Recognition of Cloth Target Based on Photograph Modeling and Handling

○学　Khaing Win Phyu (岡山大)　学　崔　禹（岡山大）
学　田　宏志（岡山大）　学　萩原　涼太（岡山大）
学　舟久保　龍希　（岡山大）　正　矢納　陽（岡山大）
正　見浪　護　（岡山大）

Khaing Win Phyu, Okayama University, p6nn3hpz@s.okayama-u.ac.jp
Yu Cui, Okayama University
Hongzhi Tian, Okayama University
Ryota Hagiwara, Okayama University
Ryuki Funakubo, Okayama University
Akira Yanou, Okayama University
Mamoru Minami, Okayama University

Human can recognize easily the different kinds of object. However, it is difficult for robots recognize the objects such as cloth, string, moving objects and etc. This paper is to develop the cloth recognition scheme for visual servoing to apply it in industries where workers recognize clothes and package them to send internet users who have purchased. In this study, the 7-link manipulator equips with two cameras that detect position and orientation (pose) of cloth using 3D-MoS (Move on Sensing). The proposed system introduces a new model generating method for model-based cloth recognition that utilizes a photograph of cloth as a model to recognize the cloth, which enable that the robotic system does not need to prepare a definition of the target cloth as a preparation of the recognition. It has been confirmed that cloth can be recognized by the system through 1000 times recognition experiment.

Key Words: Visual Servoung, Genetic Algorithm, Move on Sensing

1 Introduction

In this system, three cameras as vision sensors, one PA-10 robot and two personal computers are used. The first camera, that is fixed in the workspace for capturing photos, save the photo with the BMP (bitmap file) and then it generates the model in the sensor PC. The other two cameras set up at the end-effector of a PA-10 robot are for making the recognition based on the model. Finally, the robot PC controls the PA-10 controller to pick up the cloth after recognition successfully and set the cloth into the box for which human send to the client. The advantages of this research system is to reduce the cost of human operators. Moreover, the cameras, robot, PC and robot controller make more effective time, specific location and saving cost than staffs. Fig. 1 shows the configuration of the system. In addition, the light environment effects on not only recognition but also the handling performance. In this paper, we did experiment under two light conditions and recognition are calculated by error estimation and also these two results are demonstrated by histogram. One condition is calculated from the normal lighting condition in the factory. Another one is adding some lighting effect using a fluorescent table lamp with 1000 Lx. The experiment confirmed that the system can handle for both two conditions. Moreover, we can get the best performance of handling in which circumstance by comparing two light environments. This research emphasis on various light condition for handling of different kinds of cloths by the robot with high performance.

2 Recognition

2.1 Model Generation Process

This section explains how to generate such a model relationship between background and the target object. Fig.2 shows the process of model generation that represents the best matching condition with the black points group on the target object.

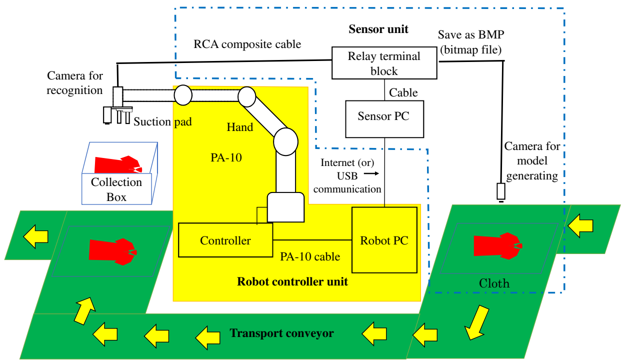


Fig.1 System Configuration

2.2 Fitness Function

Fitness distribution is decided by hue value. The fitness function value is high in the care when the system recognizes the object clearly. We collect the maximum fitness valuate and weed out the minimum value by finding the average as the following equations;

$$F_L = \frac{\sum_{i=0}^{num_{in}} I_{L_{in}}(x_i, y_i) + \sum_{j=0}^{num_{out}} I_{L_{out}}(x_j, y_j)}{num_{in} + num_{out}} \quad (1)$$

Fig.2 Model Generation Process

$$F_R = \frac{\sum_{i=0}^{num_{in}} I_{R_{in}}(x_i, y_i) + \sum_{j=0}^{num_{out}} I_{R_{out}}(x_j, y_j)}{num_{in} + num_{out}} \tag{2}$$

$$F = \frac{F_L + F_R}{2} \tag{3}$$

2.3 Projection

The relationship between the world coordinate system of the manipulator and the hand coordinate system is shown in Fig.3. It can also be named as the coordinate system of 3D-MoS (Move on Sensing) robot. 3D-MoS uses forward kinematic for automatic handling of clothes. Fig. 4 shows the coordinate of the dual-eye vision system, where Σ_M is the target object's coordinate system. The left and right input images from the two cameras are directly matched by the left and right searching models, which are projected from 3D model onto 2D image plane. Σ_{CR} and Σ_{CL} are the coordinate systems of the left and right cameras, and Σ_{IR} and Σ_{IL} are the coordinate systems of the left and right camera's images. In Fig. 4, two cameras are fixed on the robotic hand that it represents as desired end-effector coordinate (Σ_E). Using the perspective projection as projection, transformation matrix is shown in equation(4).

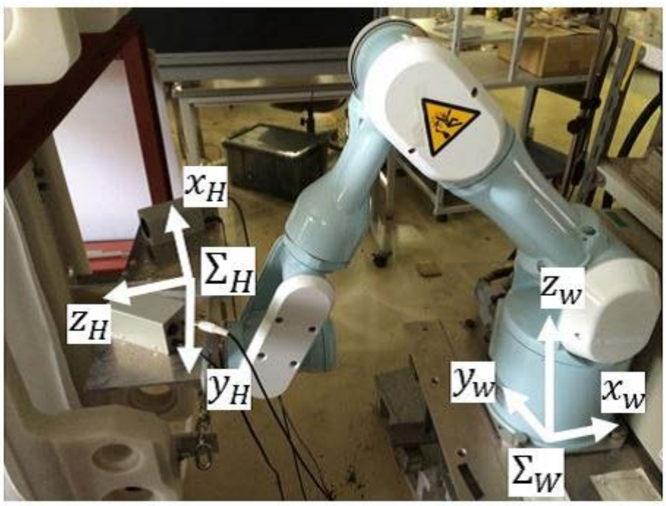


Fig.3 Coordinate System of 3D-MoS Robot

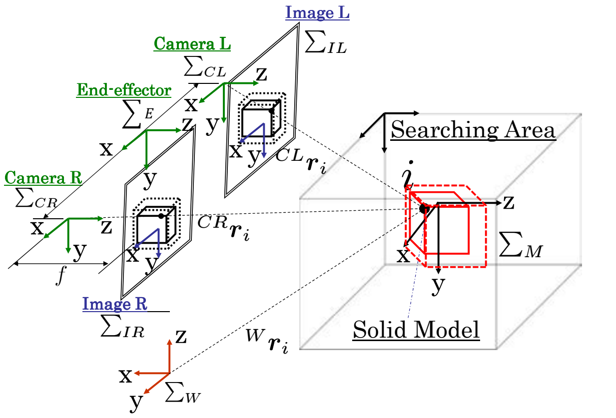


Fig.4 Coordinate System of dual-eyes

$$P = \frac{1}{^{C}z_i}\begin{bmatrix} \frac{f}{\eta_x} & 0 & ^{I}x_0 & 0 \\ 0 & \frac{f}{\eta_y} & ^{I}y_0 & 0 \end{bmatrix} \tag{4}$$

2.4 Genetic Algorithm (GA)

Genetic algorithm (GA) includes two portions. They are individual GA and GA evolution. The individual GA has the same color and shape with the target object but different position and orientation. It has 72 bits and characteristic as shown in the following;

$$\overbrace{\underbrace{011000100111}_{12bits}}^{t_x}\overbrace{\underbrace{000011000111}_{12bits}}^{t_y}\overbrace{\underbrace{001100111101}_{12bits}}^{t_z}$$

$$\overbrace{\underbrace{110101001001}_{12bits}}^{\epsilon_1}\overbrace{\underbrace{000101111001}_{12bits}}^{\epsilon_2}\overbrace{\underbrace{001101111001}_{12bits}}^{\epsilon_3}$$

GA evolution process can get the best matching between the target object and model quickly and accurately. Fig. 5 shows the improvement of evolution process in each generation.

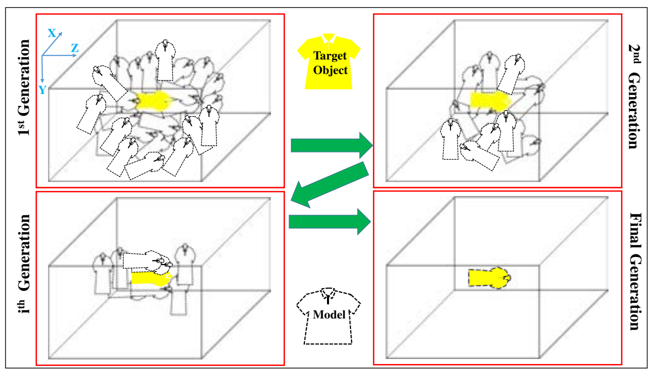


Fig.5 GA Evolution Process

3 Experiment of handling

There are different kinds of gripper such as mechanical, magnetic, vacuum and so on, which are divided by depending on the grasp application of the objects. Most of end-effector types are mechanical grippers. However, in this experiment, the PA-10 robot attached 4-section cups (vacuum type) that allow it to perform pick and place application. According to the ability of handling object that are the smooth, flat surface (cloth package), vacuum type gripper is selected. Experiments under two light conditions (700[Lx], 1000[x] are conducted and analyzed. Table 1 and 2 show the average error and standard deviation. It can be seen from table 2 (100 times handling experiment) that handling error is increased under 1000[Lx] comparing to the error under 700[Lx]. The increasing error at approximately 1000[Lx] become worse performance than general light environment at approximately 700 [Lx]. Not only in comparison of the table but also in each position error and angle error of the histogram, it can be seen that the changing data approaches the bad one. The histogram shows how error has taken place. With respect to the comparison of two histogram, the number of frequencies (times) caused by position error (position x [mm] in 100 times handling experiments between normal optical environment) and approximately 1000[Lx]. In Fig.6 (a), there are 40 times frequency from zero position error, over 5 times frequency from 0.5 position error, same time frequency of 2 and 2.5 position error. On the other hand, as Fig. 6 (b), there are 30 times frequency from zero position error which can be understood clearly that zero position error decrease 10 times less than Fig. 6 (a). Moreover, the other position error of frequency times was increased and appeared which the position error lack in Fig. 6 (a). Also, we can see it in the position error (position y [mm]) and angle error (angle θ [deg] from Fig. 7 (a), (b) and Fig. 8 (a), (b). There have been confirmed and evaluated by the above reasons which the normal optical environment is better than approximately 1000 [Lx]. In factory with general light condition (normal optical environment), this experiment has established to recognize and handle the clothes with less error in every pose and model.

Table 1 Average error and standard deviation

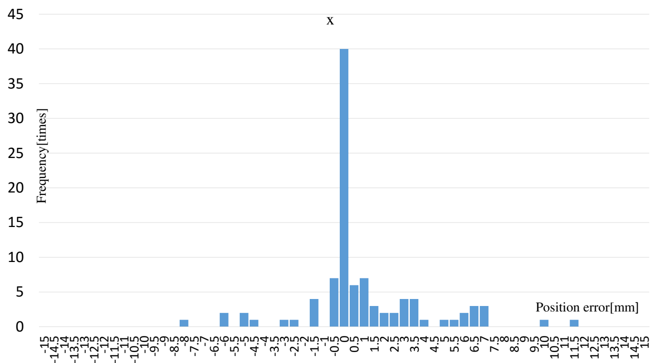
	x[mm]	y[mm]	z[mm]	θ[deg]
Average error	0.73242	0.666015	12.971025	-0.044037
Standard deviation (σ)	2.997442209	2.305696352	5.69075513	3.746834401

	x[mm]	y[mm]	z[mm]	θ[deg]
Average error (-3σ)	-8.259906627	-6.251074055	-4.10115154	-11.2845402
Average error (-2σ)	-5.262464418	-3.945377703	1.589573974	-7.537705801
Average error (-1σ)	-2.265022209	-1.639681352	7.280299487	-3.790871401
Average error	0.73242	0.666015	12.971025	-0.044037
Average error (+1σ)	3.729862209	2.971711352	18.66175051	3.702797401
Average error (+2σ)	6.727304418	5.277407703	24.35247603	7.449631801
Average error (+3σ)	9.724746627	7.583104055	30.04320154	11.1964662

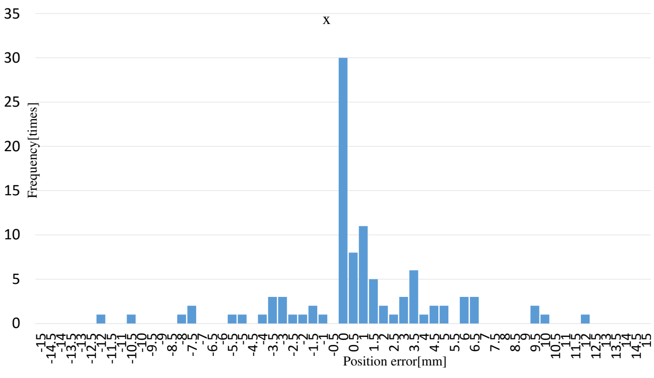
Table 2 Average error and standard deviation

	x[mm]	y[mm]	z[mm]	θ[deg]
Average error	0.525392	0.501954	9.961352	0.464328
Standard deviation (σ)	3.771772775	2.427697503	7.164103336	3.159272379

	x[mm]	y[mm]	z[mm]	θ[deg]
Average error (-3σ)	-10.78992632	-6.78113851	-11.53095801	-9.013489138
Average error (-2σ)	-7.018153549	-4.353441007	-4.366854672	-5.854216758
Average error (-1σ)	-3.246380775	-1.925743503	2.797248664	-2.694944379
Average error	0.525392	0.501954	9.961352	0.464328
Average error (+1σ)	4.297164775	2.929651503	17.12545534	3.623600379
Average error (+2σ)	8.068937549	5.357349007	24.28955867	6.782872758
Average error (+3σ)	11.84071032	7.78504651	31.45366201	9.942145138

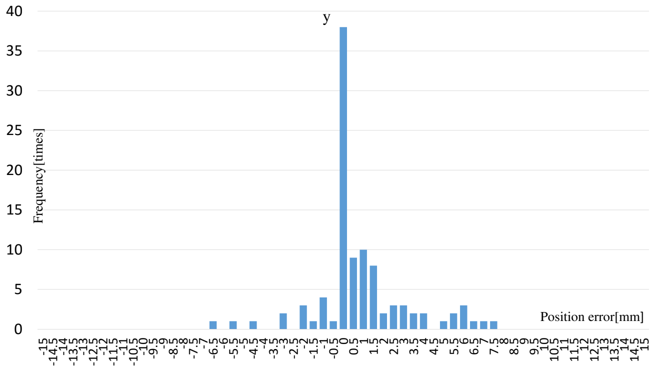


(a)

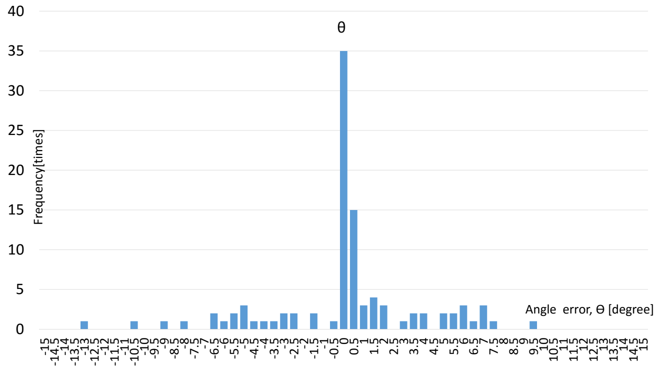


(b)

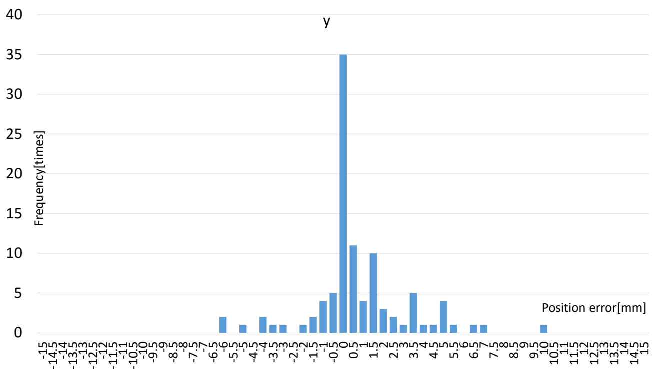
Fig.6 Histogram of position error (a)approximately 700[Lx] and (b) approximately 1000[Lx] (position x[mm])



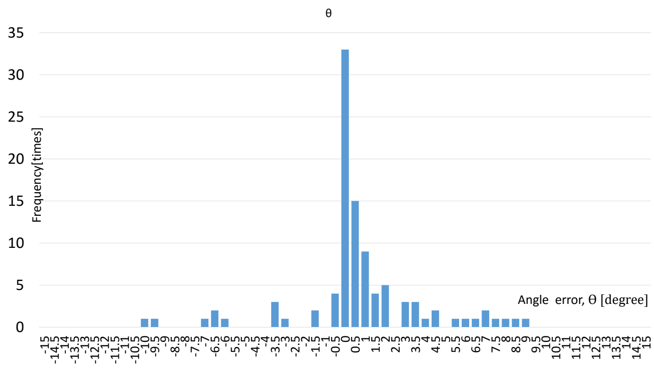
(a)



(a)



(b)



(b)

Fig.7 Histogram of position error (a)approximately 700[Lx] and (b) approximately 1000[Lx] (position y[mm])

Fig.8 Histogram of angle error (a)approximately 700[Lx] and (b) approximately 1000[Lx] (angle θ[deg])

4 Conclusion

In conclusion, we discuss handling performance under the two different optical environment conditions by calculating error and compare the result by histogram. In 100 times handling experiment, the cloth in each time can recognize and handle at both about 700[Lx] and 1000[Lx]. We concluded that the histogram of position errors (x, y and θ) at approximation of 1000[Lx] is larger than the normal optical environment (about 700[Lx]). The most beneficial result is getting at approximation 700[Lx]. This condition can be concluded by the experiments of same 100 times recognition, which can be known that practical experiment's result can approach to the best goal in handling performance.

References

[1] Ryohei ARAKI, " Simulated Decontamination Experiments by Mobile Manipulator with Visual Recognition " , Intermediate Paper

[2] W. Song, Y. Fujia, M. Minami, 3D Visual Servoing by Feedforward Evolutionary Recognition, Journal of Advanced Mechanical Design, Systems, and Manufacturing (JSME), Vol.4, No.4, pp.739-755, 2010.

[3] Wei Song, Mamoru Minami, Fujia Yu, Yanan Zhang and Akira Yanou : " 3-D Hand and Eye-Vergence Approaching Visual Servoing with Lyapunouv-Stable Pose Tracking" , IEEE International Conference on Robotics and Automation Shanghai International Conference Center, May 9-13, 2011, Shanghai China.

[4] 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.

[5] 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 System " SICE Annual Conference pp.101-107, 2012.

[6] Hongzhi Tian, yu Cui, Mamoru Minami and Akira Yanou, " Enhanced Tracking Performances of Eye-Vergence Visual Servoing with Evolutionary Pose Tracking". 2015