3차원 물체의 인식 성능 향상을 위한 감각 융합 시스템

Sensor Fusion System for Improving the Recognition Performance of 3D Object

Ji Kyoung Kim, Yeong Jae Oh, Kab Sung Chong, Jae Woo Wee, Chong Ho Lee
Dept. Information Technology & Telecommunication, Inha University

Abstract - In this paper, authors propose the sensor fusion system that can recognize multiple 3D objects from 2D projection images and tactile information. The proposed system focuses on improving recognition performance of 3D object. Unlike the conventional object recognition system that uses image sensor alone, the proposed method uses tactual sensors in addition to visual sensor. Neural network is used to fuse these informations. Tactual signals are obtained from the reaction force by the pressure sensors at the fingertips when unknown objects are grasped by four-fingered robot hand. The experiment evaluates the recognition rate and the number of learning iterations of various objects. The merits of the proposed systems are not only the high performance of the learning ability but also the reliability of the system with tactual information for recognizing various objects even though visual information has a defect. The experimental results show that the proposed system can improve recognition rate and reduce learning time. These results verify the effectiveness of the proposed sensor fusion system as recognition scheme of 3D object.

Key Words :Intelligent System, Neural Network, Sensor Fusion System, Object Recognition, Haptic Recognition

1. Introduction

A great deal of research has been introduced in relation to object recognition. There are two main approaches in object recognition. One method is model-based recognition. This scheme uses the 3D object representation in an object-centered coordination system [1], [2]. This method recognizes object by matching an input image with the model image that is obtained from the 3D object representation. However, some problems of model-based recognition are to consider that the 3D object model must be built in advanced and search range is extremely large during execution.

Second method is the view-based recognition that uses an object representation based on various 2D projection images of the 3D object [3], [4]. This method recognizes an object by matching an input image with multiple 2D projection images that are provided in advance. One problem of view-based recognition is that matching with learning image must be flexibly performed so that leaning images can correspond to the various 2D projection images of the 3D object.

In this paper we propose the sensor fusion system for improving 3D object recognition rate and decreasing the number of learning iteration. Tactual senses are obtained from reaction forces measured by grasping unknown object with four-fingered robot hand including pressure sensors at its fingertips. And visual sense is obtained from CCD camera. This sensory information is fed to neural network for learning of each object.

2. System Configuration

Fig. 1 shows the data flow of the proposed system. It has three components: visual perception component that consists of CCD camera and vision frame grabber, tactual perception component that consists of robotic hand and tactile sensor in it and sensor fusion component that uses neural network (NN).

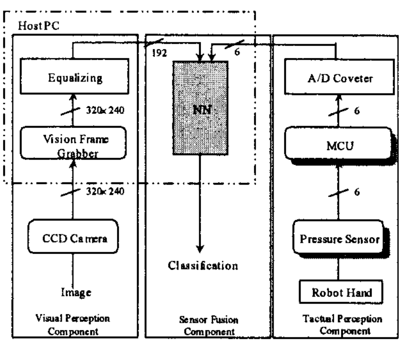

Figure 1. Data flow of sensor fusion system.

Size of the image information in vision frame grabbers is 320×240 pixel and this information is transmitted with format of 8 bit gray scale. In our experimentation, size of this information is too large for use. So this image information is transformed. Transformed size 16×12 unit. Simultaneously this signal is digitalized. That is, because has 0~255 pixel intensity, it is transformed into 0 or 1 value according to the threshold value which is determined as 123 through experiment.

For haptic recognition, tactual perception component is consisted of seven pressure sensors in four-fingered gripper and AVR micro-control unit (MCU).

Tactile sensors can treat the value of the pressure force that have range of 0~10 lbs. For receiving tactile information of the pressure sensors, the internal cache in MCU is used. It receives analog sensory inputs and temporally stores in the cache and then converts analog signal into digital signal. After conversion, it transfers digitalized pressure information to NN.

If w is the length of one side of the object, finger A is firstly located on $w/3$ or $w/5$ of the one side for training and test. Once the finger A is placed on these positions, the other finger B, C, D are automatically relocated because of fixed inter-contact distance. Then finger A moves to next position where is $2w/3$ or $2w/5$. Incremental value $1/3$ or $1/5$ is applied to all sides of all objects for making force patterns. These patterns form the train set and test set that are fed to sensor fusion component. And they are used to train and test the shape recognition. Fig. 2 presents some examples of rotational grasp position along the object side and Fig. 3 illustrates the shapes of eleven objects.

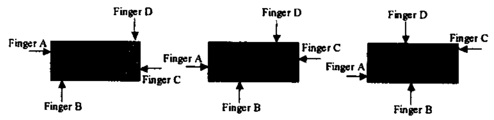

Figure 2. Rotational grasp positions.

Sensor fusion component consists of vision and tactile information as input of NN. But when visual perception component has insufficient information or malfunction, NN does a function as recognition system with a few tactile sensors. So, visual input signal and tactile input signal of NN is superimposed. Vision input signal is connected with vision frame grabber and its input size for NN is 192. And tactile input signal for NN is 6. Presser sensors in finger B, C and D are used for input signal of NN. Presser sensor 0 is except from input signal. This pressure sensor only has a role as guide of contact position when robotic hand grasps objects. This pressure sensor only has a role as guide of contact position when robotic hand grasps objects. This component for recognizing object follows multi-layered NN in the form of a NN structure. Specification used in this component is illustrated in Table I.

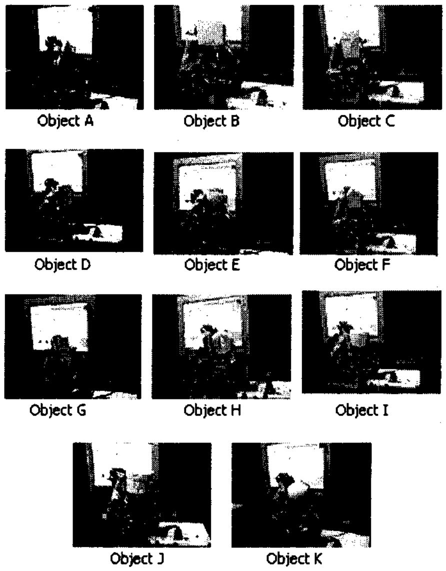

Figure 3. Shapes of eleven objects.

Table I. NN Specification

The number of input neuron	Image(192)+pressure(6)
The number of hidden neuron	Image(82)+pressure(10)
The number of output neuron	11
Learning rate	0.45
Tolerance	0.5
Max iteration	2000
Learning method	BP algorithm

3. Experimental Method and Results

3.1. Processing of Sensor Fusion Data

Training set of image information is as follow. The spherical observation surface is divided into both and according to the mesh that is equally spaced with latitudinal and longitudinal line per 60 degree. This data set consists of 2D projection image at the intersections (6×3=18). Test set of image information is as follow. The spherical observation surface is divided into both and according to the mesh that is equally spaced with latitudinal and longitudinal lines per 15 degree. This test set consists of 2D projection image placed at the intersections (12×6=72).

For haptic making training data and test data, finger A is located on $w/3$, $2w/3$, $3w/3$ and $w/5$, $2w/5$, $3w/5$, $4w/5$, $5w/5$ of all sides for all objects. Once the finger A is placed on these position, the other finger B, C, D are automatically relocated because of fixed inter-contact distance. So, object A has 9 training patterns and 15 test patterns. Object B, C, D, E, G, I, J have 12 train patterns and 20 test patterns. Object H has 2 training patterns and 2 test patterns. Object K has 1 training patterns and 1test patterns. Table Ⅱ sows some examples of classification with value of six pressure sensors for all objects.

Table Ⅱ. Examples of Classification by Test Set

P1	P2	P3	P4	P5	P6	object
3.21	3.23	4.81	4.84	4.49	4.51	A
3.39	3.41	4.41	4.43	4.38	4.71	A
2.24	2.25	2.29	2.31	2.21	2.19	B
2.13	2.16	2.15	2.14	2.12	2.13	B
2.49	2.50	4.10	4.12	2.48	2.47	C
2.62	2.63	5.13	5.13	2.60	2.63	C
3.76	3.77	5.54	5.52	3.80	3.80	D
3.72	3.71	5.49	5.50	3.82	3.83	D
5.53	5.53	7.21	7.20	5.48	5.49	E
5.63	5.66	7.15	7.17	5.66	5.68	E
1.92	1.93	3.34	3.38	2.00	2.04	F
1.96	1.99	3.29	3.31	1.96	1.97	F
7.71	7.70	8.02	8.05	7.69	7.71	G
7.64	7.67	8.11	8.15	7.64	7.62	G
5.55	5.56	5.55	5.56	5.55	5.56	H
5.59	5.61	5.59	5.61	5.59	5.61	H
4.02	4.06	4.56	4.58	4.04	4.07	I
4.96	4.95	5.01	5.05	4.95	4.97	J
6.23	6.26	6.23	6.26	6.23	6.26	K

3.2. Results

Fig. 4 shows the recognition rate for eleven objects. The figure indicates that recognition of sensor fusion is better than that of vision sensing. And haptic recognition is good performance even rather than that of visual recognition for similar shape of object. That is, classification results about object C, D, E, G and object I, J are better than that of visual recognition. Recognition using visual sense for these six objects is less than 73 % but recognition using tactile senses is more than 85 %.

And classification for all eleven objects by using sensor fusion is more than 93 %. This result verifies that proposed sensor fusion system in this paper has the most efficiency for object recognition. This system is especially efficient for similar-shaped object.

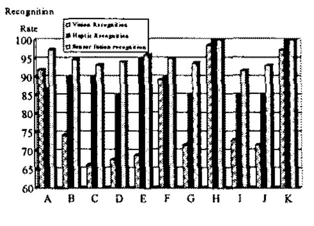

Figure 4. Comparison of recognition rate.

Fig. 5 illustrates that sensor fusion system reduces the number of learning iterations for each object. This figure verifies that sensor fusion system can improve learning speed in comparison with visual recognition.

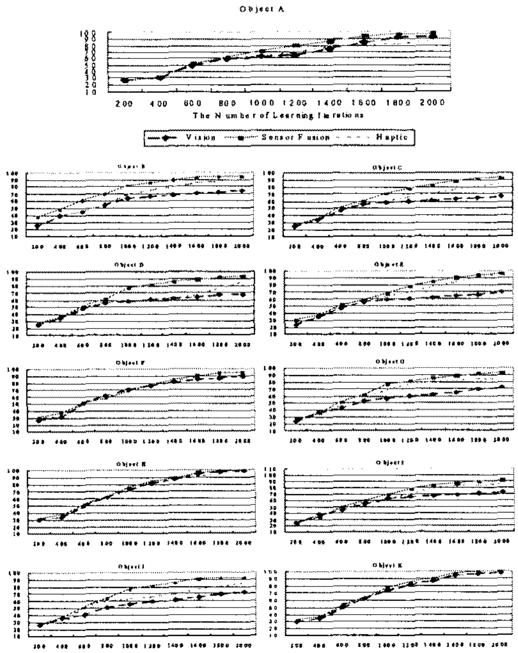

Figure 5. Comparison of the number of learning iterations.

4. Conclusion

The estimated performance indicates that this system can improve accuracy of recognition and speed of learning time with sensor fusion. Recognition using tactile sensors is better than that of single vision sensors for similar-shaped objects. It is anticipated that we will have to study how to recognize the colors of objects and apply to practical objects. And we must also study how to exchange information among different hierarchical levels if the number of hierarchical sensor levels is increased.

References

[1] Y. Sato, S. Tamamura, "3D object recognition using parameterized model", Trans. Inf. Soc. Japan, vol. 31, pp. 56-67, 1990.

[2] S. Yamane, M. Izumi, K. Fukunaga, "Model-based object position and orientation prediction", IEICE Trans. vol. J79-D-Ⅱ, pp. 165-173, 1996.

[3] H. Murase, SK. Nayar, "3D object recognition from appearance", IEICE Trans. vol. J77-D-Ⅱ, pp.2179-2187, 1994.

[4] T. Amano, S. Hiura, A. Yamaguchi, "Eigen space approach for pose detection with range images", IEICE Trans. vol. J80-D-Ⅱ, pp. 1136-1143, 1997.