



KR9700352

*'97 Workshop on Spent Fuel Management Technology,
Nov. 13-14, KAERI, Taejon, Korea*

An Object Recognition Using Structured Light and Neural Networks

Byeong-Gab Kim*, Dong-Gi Kim*, E-Sok Kang*, Ji Sup Yoon**

* Chungnam National University, 305-764, Taejon, Korea, ** KAERI
E-mail: eskang@hanbat.chungnam.ac.kr

Abstract

This paper presents a 3D image processing which uses neural networks to combine a 2D vision camera and a laser slit beam. A laser slit beam from laser source is slitted by a set of cylindrical lenses and the line image of the slit beam on the object is used to estimate the object parameters. The neural networks allow to get the 3D image parameters such as the size, the position and the orientation from the line image without knowing the camera intrinsic parameters.

Keywords - Object recognition, Structured light, Neural networks

1. Introduction

Various sensors have been used for recognizing environments. However, each sensor shows different characteristics in measuring range and accuracy, and in application areas. Ultrasonic sensors have been most widely used[1], because the range map can be acquired directly from the sensor output. However, these sensors have a low resolution due to a wide measurement angle. Stereo vision has been used as another method of object recognition[2]. The stereo vision utilizes disparity between right and left images. This method has a drawback that finding the correspondence of two images requires a computationally expensive procedure.

Structured beam has been used as a reliable object recognition method. An active light source projects a stripe pattern of light onto the object surface, while the camera observes the projected pattern on the object surface. The method has a good accuracy and camera calibrations are quite simple. Yuta et al. have developed an optical range sensor using laser slit beam for indoor mobile robots. Their system has a wide viewing angle of 60° and does not require mechanical scanning[3, 4].

In this paper, an object recognition method using structured light and neural

networks is proposed. The proposed method utilizes line segment interpretation from the image level and the local map is acquired efficiently. The performance of the proposed method is demonstrated in detail through the experiments.

2. System Description

An object recognition system has been constructed, as is shown in Fig. 1. The system consists of a laser slit unit, a CCD camera, a pan/tilt device, a vision processing unit and the computer. The laser unit consists of a semiconductor laser and a cylindrical lens that converts a laser spot into a stripe. To get the widest viewing angle, the laser stripe is projected laterally. The camera is located at an offset distance from the laser unit. To enhance the laser stripe pattern in the acquired image, an optical band-pass filter that passes only the narrow wave-length band of the laser light is used. Both the laser unit and the camera are mounted on the pan/tilt device.

The pan/tilt device is a 2-axis driven sensor scanning device. Panning is defined to be a rotation about the vertical axis, while tilting is defined to be a rotation about the axis perpendicular to the pan axis. The vision processing unit processes 512×512 gray level images. Laser slit beam generated by a laser diode and a cylindrical lens forms a light plane in 3D space. At the intersection of the object surface plane and the light plane, a light stripe is formed which is represented as a line in a camera image. Thus, a line segment in the 2D camera image implies an object surface plane.

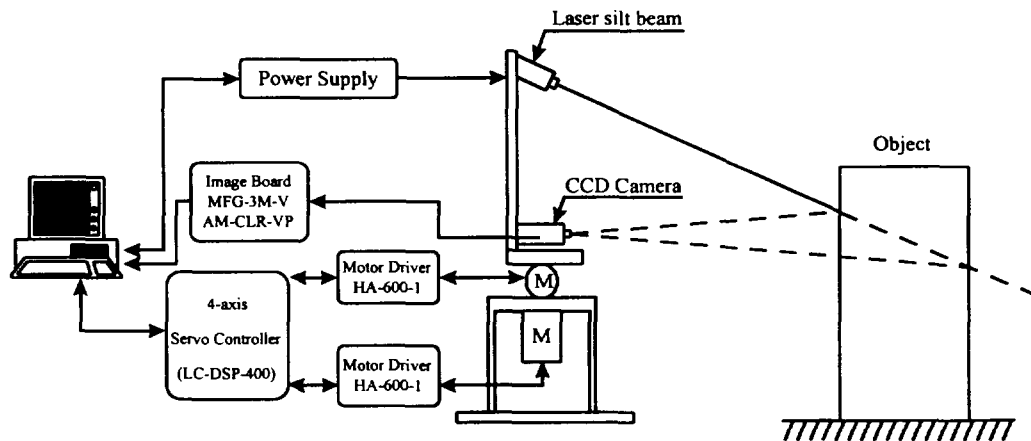


Fig. 1 Schematic Diagram of the Object Recognition System.

3. Object Recognition

Fig. 2 shows the parameters of the cylinder type and the box type models. Let the image plane coordinates be u and v and an image plane point in homogeneous coordinates is (U, V) . Similarly the world point in homogeneous coordinates is (X, Y, Z) . For cylinder type objects, we get the typical lengths, l_1 and l_2 , the top point, $A(U_A, V_A)$, from the image of the CCD camera. We can also get the typical parameters such as inclined angles, α_1 and α_2 , the lengths, l_1 and l_2 , and the top point, $A(U_A, V_A)$ for box type objects. 3D recognition of a cylinder type objects needs the diameter, d , and the center position, (X_C, Y_C, Z_C) and the recognition of a box type object needs the side lengths, l_a and l_b , the orientation angle, θ , and the corner position, (X_B, Y_B, Z_B) . So we need an algorithm which maps the parameters of the image plane onto the parameters of the world plane.

Recent successes in employing neural network models for solving various computationally difficult problems have inspired the research in the area[5]. Numerous researches are reported in neural networks and introductory texts on the subject are commonly available[6]. A neural network typically consists of a relatively large number of computational units connected with a larger number of communication links. In essence, the neural network is a self-adaptive structure which incrementally alters its characteristics corresponding to external conditions until it achieves the desired performance.

In this study, the generalized delta rule for the semilinear feedforward net with back propagation of error is implemented to assimilate the mapping characteristics. The structure of the neural network is shown in Fig. 3. The net input to the j th node in the next layer is computed as

$$net_{pj} = \sum_i w_{ji} o_{pi} \quad (1)$$

where the subscript p denotes the individual training data and w_{ji} denotes the weight between nodes, and o_{pi} is the output of the input layer. When a sigmoidal activation function is used, the corresponding output of a node j is

$$o_{pj} = \frac{1}{1 + e^{-net_p + \theta_j}} \quad (2)$$

where θ_j serves as a threshold or bias. For each pattern, the square of the error is defined as

$$E_p = \frac{1}{2} \sum_k (t_{pk} - O_{pk})^2 \quad (3)$$

In the training phase, the set of weights and biases is iteratively altered in a manner to reduce the error E_p as rapidly as possible. The learning procedure is accomplished in an iterative manner as

$$\Delta w_{ji} (n+1) = \eta \delta_{pj} o_{pi} + \alpha \Delta w_{ji} (n) \quad (4)$$

where η is the learning rate concerned with the speed of convergence of the error and α is the momentum rate to avoid the oscillating phenomena. Through preliminary simulation tests, the learning rate and the momentum

with frame processor board by Imaging Technology Inc. Tilting angle of laser light was set to about 25° from the horizontal axis, considering the visibility of reflections and protection of human eyes nearby.

Fig. 4 and Fig. 5 shows the preprocessing of the images of a cylinder and a rectangular box to calculate input values of the neural network. (a) shows the scaled image which was reduced to 256×240 image from 512×480 CCD image by the pyramid method to reduce the computing time, and (b) shows the image after edge extraction where the threshold value was 30. (c) and (d) show the images after the object extraction and thinning.

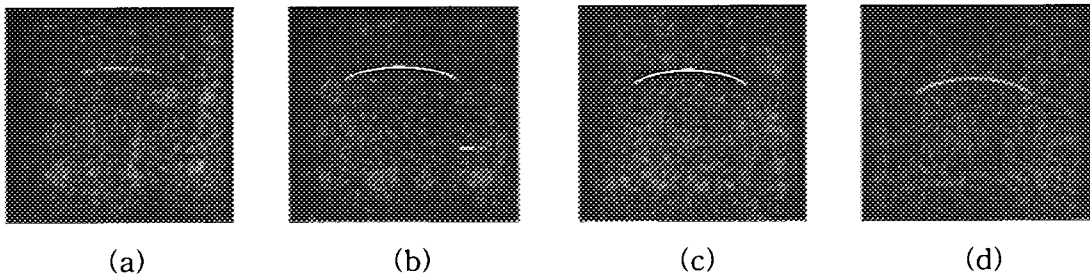


Fig. 4 Procedures of the Image Processing to Calculate Input Values for the Neural Network(for Cylinder Type): (a) Scaled Image by 256×240, (b) Edge Extraction, (c) Object Extraction, (d) Thinning.

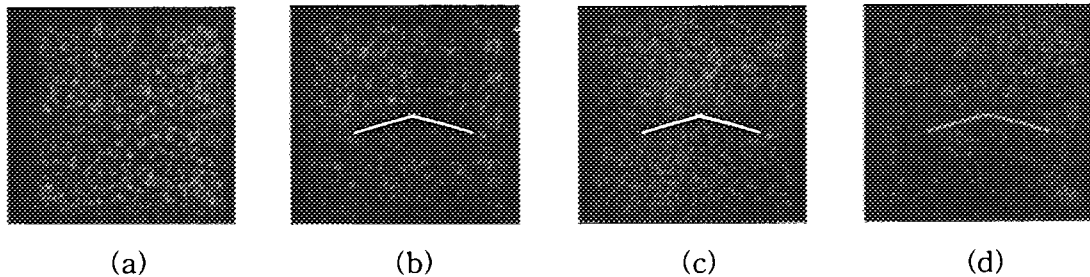


Fig. 5 Procedures of the Image Processing to Calculate Input Values for the Neural Network(for Box Type): (a) Scaled Image by 256×240, (b) Edge Extraction, (c) Object Extraction, (d) Thinning.

Table 1 illustrates the mapping performance of the neural network for cylinder type models, and Table 2 and Table 3 illustrate the mapping performance for box type models, where the deviation was calculated by

$$deviation(\epsilon_x) = \frac{|X_B - X_E|}{\sqrt{X_B^2 + Y_B^2 + Z_B^2}} \times 100 (\%) \quad (5)$$

$$deviation = \frac{|actual\ value - estimated\ value|}{Training\ range\ of\ the\ value} \times 100 (\%) \quad (6)$$

As it can be seen in the tables, deviations of the errors were mostly in the 1% range for the cylinder type models and in the 4% range of actual lengths

Table 1 Evaluations of the Neural Network for Cylinder Type Models.

Actual value (mm)				Estimated value (mm)				deviations (%)			
X_B	Y_B	Z_B	d	X_E	Y_E	Z_E	d_E	ϵ_X	ϵ_Y	ϵ_Z	Δd
1000.0	-40.0	1127.0	290.0	997.5	-42.6	1126.7	290.1	0.16	0.17	0.02	0.01
1100.0	0.0	1100.0	290.0	1083.3	-5.6	1104.3	290.5	1.07	0.36	0.28	0.05
900.0	-50.0	1151.0	290.0	901.7	-55.5	1151.4	288.3	0.12	0.38	0.03	0.19
950.0	-40.0	1140.0	290.0	949.6	-41.0	1139.6	290.0	0.03	0.07	0.03	0.0
1050.0	40.0	1112.0	290.0	1051.2	44.9	1113.0	290.2	0.08	0.32	0.07	0.02
900.0	0.0	1151.0	146.0	902.4	3.8	1151.3	146.4	0.16	0.26	0.02	0.02
1050.0	0.0	1112.0	146.0	1052.1	3.6	1112.9	147.9	0.14	0.24	0.05	0.09
1100.0	-40.0	1100.0	146.0	1088.1	-31.9	1103.3	144.0	0.76	0.52	0.21	0.09
950.0	0.0	1140.0	146.0	943.5	-3.3	1141.0	146.0	0.44	0.22	0.07	0.00
1000.0	80.0	1127.0	146.0	999.2	73.8	1126.3	145.9	0.05	0.41	0.05	0.01

Table 2 Evaluations of the Neural Network for Box Type Models.

Actual value						Estimated value					
X_B (mm)	Y_B (mm)	Z_B (mm)	θ ($^\circ$)	l_a (mm)	l_b (mm)	X_E (mm)	Y_E (mm)	Z_E (mm)	θ_E ($^\circ$)	l_{aE} (mm)	l_{bE} (mm)
1150.0	80.0	1090.0	15.0	230.0	245.0	1139.6	71.3	1092.4	20.1	230.1	245.1
1250.0	40.0	1058.0	75.0	230.0	245.0	1257.2	37.0	1058.4	74.7	230.1	245.2
1350.0	-80.0	1035.0	75.0	230.0	245.0	1344.5	-83.0	1035.5	68.4	229.9	244.9
1150.0	0.0	1090.0	75.0	230.0	245.0	1147.2	-12.0	1090.0	79.8	230.1	245.3
1350.0	40.0	1035.0	30.0	230.0	245.0	1339.0	35.4	1036.9	24.1	230.0	245.1
1150.0	-80.0	1090.0	30.0	205.0	185.0	1160.4	-83.5	1084.8	21.3	205.4	186.0
1350.0	0.0	1035.0	30.0	205.0	185.0	1343.2	4.3	1035.7	35.8	205.0	185.0
1200.0	40.0	1075.0	30.0	205.0	185.0	1204.5	35.5	1071.8	29.3	205.0	185.0
1250.0	80.0	1058.0	60.0	205.0	185.0	1242.0	74.6	1060.7	62.0	205.0	184.9
1300.0	40.0	1048.0	60.0	205.0	185.0	1291.4	37.7	1048.1	64.7	205.0	185.1

Table 3 Errors of the Neural Network Output for Box Type Models.

Deviation (%)					
ϵ_X	ϵ_Y	ϵ_Z	$\Delta \theta$	Δl_a	Δl_b
0.65	0.55	0.15	6.8	0.04	0.04
0.44	0.18	0.02	0.4	0.04	0.08
0.32	0.18	0.03	8.8	0.04	0.04
0.18	0.76	0.0	6.4	0.04	0.12
0.65	0.27	0.11	7.9	0.00	0.04
0.66	0.22	0.33	11.6	0.20	0.05
0.40	0.25	0.04	7.7	0.00	0.00
0.28	0.28	0.20	7.3	0.00	0.00
0.49	0.33	0.16	2.6	0.00	0.05
0.51	0.14	0.0	6.3	0.00	0.05

and in 10% range of actual orientation for the box type models.

5. Conclusions

An object recognition method using a camera and a laser slit beam with neural networks has been presented. The proposed method utilizes the light stripe image on objects, and then constructs a 3D local map. Image enhancement using edge extraction, object extraction and thinning with thresholding is described. The performance of the proposed method has demonstrated in detail through the experiments for various type objects. Experimental results show that the method has a good position accuracy, effectively eliminates optical noises in the image, greatly reduces memory requirement, and also greatly cut down the image processing time for the 3D object recognition compared to the conventional object recognition.

References

- [1] L. Kleeman and R. Kuc, "Mobile Robot Sonar for Target Localization and Classification", *Int'l J. of Robotics Research*, Vol. 14, pp. 295-318, 1995
- [2] K. H. Kim, H. R. Beom and H. S. Cho, "An Obstacle Detection Method for Mobile Robots Using Structured Light" *Proc. ISMCR*, Slovakia, pp. 435-440, June 1995.
- [3] Y. C. Cho and H. S. Cho, "A Stereo Vision-based Obstacle Detection Method for Mobile Robot Navigation", *Robotica*, Vol. 12, pp. 203-216, 1994.
- [4] S. Yuta, S. Suzuki, Y. Saito and S. Iida, "Implementation of an Active Optical Range Sensor Using Laser Slit for In-Door Intelligent Mobile Robot", *Proc. IROS*, Osaka, Japan, pp. 415-420, Nov. 1991.
- [5] A. Busboom and R. J. Schalkoff, "Direct Surface Parameter Estimation Using Structured Light: A Predictor-corrector Based Approach" *Image and Vision Computing*, Vol. 14, pp. 311-321, 1996.
- [6] E. S. Kang and H. S. Cho, *Robotica*, "Vibratory Assembly of Prismatic Parts Using Neural Network-based Positioning Error Estimation", Vol. 13, pp. 185-193, 1995.
- [7] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning Internal Representations by Error Propagation", In: *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*, MIT Press, 1986.