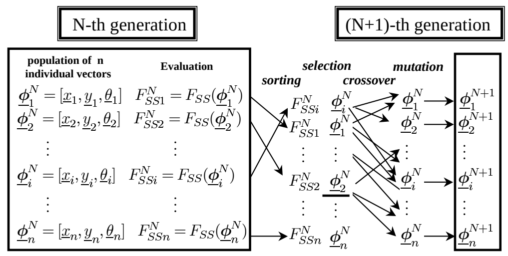

Fig. 8. Evolution process of GA

Model-based matching system using dual-eyes cameras. Target object is the human recognition model that be introduced in last chapter. Knowing the information of the target and predefined relative pose to the dual-eyes cameras, the solid model of the target is predefined and projected to 2D images. Comparing the projected solid model image with the captured 2D images by dual-eyes cameras, the relative pose difference is calculated.

Genes representative to 2 position parameters as shown in Fig.7 are initiated randomly. The gene that has the highest fitness function value represents the position of the real target.Therefore the problem of position recognition addresses to the searching problem. The solution is GA with promising speed and accuracy of performance. According to the performance in time-domain, GA is selected in this work even though there are advanced optimized techniques. The effectiveness of 1-step GA was confirmed in robots especially manipulators and reported in previous work [11]-[13]. Through the steps of GA (Selection, Cross over and Mutation), a number of genes that represent different poses are evaluated by the defined fitness function to get the best gene with the most truthful estimated pose. A evalute function representing a

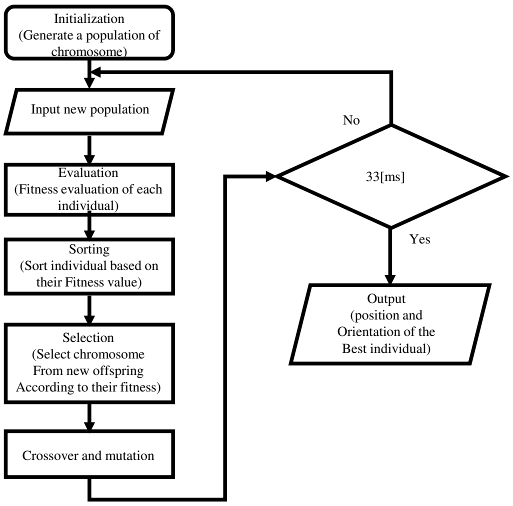

Fig. 9. Flow chart of GA

matching degree of human recogintion model against the real target in the image, which is a correction function of real target projected in camera images with the assumed model represented by poses in genes, is used as a fitness function in GA process. The convergence of GA is realized in the sequences of dynamic images input by video rate [30 fps]. Detail discussion about 1-Step GA and fitness function are explained in [11].

3.5 Real-time position tracking using "multi-step GA" method

For real-time visual control purpose, GA has been employed in a way denoted as " 1-Step GA " evolution. The used cameras 'frame rate is 30 fps. That means every 33 ms 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). Accuracy has also been improved. Now it is renamed as " Multi-Step GA ".

4 RECOGNITION EXPERIMENT

4.1 Human Recognition Full Search

In order to confirm the effective of GA, we performed the full search experiment in frontlighting and backlighting conditions as shown in Figs.10 and 11. The full search is a method to evaluate the result of multi-step GA by analysing the specific image where multi-step GA get the corresponding fitness value. The main idea of full search is to calculate the fitness of every points which are 1[mm] apart in the entire searching area.

The results of the full search show, in both light environments, that there is a significant peak more than 0.8, which is markedly distinct from the environment, can be observed. In this experiment, GA recognition accuracy needed to have the value of 0.7 or more for good performance. So according to the experiment, it can be known that GA recognition is effective no matter in backlight or forward light condition.

4.2 Accuracy examination of position

Accuracy examination of position was conducted under backlight and forward light condition. We selected seven specific positions as measuring points within the scope of effective recognition (1.5[m]-4.0[m]), and then use GA method for distance measurement. Take the average of the 10 measurements as the final result. As shown in Table.I, in both backlight and forward light conditions, the error between the measured and actual values is controlled within 0.2[m]. In addition, the recognitive system has higher accuracy in the range of 2[m] to 3[m]. We think this is because when the actual distance is less than 2[m], the camera cannot capture the head so that the error of GA be increased, and when the actual distance is more than 3[m], the region of the human in image is too small to let the system compute the actual distance accurately using model-based maching method.

Table 1. Distance recognition

Position[m]	Backlight Error[m]	Forward light Error[m]
1.5	-0.14	-0.14
2.0	-0.04	-0.01
2.5	-0.02	-0.05
3.0	-0.05	0.03
3.5	0.05	0.14
4.0	-0.12	-0.09

5 HUMAN TRACKING EXPERIMENT

In order to confirm effectiveness of the recognition and tracking performance of the robot, we conducted human tracking experiment in complex light environment.

5.1 Experiment Condition

As shown by the red arrow in Fig. 11, the route of the tracking experiment can be divided into nine parts. Among them, ①, ②, ⑧, ⑨are performed in the state of backlight, ④, ⑤, ⑥ are in the state of forward light.

5.2 Experiment Results

The situation during the experiment is shown in Fig.13. The left side of Fig.13 shows the screen of the recording video camera, and the right side shows the image from the left and right camera of the mobile robot visual sensor. Experimental results graphs are shown in Figs.14 and 15.

Fig.14 shows the moving direction position x_c and the commanded speed v_d during the experiment. A large recognition error can be seen around 85 s and 140 s. The cause may be the influence of the ground on the running of the

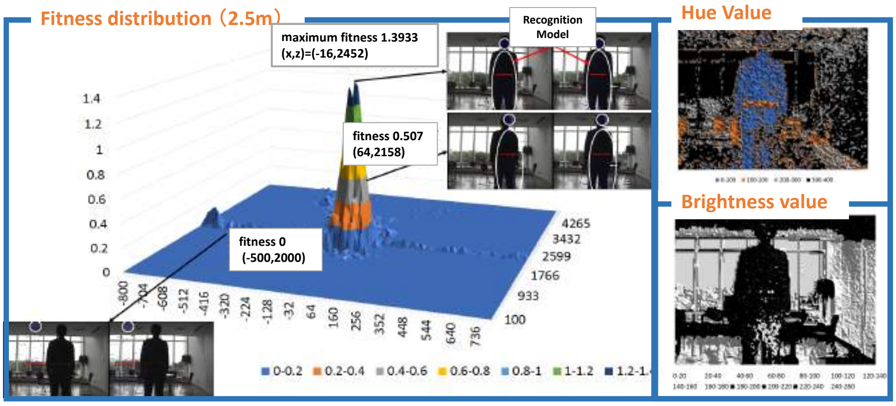

Fig. 10. Fitness distribution under backlight condition

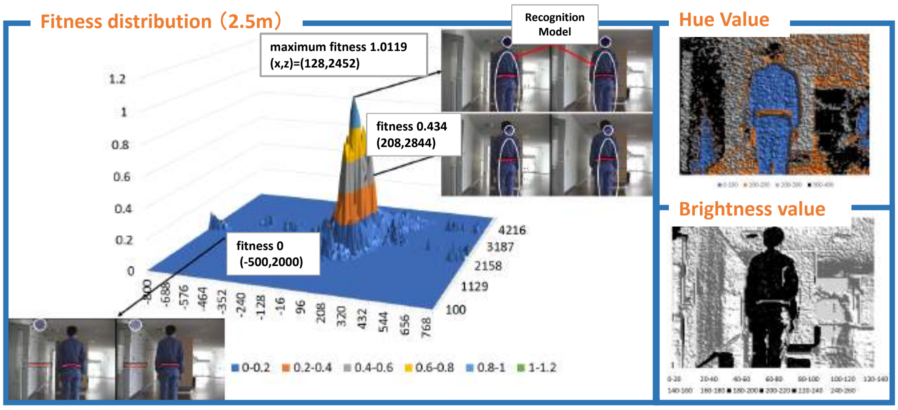

Fig. 11. Fitness distribution under forward light condition

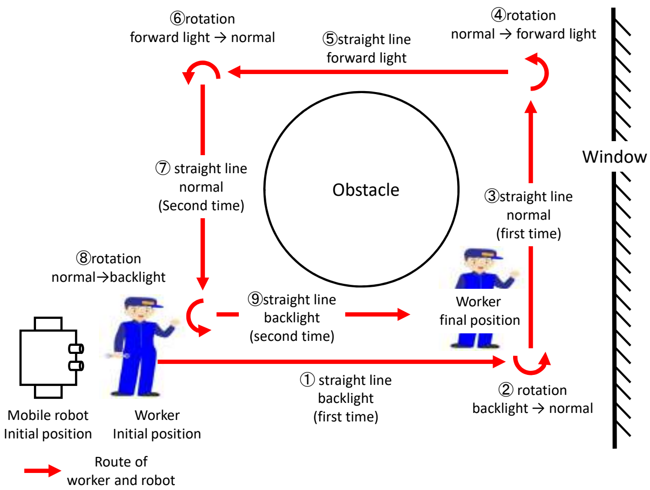

Fig. 12. The route of the tracking experiment

mobile robot (idling of the wheels) and the influence of the backlight environment (the color disappears on the camera screen). Fig.15 shows the deviation Δy in the y-axis direction and the rotation speed ω_d during the experiment. It can be seen that the movement of the mobile robot is greatly affected by rough terrain around 100 s.

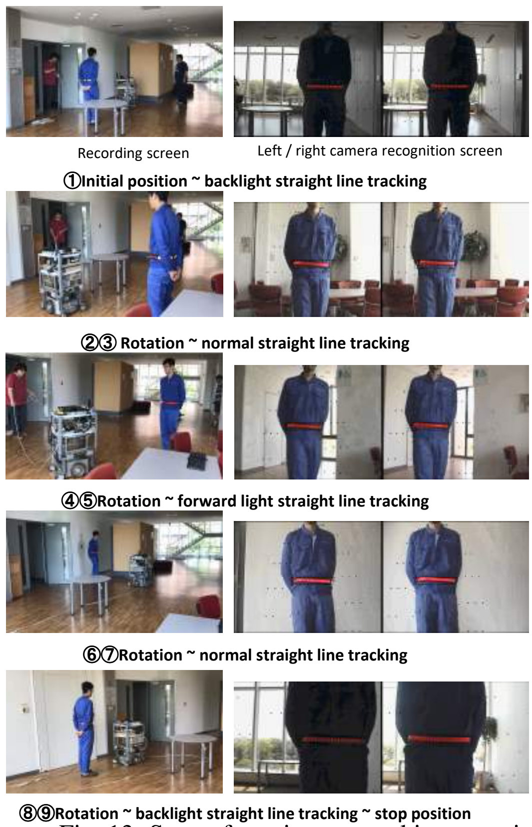

Fig. 13. State of continuous tracking experiment

6 CONCLUSION

In this study, a robust real-time human recognition system by using dual-eyes cameras has been presented. In order to

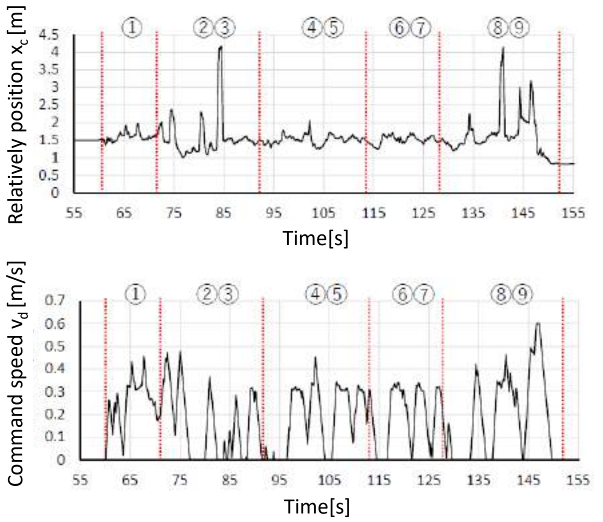

Fig. 14. During experiments, the moving direction position x_c and the commanded speed v_d

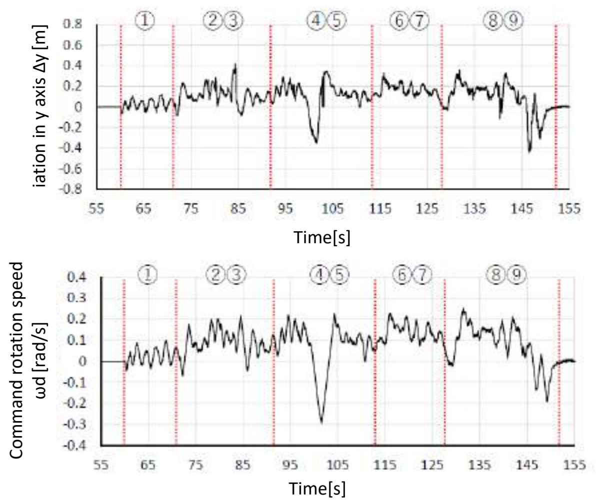

Fig. 15. During the experiments, the deviation Δy in the y-axis direction and the rotation speed ω_d

perform the recognition of a target human adapting to various lighting environment including backlight condition, we propose a multi-step GA, which can gradually take the position of human on basis of HSV color value. Results of the experiment to track a target human whose position in different lighting environment, have shown the robustness of proposed GA search method.

As the next experiment, we plan to further verify the effectiveness about the method suggested above by the recognition experiment of more poeple. Furthermore, we will try to enhance the accurancy of recognition when the distance is less than 1.5[m] by consummating the model and fitness evaluate function.

REFERENCES

[1] Brady, M., Hollerbach, J. M., Johnson, T. L., Lozano-Perez, T. and Mason, M. T., Robot motion: planning and control (1982), pp. 51-71, The MIT Press.

[2] Hooker, W.W., Margulies, G., The Dynamical Attitude Equations for an n-body Satellite, Journal of Astronautical Sciences, Vol. 12 (1965), pp. 123-128.

[3] Stepanenko, Y. and Vukobratovic, M., Dynamics of Articulated Open- chain Active Mechanisms ', Mathematical Biosciences, Vol. 28, Iss. 1-2 (1976), pp. 137-170.

[4] Walker, M. W. and Orin, D. E., Effcient Dynamic Computer Simulation of Robotic Mechanisms, Transactions of ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 104, No. 3 (1982), pp. 205-211.

[5] Huang, Y. F. and Lee, C. S. G., Generalization of Newton-Euler Formulation of Dynamic Equations to Nonrigid Manipulators, Transactions of ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 110, No. 3 (1988), pp. 308-315.

[6] D.E.Goldberg, " Genetic algorithm in Search, Optimization and Machine Learning. Reading ", Addison Wealey, 1989

[7] Anitescu M. and Potra A. F., A time-stepping method for stiff multibody dynamics with contact and friction, International Journal for Numerical Methods in Engineering, Vol. 55 (2002), pp. 753-784.

[8] Myo Myint, Kenta Yonemori, Akira Yanou, Khin Nwe Lwin, Mamoru Minami and Shintaro Ishiyama," Visual-based Deep Sea Docking Simulation of Underwater Vehicle Using Dual-eyes Cameras with Lighting Adaptation ", MTS/IEEE OCEANS ,Shanghai International Convention Center, April 10-13, 2016

[9] M. Minami, J. Zhu, M. Miura, Real-time Evolutionary Recognition of Human with Adaptation to Environmental Condition, The Second International Conference on Computational Intelligence, Robotics and Autonomous Systems (CIRAS), Proceeding CD-ROM PS010103 (2003.12)

[10] Mamoru Minami, Fujia Yu and Akira Yanou: " Lyapunov-stable Position/Force Control Based on Dual Nature in Constraint Motion " 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2013), pp.2894-2901, 2013.

[11] Yu F., Minami M., Song W., Zhu J. and Yanou A., On-line head pose estimation with binocular hand-eye robot based on evolutionary model- based matching, Journal of Computer and Information Technology, Vol.2, No.1, pp.43-54, 2012.

[12] Song W. and Minami M., 3-D Visual Servoing Using Feedforward Evolutionary Recognition, Journal of the Robot Society of Japan, Vol.28, No.5, pp.591-598 (in Japanese), 2010.

[13] Yu Cui, Kenta Nishimura, Yusuke Sunami, Mamoru Minami, Takayuki Matsuno and Akira Yanou, Analyses about Trackability of Hand-eye-vergence Visual Servoing in Lateral Direction, OPTIROB Conference, Romania, 2015.