World Automation Congress © 2010 TSI Press.

ACTIVE SENSING FOR INTELLIGENT ROBOT VISION WITH RANGE IMAGING SENSOR

HIRONOBU SASAKI[1], TOMOYA FUKUKAWA[1], CHEN FEI[1], SUN BAIQING[1], NAOYUKI KUBOTA[2], TOSHIO FUKUDA[1]

1: NAGOYA UNIVERSITY
2: TOKYO METROPOLITAN UNIVERSITY

ABSTRACT— Recently, various types of robots have been researched and developed for supporting our life. Also, the perceptual system for the robot is researched. Visual perception includes a lot of valuable information and it is useful for all intelligent robot system. In this research, we discuss about 3-D information structuring for intelligent robot vision. There are many visual sensors and we use the range-imaging camera to detect distance and image data. We propose a method for perceive moving target by using Growing Neural Gas. In the experimental results, we show the potency of our method

Key Words: Robot Vision, Range Imaging Sensor, Growing Neural Network

1. INTRODUCTION

Various types of robots have been used in industrial site, and automated machinery manufacturing process is increasing. On the other hand, flexible parts assembling or adaptable assembling according to the situation is still difficult for robot. Because of development of the robotics technology, the assembling scene with people are will be decrease. But for the final checks or error recovery, people will be still needed. Here, we can expect the human-robot collaboration improves the work efficiency. An active sensing and Information Structuring is necessary to achieve the system that works as human cooperates with the robot in confined environment. The sensing of human and the environment is important elements to achieve the cooperation working of human and the robot. However, fixed sensor makes some blind spot, part that more detail information required or unimportant part, is exist. In this paper, we develop active sensing system that can change searching position or change the priority of searching point to get blind position and considerable position information. Off course, collected information is enormous quantity of data, we should organize information. In this instance, we consider about Active Sensing and Information Structuring for effective information use. Off course, collected information is enormous quantity of data, we should organize information. In this instance, we consider about Active Sensing and Information Structuring for effective information use. We will be develop the cooperated work system of human and the robot with an effective teaching for the robot or human state detection, etc. by using above-mentioned concept.

2. ROBOT VISION

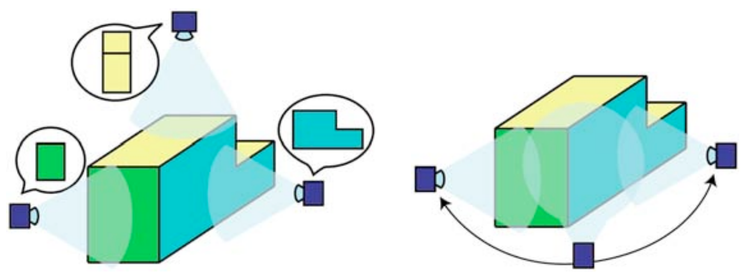
The mobile robot have vision sensors for sensing ambient environmental information. In the visual image, there are so many useful information to extract from the color information. The research of the part of “Perception" in robot vision is still insufficient. The individual identification of the object, detection and identification of the movement, grouping of the objects are valuable for the intelligent robot vision. Labeling for the recognized object is useful for "Perception". And we should select information by the task with Digital Image processing(Fig.2). However, it is difficult to extract the 3-D information from visual image, that is necessary for mobile robot, and we need to put in another sensor to the robot. Range imaging camera can easily get a lot of information(Table.1, Fig.3). By using this sensor data, GNG is useful for structuring the information group. Growing Neural Gas can dynamically change the adjacent relation (edge) referring to the ignition frequency of the adjacent node(Fig.4). And we apply GNG for unsupervised clustering of the distribution of radiance value and the distance data (Fig.5). Proposed method will be helpful for labeling the clusters, and utilize the information.

2.1 Environmental Sensing

The static image processing is useful for robot vision, such as edge detection by laplacian filter, noise reduction by gaussian filter, and so on. And we can process sequential image to processing videos, such as by moving object detection by using difference picture or optical flow. However the vision sensors are easily influenced of environment light condition. And it is difficult to understanding physical relationship between objects.

There are various methods to gather three dimensional information on the object. And, information from various viewpoint is necessary to gather the in-depth information. It is able to get 3-dimensional information by using multiple vision sensors. Because the sensor position is already-known as the merit, the information integration is easy. But the sensor area is fixed and it is difficult to setting cameras with accurate positional relationship. Another method is moving the sensor actively to gather the consecutive view point information. It has the robustness to the environmental changes, but the movement mechanism and position estimation of the sensor is needed(Fig I).

On the other hand, we can get the distance data by ultra sonic sensor or laser range finder, and so on. However, these sensors only can get a point information or linear information. So if we wont to get the 3 dimensional distance data, we have to move or rotate these sensors. In general, these kind of sensors are uses for mapping or localize ones position.



(a)Multiple view (b)Consecutive view

Fig I. Multiple information Structuring

2.2 The range-imaging camera

In the case of gathering image data and 3 dimensional surface data simultaneously, we generally uses cameras and range sensors. However it is difficult to synchronize or we have to compute differences of physical relationship between sensors. In this paper we use SR-3000 range imaging camera(Fig II). Range imaging camera can gather image data and 3 dimensional surface data simultaneously. The SR-3000 is an optical imaging system which offers real time 3-dimensional image data. It has infrared camera so the gathering data is almost free of the influence of environment light, but alternatively, it can't gather color image.

(a) Value input (b) Distance input

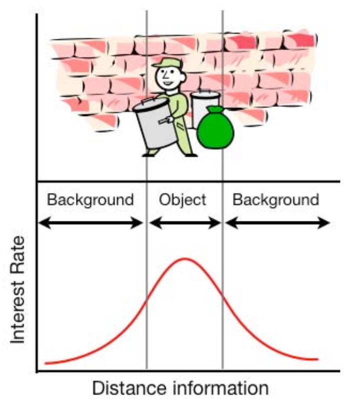
Fig II. SR3000 (MESA Imaging AG)

By using this sensor, we don't have to prepare the transform formula for sensor fusion of image and distance data. And we can clustering these data for detecting objects, reintegrate the 3 dimensional image by radiance value, or labeling to the objects.

3. OBJECT DITECTION AND INFORMATION STRUCTURING

3.1 Digital Image Processing

The camera image include a copious information, so the digitization method has some availableness. However, averagely digitization is sometimes un-useful in case of many objects are observable. Because important information is not uniformly distributed in the space but biased. Therefore, we can say that better information can be collected by focused digitalization. Because volume of information is too large, calculation cost reduction method or the information extraction method is necessary while the data obtained with a camera contains various information. Therefore, we use the concept of ROI(Region of Interest) to get the gazed area information and reduce the calculation of other area.



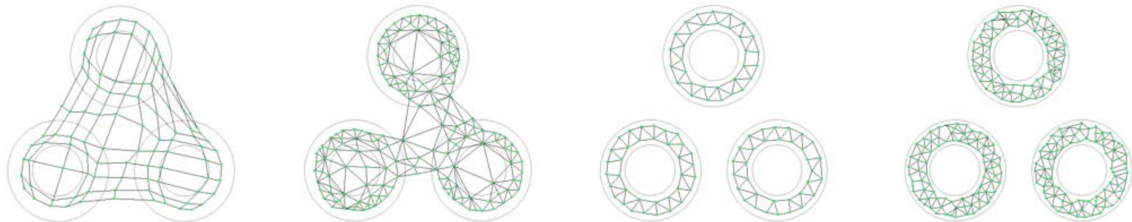
(a)Region of interest (b)Digital image processing

Fig III. Information extraction from ROI by the filter

3.2 Growing Neural Gas

Various types of pattern matching methods such as template matching, cellular neural network, recognition, and dynamic programming (DP) matching, have been applied for human detection problems. In general, pattern matching is composed of two steps of target detection and target recognition. The aim of target detection is to extract a target candidate from an image, and the aim of the target recognition is to identify the target from classification candidates. In this paper, we focus on the target detection, because the main aim of this paper is to discuss on growing neural gas.

Unsupervised learning is performed by using only data without any teaching signals [8-12]. Self-organized map (SOM), neural gas (NG), growing cell structures (GCS), and growing neural gas (GNG) are well known as unsupervised learning methods. Basically, these methods use the competitive learning. The number of nodes and the topological structure of the network in SOM are designed beforehand [8]. In NG, the number of nodes is fixed beforehand, but the topological structure is updated according to the distribution of sample data [9]. On the other hand, GCS and GNG can dynamically change the topological structure based on the adjacent relation (edge) referring to the ignition frequency of the adjacent node according to the error index. However, GNG does not delete nodes and edges, while GNG can delete nodes and edges based on the concept of ages [10]. Furthermore, GCS must consist of k-dimensional simplifies whereby k is a positive integer chosen in advance. The initial configuration of each network is a k-dimensional simplex, e.g., a line is used for k=1, a triangle for k=2, and a tetrahedron for k=3 [11]. GCS has applied to construct 3D surface models by triangulation based on 2-dimensional simplex. However, because the GCS does not delete nodes and edges, the number of nodes and edges is over increasing. Furthermore, GCS cannot divide the sample data into several segments. Fig IV shows how to cluster the data by GNG. GNG cluster the data by nodes and edges. GNG is topological clustering method, and it cluster the data with the shape (Fig IV).



(a)Self Organizing Map (b)Growing Cell Structure (c)Neural Gas (d)Growing Neural Gas

Fig IV. Data Clustering Algorithm

GNG can dynamically change the adjacent relation (edge) referring to the ignition frequency of the adjacent node. We apply GNG for unsupervised clustering of the distribution of radiance value and the distance data. The learning algorithm of GNG is shown as follows. The nth dimensional reference vector of the ith node is w_i; a set of nodes is A; a set of nodes connected to node i is N_i; a set of edges is C; and the age of the edge between the ith and jth node is $age(i,j)$.

Step 1. *Select two units $c1$, $c2$ at random position w_{c1}, w_{c2} in $\mathbf{R}^n$. Initialize the connection set.*
Step 2. *Determine the nearest unit $s1$ and the second-nearest unit $s2$ according to input signal ξ by*

$$s_1 = \arg\min_{c\in A} \|\xi - w_c\| \quad \text{and} \quad s_2 = \arg\min_{c\in A\backslash\{s_1\}} \|\xi - w_c\| \tag{1}$$

where ξ is composed of the position (x, y) and radiance value on the image (Fig.6(a)).
Step 3. *If a connection between $s1$ and $s2$ does not yet exist , create it. Set the age of the connection between $s1$ and $s2$ to zero.*

$$age_{(s1,s2)}=0 \tag{2}$$

Step 4. *Add the squared distance between the input signal and the winner to a local error variable E_{s1}.(Fig.6(b)).*

$$E_{s1} \leftarrow E_{s1} + \|\xi - w_{s1}\|^2 \tag{3}$$

Step 5. *Adapt the reference vectors of the winner and its direct topological neighbors by the learning rate ε_b and ε_n, respectively.*

$$\nabla w_s = \varepsilon_b\left(\xi - w_s\right) \quad \nabla w_n = \varepsilon_n\left(\xi - w_n\right) \tag{4}$$

Step 6. *Increment the age of all edges emanating from $s1$.*

$$age_{s1} \leftarrow age_{s1} + 1 \tag{5}$$

Step 7. *Remove edges with the age larger than a_{max}. If units have no more emanating edges after this, remove those units.(Fig.6(c))*
Step 8. *If the number of input signals generated so far is an integer multiple of a parameter λ, insert a new unit as follows.(Fig.6(d))*

In addition, the node is generated based on the distance of the intent objects. Paying attention to the object is useful for the reduction of nodes that indicate the useless obstacles. And also, we can separate off the adjacent object by paying attention to the object area. We set the node additional space threshold by the position of intent objects. In future we will be using normal distribution or gaussian membership function for node additional probability.

In this way, the radiance value distribution can be extracted from the image by using GNG. Moreover, if the Delaunay triangulation consists of GNG, three dimension information can be restored by using the texture of brightness.

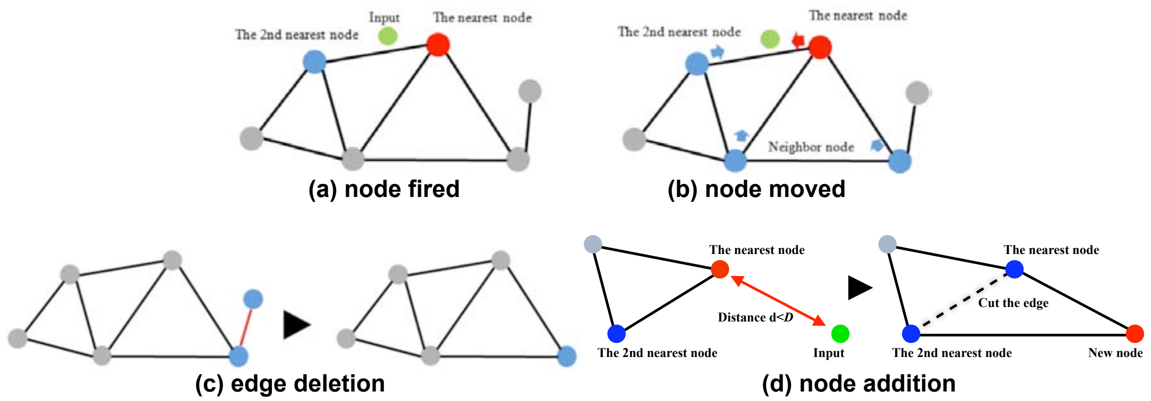


Fig V. How to learn GNG nodes and edge

4. EXPERIMENTAL RESULT

By moving sensor, we implemented the active sensing method for separation from background and the information structuring of the object. The range-imaging camera was installed in the point of the arm robot, and the information gathering was done while rotating surroundings of the object. The sensor works with the arm robot by tracks of circle that centers on the object to gather the consecutive view point information.

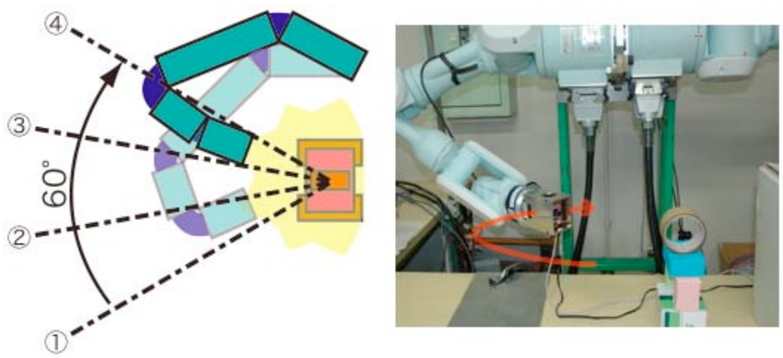


Fig VI. Experimental Environment : SR3000 with arm robot

In the experimental results, we show the results of classification for detecting object by using GNG for the range imaging camera data. The generation of the node and the edge was limited according to the output result of the digital image processing with concept of ROI.

Fig VII shows the result of sensor information at 0 degrees, 20 degrees, 40 degrees, and 60 degrees, and clustering result of GNG. We can see that the object is separated from background as the calculation advances. However, we couldn't get the accurate shape of the object. However, because of a problem that information on the edge is omitted during clustering, we couldn't get the accurate shape of the object.

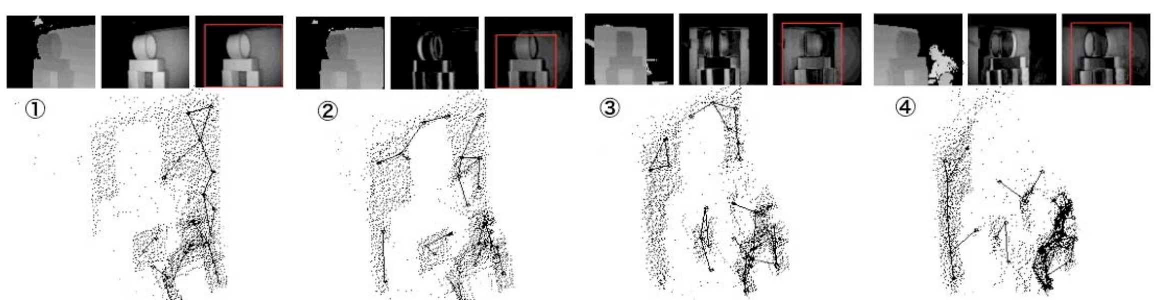


Fig VII. Consecutive data clustering

5. CONCRUTION

In this research, we aimed at the construction of the intelligent robot vision. We using GNG for self-organized information of the object. And we discussed about the usage of the structured information. We discussed about the effectiveness of structurizing method for three dimension information by GNG. Three dimension information on the object was structurized by range-imaging camera with consecutive information. In the experiment result, the boundary of the background and the object was able to be clarified.As future tasks, we improve GNG to preserve more detailed three dimension shape. Also, we should structurize the information on the object in detail. To improve the utilization efficiency of information, we make some groups of the information by applying upper and lower level of the GNN layer.

As mentioned before, range imaging camera and GNG is useful for active sensing. Range imaging camera can easily get a lot of information. However, proposal algorithm is unsatisfactory. So in the future tasks, we have to propose more effective algorithm such as difference filter, long-term memory filter, radicalization.

The research of the part of “Perception" in robot vision is still insufficient. The individual identification of the object, detection and identification of the movement, grouping of the objects are valuable for the intelligent robot vision. Here, we can assume that the labeling for the recognized object is useful for "Perception". And proposed method will be helpful for labeling the clusters, and utilize the information.

REFERENCES

1. Anderson, J. A. & Rosenfeld, E. (1988). Neurocomputing, The MIT Press, Cambridge, Massachusetts, US
2. David Marr (1982), Vision, W. H. Freeman, San Francisco.
3. K.Fukushima (2003). Neural network model restoring partly occluded patterns, Knowledge-Based Intelligent Information and Engineering Systems, (7th International Conference, KES 2003), Part II, eds: V. Palade, R. J. Howlett, L. Jain, Berlin - Heidelberg: Springer-Verlag, pp. 131-138
4. H. H. Bulthoff, S..W. Lee, T.A. Poggio, & C. Wallraven (2002), Biologically Motivated Computer Vision, Springer-Verlag.
5. J. H. Holland, “Adaptation in Natural and Artificial Systems“, First MIT Press Ed., Massachusetts: The MIT Press, 1992.
6. D. Fogel, “Evolutionary Computation“, New York: IEEE Press, 1995.
7. J. Kennedy and R. Eberhart, “Particle Swarm Optimization” Proc. IEEE Int. Conf. Neural Netw., Perth, Australia, pp. 1942-1945, 1995.
8. T. Kohonen: Self-Organization and Associative Memory; Springer-Verlag, 1984.
9. T. M. Martinetz and K. J. Schulten: A “neural-gas" network learns topologies; Artificial Neural Networks, Vol. 1, pp. 397-402, 1991.
10. B. Fritzke: A growing neural gas network learns topologies; Advances in Neural Information Processing Systems, Vol. 7, pp. 625-632, 1995.
11. B. Fritzke: Growing cell structures – a self organizing network in k dimensions; Artificial Neural Networks, Vol. 2, No.2, pp. 1051-1056, 1994.
12. K.A.J. Doherty, R.G. Adams, N. Davey: Hierarchical Growing Neural Gas; Adaptive and Natural Computing Algorithms, pp. 140-143, 2005.
13. N. Kubota, “Visual Perception and Reproduction for Imitative Learning of A Partner Robot“, WSEAS Trans. Signal Processing, vol. 2, no.5, pp. 726-731, 2006.
14. N. Kubota, “Computational Intelligence for Structured Learning of A Partner Robot Based on Imitation“, Info. Sci., vol. 171, no. 4, pp. 403-429, 2005.
15. I. A. Sulistijono and N. Kubota, “Particle Swarm Intelligence Robot Vision for Multiple Human Tracking of A Partner Robot“, Proc. Society of Instrument and Control Eng. Annual Conf., pp. 604-608, 2007.
16. I. A. Sulistijono and N. Kubota, “Evolutionary Robot Vision and Particle Swarm Optimization for Multiple Human Heads Tracking of A Partner Robot“, Proc. IEEE Cong. Evolutionary Comp., pp. 1535-1541, 2007.