SICE Annual Conference 2008
August 20-22, 2008, The University Electro-Communications, Japan

Cognitive Resource Allocation Optimization for Real-time Multiple Object Recognition

Ryohei Ozawa and Mamoru Minami

Graduate School of Engineering, University of Fukui, Fukui, Japan
(Tel : +81-776-27-8527;Email: {ozawa,minami}@rc.his.fukui-u.ac.jp)

Abstract: This paper proposes a method of allocation of cognitive resource for efficient real-time multiple object recognition in robot vision. For human beings and even high-performance computers, the number of recognizable objects is limited at an instant time because cognitive and computational resources are limited. Thus, real-time reallocation of computational resources to the multiple objects is logical for robot to recognize them in real time under dynamically changing environment. We propose such dynamical reallocation scheme for robots recognize multiple objects effectively in real time. Positions of object models are represented by the genes of GA. The individuals of GA are allocated for recognizing multiple objects depending on the situation and priority of the object recognition. The proposed method is able to improve the speed of multiple object recognition. Some simulations of multiple object recognition will show the effectiveness of the method.

Keywords: Cognitive Resource Allocation, Visual Attention, Gazing-GA

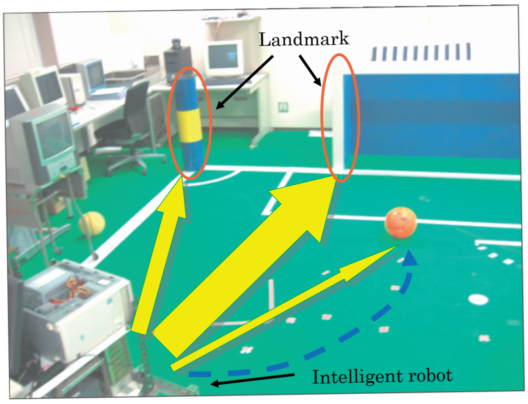

Fig. 1 Resource Allocation of Visual Attention for Localization of Mobile Robots

1. INTRODUCTION

Vision is one of important cognitive functions for both human beings and intelligent robots to act in a complex and dynamically changing environment [1]. In the vision field, recognition of multiple objects in real time is especially essential for adaptive and safe actions. Human vision can recognize multiple objects at an instant time, but the number of simultaneous recognizable objects is limited because of the restricted cognitive resource. Observing ability of human shows a cue to reconsider about cognitive function of robots whose computational resources are limited. There are several methods to recognize multiple objects [2]-[5], however, their calculation burden and processing time swell along with the increase of the number of recognition objects.

There is a well-known fact that humans move their eyes fast and focus attention on different objects to understand a scene. Based on the fact that the cognitive resource of humans at an instant time is limited, it can be assumed that humans recognize multiple objects in a dynamic scene by unconsciously and instantaneously integrating objects' information acquired in a short time-series. And we think that humans distribute cognitive resource for recognition of multiple objects by paying attention to the objects according to the priority order that depends on situations.

Based on the above consideration about human's vision, we consider that efficient allocation of cognitive resource is necessary for intelligent robots to recognize multiple objects at an instant time. In robot vision, a methodology named active vision has been proposed [6]-[8], which realizes a function of selective attention that can provide an ability of multiple object recognition for robots. We also think that the cognitive resource of robot would be better to be distributed according to the recognition priority of each object. There have been several proposed methods of cognitive resource distribution for object recognition and tracking. These methods are roughly divided into probabilistic method [9], [10] and nonprobabilistic method [11], [12]. In probabilistic method as particle filters, the number of particles represents the cognitive resource for recognition. And in nonprobabilistic method, the number of genes of Genetic Algorithms for object detection is used as the cognitive resource. However, neither of them deals with real-time reallocation of cognitive resource for multiple object recognition, as human beings.

In the presentation, we propose a method for real-time distribution of cognitive resources to make the robot recognize multiple objects effectively in real time. In the proposed method, recognition process is performed by model-based matching using Genetic Algorithm (GA). Positions of object models are represented by the genes of GA. The total number of genes represents the total number of cognitive resource. The balance of searching genes for each object is determined by the recognition priority of the objects. The proposed method improves recognition speed of multiple objects. Some simulations of multiple object recognition will show the effectiveness of the

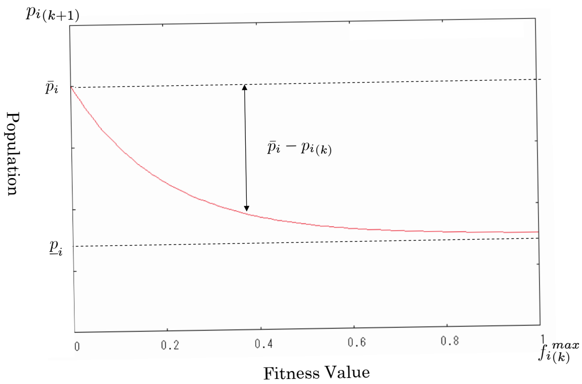

Fig. 2 Exponential Graph of This Method

method.

2. COGNITIVE RESOURCE ALLOCATION

We propose two recognition methods whose features are different on the point of how to allocate the cognitive resource as follows.

In the first case, the number of searching resources varies depending on whether a robot is searching for the object or tracking the object. When a robot is searching for an object, the maximum number of searching resources will be used since the more genes are used, faster the recognition is. Once the object is detected, less genes is enough to keep tracking object, because the genes for tracking will be converged in a narrow searching area. Gazing GA [11] is used to adjust the searching range in our method. The reduced resources can be used for realizing other objects not found.

In the second case, the number of searching resources is determined by recognition priority of the multiple objects. The priority depends on the situation of the robot. For example, when a soccer robot moves to catch a ball as shown in Fig.1, recognition for the ball is more important than the other objects. And, when the robot dribbles the ball to the goal, both of the ball recognition and the goal are important. In this case a priority varying with the robot's situation determines the allocation of the recognition resource for the ball and the goal. In this research, the priority of the objects in various situations is defined beforehand, and the number of searching resources will be changed according to the priority of the object recognition.

2.1 Allocation by Exponent Function and Priority

Allocation of cognitive resource for real-time recognition is realized by changing the number of searching genes. In this research, cognitive resources (population of GA) at $(k+1)$-th generation time are determined by exponent function according to fitness value of recognition at k-th generation time as shown in Fig.2. And surplus resources are allocated again according to priority of the object recognition. Equations for determining the number of resources is shown by follows.

In case of n objects to be recognized, maximum number of resources for ith object is defined as $\overline{p_i}$ $(i = 1, 2, 3, ..., n)$. And, minimum number of resources for ith object is defined as $\underline{p}_i$.

First, all resources defined as P_{max} are divided equally into n parts as follows.

$$\overline{p_i} = \frac{P_{max}}{n} \tag{1}$$

Second, by using $\overline{p}_i$, $\underline{p}_i$, the number of allocated resources defined as $p_i(k+1)$ is expressed by an exponent function of the maximum fitness value for i-th object at k-th generation $f_i{}^{max}(k)$ as follows.

$$\begin{aligned} & p_i(k+1) \\ & \qquad = (\overline{p}_i - \underline{p}_i) exp(-a f_i{}^{max}(k)) + \underline{p}_i \end{aligned} \tag{2}$$

,where a is a parameter adjusting the scale of the horizontal axis. The number of the surplus resources denoted as $P_{rest}(k+1)$ is given by

$$\begin{aligned} P_{rest}(k+1) &= \sum_{i=1}^{n} (\overline{p}_i - p_i(k+1)) \\ &= P_{max} - \sum_{i=1}^{n} p_i(k+1) \end{aligned} \tag{3}$$

.

Next, resources are allocated again according to priorities $w_i(\Sigma_{i=1}^{n} w_i = 1)$. w_i is updated to $\hat{w}_i$ by a given threshold of fitness value $\overline{f}$ that is to judge whether the object has been recognized. $\hat{w}_i$ is given by

$$\hat{w}_i = \left\{ \begin{array}{ll} w_i & (f_i{}^{max}(k) < \overline{f}_i) \\ 0 & (f_i{}^{max}(k) \geq \overline{f}_i) \end{array} \right\} \tag{4}$$

.

Moreover, in order to satisfy the reallocation condition $\Sigma_{i=1}^{n} \hat{w}_i = 1$, the renewed priority defined as $\bar{w}_i$ is given by

$$\bar{w}_i = \frac{\hat{w}_i}{\sum_{j=1}^{n} \hat{w}_j} \quad (i = 1, 2, 3, ..., n) \tag{5}$$

.

Finally, the number of resources $q_i(k+1)$ is determined by using priority $\bar{w}_i$ as follows.

$$q_i(k+1) = p_i(k+1) + P_{rest}(k+1)\bar{w}_i \tag{6}$$

3. RECOGNITION OF TARGET OBJECTS

3.1 Recognition by Model-based Matching

In this paper, a target object is recognized by the Model-based matching method, which is to correlate the input dynamic images through a known geometric model. The model of the target object is defined based on the shape of the target in the image, where correlation calculation includes shape-based derivation and integration of brightness and color distribution. We call the method surface-strips model [11]-[13]. We can give such condition of the environment that the maximum value of the correlation corresponds to exact position/orientation of the target object. Thus, the recognition problem will be

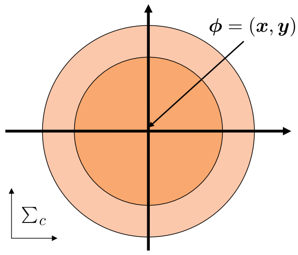

Fig. 3 Ball search model

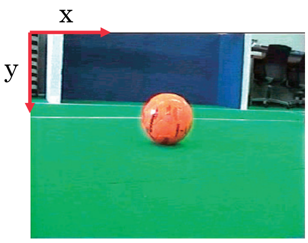

Fig. 4 Input image of ball search

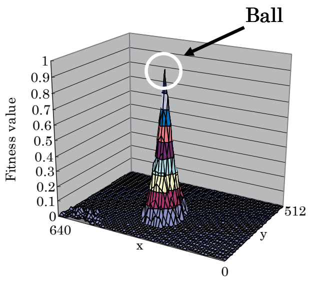

Fig. 5 Searching result by model of Fig.3

changed into an optimization problem of finding the maximum of the fitness function.

Target object for recognition in our simulation is a ball whose color and size are assumed to be known. The color information is used for calculating the fitness function. The details of the definition of the object model and the fitness function are described below.

3.1.1 Ball Model

We explain our recognition method by using ball recognition as a target example. The color of the ball used in RoboCup medium league is orange by regulation.

Here, we extract the orange area to recognize the ball by using a fixed range of hue value H. By preliminary experiments, it has been known that orange can be specified by limiting the value of H to 7-35. The image domain obtained from camera is expressed as follows:

$$\Omega_{camera} = \{\boldsymbol{r} = (x, y) \mid 0 \le x \le x_{max} , \; 0 \le y \le y_{max}\}. \quad (7)$$

Then, a set of orange point Ω_{orange} is expressed by the following equation:

$$\Omega_{orange} = \{\boldsymbol{r} = (x, y) \mid 7 < H(\boldsymbol{r}) < 35\}, \quad (8)$$

where, x_{max} and y_{max} in are positive constants representing the limited values along the axes x and y of image coordinate. The evaluation function of orange is defined as

$$h_o(\boldsymbol{r}) = \begin{cases} 1 & (\boldsymbol{r} \in \Omega_{orange}) \\ 0 & (\boldsymbol{r} \notin \Omega_{orange}). \end{cases} \quad (9)$$

Model-based matching method is used to detect the center position of the ball in input image. A searching model consists of a circle $\boldsymbol{S}_1$ and a strip $\boldsymbol{S}_2$ as shown in Fig.3. The area of $\boldsymbol{S}_1$ is denoted by n_1 [pixel] and $\boldsymbol{S}_2$ be n_2. In the Fig.3 variables $\boldsymbol{\phi} = (x, y)$ represents the center position of the model. Correlation with an input image and the searching model is defined as a fitness function given by:

$$F_{Size}(\boldsymbol{\phi}) = \frac{\sum_{\boldsymbol{r} \in \boldsymbol{S}_1(\boldsymbol{\phi})} h_o(\boldsymbol{r})}{n_1} - \frac{\sum_{\boldsymbol{r} \in \boldsymbol{S}_2(\boldsymbol{\phi})} h_o(\boldsymbol{r})}{n_2} \quad (10)$$

The filtering result by using Eq.10 with respect to the input image Fig.4 are shown in Fig.5 respectively. Filtering result in Fig.5 has a peak corresponding to the position of the ball. Then, the recognition problem is changed into an optimization problem of finding the maximum value of the filtered image, which is solved by GA that we will explain in the next section.

3.2 Real-time Recognition by Optimization using Gazing and Step GA

We employ a Genetic Algorithm (GA) for solving optimization problem because of its high performance of optimization. The GA is an optimization method since the iterative evolution process of the potential solutions toward better solutions is equivalent to the process of optimizing the objective function used as a fitness function in GA search. In our method, the positions of the searching models are coded into genes in GA.

To recognize a target object in real time, the searching models must converge to the target in the successively input images. To realize the real-time processing and detect the target position from the consecutively input images, we have proposed a new idea of an evolutionary recognition process for dynamic image, in which the evolution of GA is applied only one time to the newly input image. We named it “Step-GA”[12]. If the convergence speed of Step-GA is faster than the moving speed of targets in the dynamic image, real-time recognition can be realized.

Moreover, in order to increase the tracking performance, we employ a hybrid-searching method, which is a localized search technique of GA combined with global GA process. We call it “Gazing GA”. Gazing GA includes two main functions, that is, global and local searching, which are switched depending on the matching degree of the searching model with an input image. The effectiveness of Gazing GA has been confirmed in [14], where a hand-eye manipulator can catch a fish swimming in a pool with a net attached at the hand. By using these techniques of Step GA and Gazing GA, multiple objects, a ball, a goal, and an enemy, can be recognized in the video rate (30fps)[5]. However, this multiple object recognition method has a problem that it costs more recognition time when the number of recognition objects increases.

4. SIMULATION OF MULTIPLE OBJECT RECOGNITION

Simulation of multiple object recognition is performed to confirm the effectiveness the proposed method. Environment of simulation is shown in Fig.6 and Fig.7. Circles whose color and size are predetermined are placed arbitrarily in static image (320[pixel]×240[pixel]). Here, we perform two objects recognition and three objects

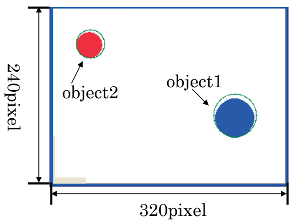

Fig. 6 Environment of 2 Object-Recognition Simulation

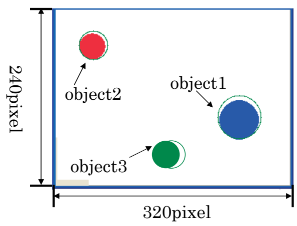

Fig. 7 Environment of 3 Object-Recognition Simulation

Table 1 Parameter of 2 Object-Recognition Simulation

Parameter	i	value
P_{max}	-	24
$\underline{p}_i$	1	4
	2	4
w_i	1	1/2
	2	1/2
$\overline{f}$	-	0.4
a	-	5

recognition separately. And in each simulation we discuss the method with cognitive resource allocation and the method without allocation by comparing their average time to recognize all objects, which is defined as T_{ave}. A threshold of fitness value $\overline{f}$ in Eq.(4) used to judge whether the object has been recognized is set at 0.4. Here, the number of recognition resources is represented by the number of genes of GA. In case of recognition with allocation, we use Eq.(6) to allocate the genes of GA. And in case of recognition without allocation, the number of genes of GA ($q_i(k+1)$) is fixed to constant as,

$$q_i(k+1) = \overline{p_i}. \tag{11}$$

4.1 Two-objects case

Firstly, we performed simulation of two objects recognition ($n = 2$). Parameters used in this simulation are shown in Table1. Here, the recognition priority for object1 and object2 are set equally as $[w_1, w_2] = [\frac{1}{2}, \frac{1}{2}]$. And, the total number of genes is set as 24.

The transient time profile of fitness value and the number of searching genes with cognitive resource allocation is shown in Fig.9, and the change of the fitness value without allocation is shown in Fig.8. Bar graph in Fig.9 represents the number of searching genes at each generation. We performed this simulation 30 times, respectively. Average time until all the objects recognized (T_{ave}) in the case of recognition with allocation and without is shown in Table 2. T_{ave} with allocation is 0.69[sec.] and without allocation is 1.20[sec.].

Comparing Fig.8 and Fig.9, we can see that cognitive resource allocation delay the rise of fitness of object1, but the average recognition speed of the two objects is faster than the case of without allocation whose in Fig. 8. From T_{ave} in Table2, we can see that multiple object recognition with cognitive resource allocation is 1.7 times faster than that of without allocation.

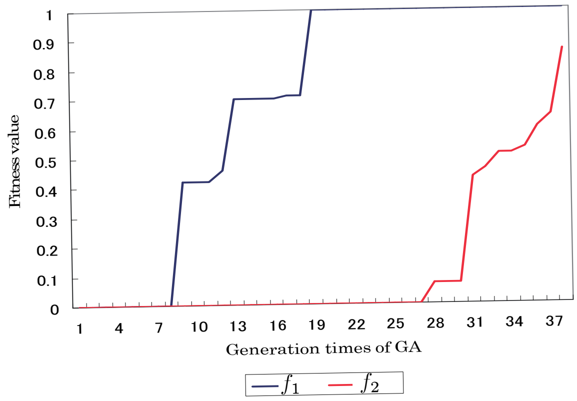

Fig. 8 Change of Fitness without Cognitive Resource Allocation Using Parameter of Table1 in 2 Object-Recognition Simulation

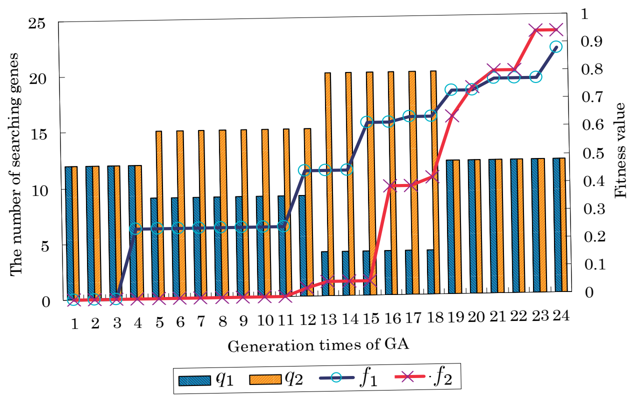

Fig. 9 Change of Fitness and Individual with Cognitive Resource Allocation Using Parameter of Table1 in 2 Object-Recognition Simulation

Table 2 Comparison of Average Processing Time with Allocation and without in 2 Object-Recognition Simulation

Cognitive resource allocation	True	False
$T_{ave}(s)$	0.69	1.20

Table 3 Parameter of 3 Object-Recognition Simulation

Parameter	i	value
P_{max}	-	24
$\underline{p}_i$	1	4
	2	4
	3	4
w_i	1	1/3
	2	1/3
	3	1/3
$\overline{f}$	-	0.4
a	-	5

Table 4 Comparison of Average Processing Time with Allocation and without in 3 Object-Recognition Simulation

Cognitive resource allocation	True	False
$T_{ave}(s)$	1.45	3.25

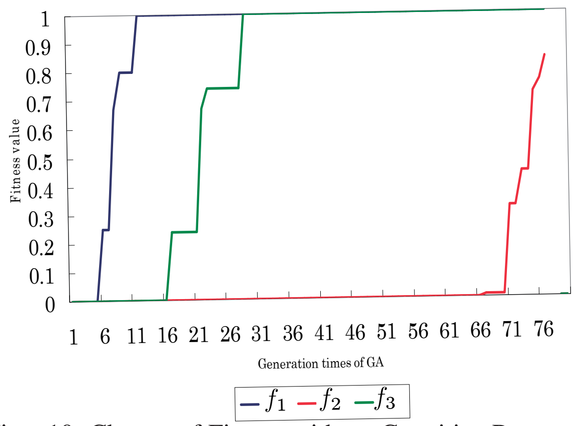

Fig. 10 Change of Fitness without Cognitive Resource Allocation Using of Table3 in 3 Object-Recognition Simulation

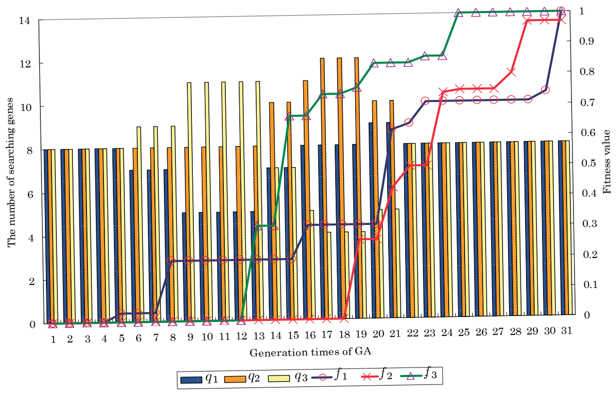

Fig. 11 Change of Fitness and Individual with Cognitive Resource Allocation Using of Table3 in 3 Object-Recognition Simulation

4.2 Three-objects case

Secondly, we performed simulation of three objects recognition ($n = 3$). Parameters used in this simulation are shown in Table 3. Here, priorities are set as $[w_1, w_2, w_3] = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ equally. And, the total number of genes is also 24.

The change of fitness value and the number of searching genes with cognitive resource allocation is shown in Fig.11, and the changing of the fitness value without allocation is shown in Fig.10. We performed this simulation 30 times, respectively. Average times until all the objects recognized (T_{ave}) in the cases with allocation and without are shown in Table 4. T_{ave} with allocation is 1.45[sec.] and without allocation is 3.25[sec.].

By comparison between Figs.10 and 11, we can see that cognitive resource allocation delays the rise of fitness of object1, but increase the speed of the multiple objects recognition efficiently, similarly to the two-objects case. From T_{ave} in Table 4, we can see that multiple object recognition with cognitive resource allocation is 2.2 times faster than that of without.

5. OPTIMIZATION OF PRIORITY

In this research, the number of searching resources is determined by the recognition priority of the object $w_i (i = 1, 2, ..., n)$. In general, bigger object can be recognized faster. Thus, we consider that efficient allocation according to the size of object will improve the average speed of recognition.

We performed simulation in the same environment as those in chapter 4 for examining the relationship between priorities and recognition speed. We wanted to estimate the probability P that n objects can be recognized when genes of GA evolving l times, with changing the ratio of priorities $\boldsymbol{w} = [w_1, w_2, ..., w_n]$,$\Sigma_{i=1}^{n} w_i = 1$. Here, the probability was calculated statistically by performing n objects recognition j times ($j = 1, 2, ..., T$). The algorithm of estimation of P is shown as follows.

First, we define the maximum fitness value of object i at k generation when jth simulation is performed as ${f_j}^i(k, \mathbf{w})$. And we define the generation number of GA when all n objects have been recognized as $p_j(\boldsymbol{w})$. A threshold of fitness that is to determine whether an object has been recognized is defined as $\overline{F}$. By using $\overline{F}$, $p_j(\boldsymbol{w})$ is defined as follows.

$$p_j(\boldsymbol{w}) = \{\mathrm{k} \mid \bigcap_{i=1}^{n} {f_j}^i(k, \boldsymbol{w}) > \overline{F}\} \tag{12}$$

Secondly, set of $p_j(\boldsymbol{w})$ which has recognized all n objects within l generation is defined as $Q(l, \boldsymbol{w})$, given by

$$Q(l, \boldsymbol{w}) = \{p_j(\boldsymbol{w}) \mid p_j(\boldsymbol{w}) < l, \quad j = 1, 2, ..., T\} \tag{13}$$

Let N represts a function that counts the number of the elements in set Q, probability of recognition P is given by

$$P(l, \boldsymbol{w}) = \frac{1}{T} N(Q(l, \boldsymbol{w})). \tag{14}$$

5.1 Simulation of Priority of Recognition

We performed two simulations by using two objects in the same environment as that of chapter 4. In each simulation, we performed 1000 times (T=1000) and set the number of all resources to 40. Additionally, the number of resources of each objects q_i is newly allocated in every evolution of GA as follows.

$$q_i = P_{max} w_i \tag{15}$$

In the first simulation, we set the ratio of the two objects' size as [object1:object2]=[4:1]. And, we examine the profile of recognition probability $P(l, \boldsymbol{w})$ with the conditions of three rates of priorities separately ($\boldsymbol{w} = [0.1, 0.9], [0.5, 0.5], [0.9, 0.1]$). Changes of $P(l, \boldsymbol{w})$ in three cases are shown in Fig.12. We can see that probability of recognition $P(l, \boldsymbol{w})$ change according to rate of priorities $\boldsymbol{w}$ from Fig.12.

In the second simulation, we set the ratio of the two objects' size as [object1:object2]=[1:1],[4:1],[6:1]. And, we discussed the probability of recognition $P(l, \boldsymbol{w})$ when $l = 40$ in case of several ratios of priorities. Change of probability of recognition $P(40, \boldsymbol{w})$ is shown in Fig.13.

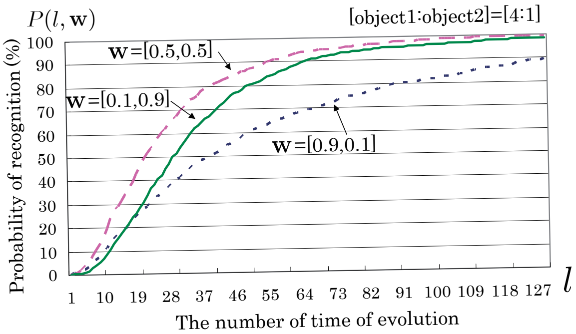

Fig. 12 Probability of recognition of 2 Object-Recognition Simulation ([object1:object2]=[4:1])

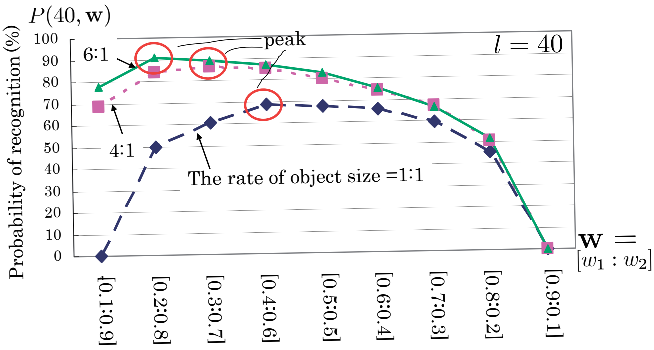

Fig. 13 Parametric Space of 2 Object-Recognition Simulation

We can see three peaks of most efficient allocation, one is in the case of $\boldsymbol{w}$=[0.4:0.6] when the rate of object size is [1:1], the second one is in the case of $\boldsymbol{w}$=[0.3:0.7] when the rate of object size is [4:1] and the third one is in the case of $\boldsymbol{w}$=[0.2:0.8] when the rate of object size is [6:1].

From these simulations, we can see that probability profile of recognition depends on the rate of priorities, and the peak of most efficient parameter combination exists.

6. CONCLUSION

We proposed a method to improve the efficiency of multiple object recognition by real-time allocation of cognitive resource determined by recognition priority of the multiple objects. Recognition process is performed by model-based matching using Genetic Algorithm (GA). Positions of object models are represented by the genes of GA. The total number of genes determines the total cognitive resources. The number of searching genes for different objects is changed according to the priority of the object recognition. The proposed method raises average recognition speed of multiple objects. Effectiveness of the method has been confirmed by simulations.

Additionally, we investigated where the peak of most efficient allocation exists. In future, we will develop the method of changing the rate of priorities for multiple object recognition in real time adaptively.

REFERENCES

[1] R.A.Brooks, *Model-Based Three-Dimensional Interpretations of Two-Dimensional Images*, IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI-5, 2, pp.140-150, 1983.

[2] S Yamamoto, Y. Mae, Y. Shirai, and J. Miura, “Realtime multiple object trackingbased on optical flows,” Proc. IEEE International Conference on Robotics and Automation, pp.2328-2333, 1995.

[3] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, ”Multi-Camera Multi-Person Tracking for Easy Living,” Proc. IEEE International Workshop Visual Surveillance, pp. 3-10, 2000.

[4] T. Darrell, G. Gordon, M. Harville, and J. Woodfill, “Integrated person tracking using stereo, color, and pattern detection,” International Journal of Computer Vision, vol.37, no.2, pp.175-185, 2000.

[5] T.Mizushima,Y.Sakamoto,W.Song,M.Minami,Y.Mae, “Generation of Adaptive Motion Using Quasi-simultaneous Recognition of Plural Targets,” Proc. Internatoinal Conference on Control, Automation and Systems, CD-ROM,2005.

[6] Y. Aloimonos and I. Weiss, “Active Vision,” International Journal of Computer Vision, vol.1, no.4, pp.333-356, 1988.

[7] R. Bajcsy, “Active Perception,” IEEE Proceedings, vol.76, no.8, pp.996-1006, 1988.

[8] D.H. Ballard and C.M. Brown, “Principles of animate vision,” In Y.Aloimonos (ed.), Active Perception, Lawrence Erlbaum Association, Publishers, pp.245-282, 1993.

[9] M. Isard and A. Blake, ”Condensation-Conditional Density Propagation for Visual Tracking,” International Journal of Computer Vision, vol. 29, no. 1, pp.5-28,1998.

[10] Dorin Comaniciu, Visvanathan Ramesh, Peter Meer, ”Kernel-Based Object Tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 564-575, 2003.

[11] M.Minami, H.Suzuki, M.Miura, “Real-Time Robust Recognition of Human using Illuminance-depending Gazing GA,” Proc. Knowledge-Based Intelligent Information Engineering Systems and Allied Technologies, pp.1377-1382, 2002.

[12] M.Minami, J.Agbanhan and T.Asakura, “Manipulator Visual Servoing and Tracking of Fish Using Genetic Algorithm,” International Journal of Industrial Robot, vol.29, no.4, pp.278-289, 1999.

[13] M.Minami, J.Agbanhan and T.Asakura, “A Method of Model-Based Object Recognition,” Japan/USA Symposium on Flexible Automation, vol.2, pp.905-912, 1996.

[14] M.Minami, H.Suzuki, J.Agbanhan, T.Asakura, *Visual Servoing to Fish and Catching Using Global/Local GA Search*, Int. Conf. on Advanced Intelligent Mechatronics, Proc., 2001, pp.183-188.