Vision Sensor and Ultrasonic Sensor Fusion Using Neural Network

Sanghoon Baek, and Se-Young Oh

Department of Electronic and Electrical Engineering, Pohang University of Science and Technology, Kyungbuk, Korea
(Tel : +82-54-279-2880; E-mail: tkeb100@postech.ac.kr, syoh@postech.ac.kr)

Abstract: This paper proposes a new method of sensor fusion of an ultrasonic sensor and a vision sensor at the sensor level. In general vision system, the vision system finds edges of objects. And in general ultrasonic system, the ultrasonic system finds absolute distance between robot and object. So, the method integrates data of two different types. The system makes perfect output for robot control in the end. But this paper does not propose only integrating a different kind of data but also fusion information which receives from different kind of sensors. This method has advantages which can simply embody algorithm and can control robot on real time.

Keywords: Ultrasonic sensor, Vision sensor, Sensor fusion, Mobile robot, Neural network

1. INTRODUCTION

Recently, autonomous robots like mobile robot are navigated by a number of methods. These methods use many external sensors for environment recognition. If we can't use these sensors, mobile robots can't navigate in the indoor environment and in the outdoor environment. Theses sensors have a role which the human's sense system does. These can measure the distance between the robot and an object and view where an object is in an environment. We can make various ways because there are various functional sensors. Also autonomous robots are continuously developed because many methods for autonomous robots are created on the base of the sensor's function. Sensors are important to autonomous robots because of these reasons.

So, there are many requirements of sensors for autonomous robot. First, accuracy is important for avoiding collision and fast motion. And sensors have a robust capacity in the various environments [1]. But most of the sensors which we are using in the present time don't have enough capacity. For achieving better environment recognition, using multiple sensors is a one of the effective ways. Using of multiple sensors is a major factor of enabling some measurement of intelligence to be incorporated into their overall operation so that they can interact with and operate in an unstructured environment without the complete control of a human operator [2].

The most popular sensors for mobile robot are vision sensor and ultrasonic sensor. The ultrasonic sensor is simple and gives distance information directly. And it is cheaper than other sensors. But it is known that the ultrasonic sensor has a big problem, difficulty to measure accurate direction [1]. The vision sensor is similar to eyes of human. The vision sensor is easy to cognize an environment. But the vision sensor is difficult to measure distance between object and robot. It can give distance information but this distance information is not accurate. This information includes large errors. This problem is compensated by using stereo vision.

Each of these systems has their own advantage and shortage. And we can utilize each system's an advantage [3,7-8]. Using these two sensors at the same time enhanced abilities of environment perception and collision avoidance. A useful categorization is to consider the fusion of multiple sensors as taking place at the signal, pixel, feature, and symbols levels of representation. The most of the sensors typically which are used in practice provide the data that can be fused at one or more of these levels [9].

2. FUSION OF ULTRASONIC SENSOR AND VISION SENSOR

2.1 Ultrasonic sensor

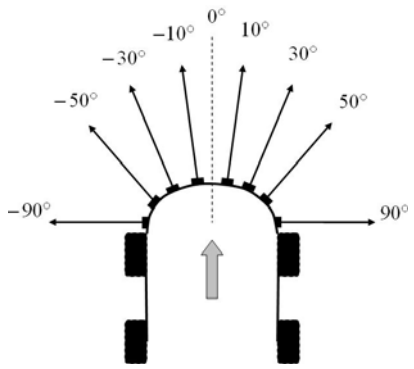
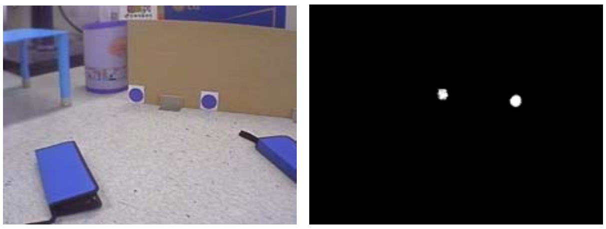


Fig. 1 The ultrasonic sensor arrangement for the Pionner-AT.

Eight ultrasonic sensors are arranged on the front surface of a robot. Ultrasonic sensors are placed symmetry on the center axis of the robot and are at the angle of 10°, 30°, 50° and 90°, respectively(see Fig. 1.).

Ultrasonic waves are reflected to objects of an environment. Ultrasonic sensor receives these waves and measures between the emitting time and the receiving time. These are used to calculate the distance between the robot and objects. Ultrasonic waves have the character which is spreading in a general environment. So the ultrasonic sensor receives the wave which is reflected at the nearest object to the robot. This means that the reflection comes from a point, not from an area, and that the reflecting position is on the surface of the wall which we don't know the accurate shape of the wall [3]. Also this implies possibility that the ultrasonic sensor don't receive the reflected wave. This case was frequently happened in the practical experiment. These data which are gathered when ultrasonic sensors don't receive the wave have an error. These data disturb the learning of neural network. A data filter of ultrasonic sensor is needed to defend this phenomenon. In this paper, we used simple method. Errors of ultrasonic sensor's data is occurred when ultrasonic sensor don't receive the ultrasonic wave. Ultrasonic sensor's data is the maximum distance in the case which the phenomenon is occurred. So the controller to control ultrasonic sensors checks continuously five times. The controller selects a minimum value among the five gathered data. The information has the higher accuracy than other data.

2.2 Vision sensor



(a) (b)

Fig. 2 Image Processing result for detecting the landmark: (a) before the image processing for detecting the landmark, (b) after the image processing for detecting the landmark.

An image which is gotten by a single camera is the same image as you can see when you close a single eye. This means that it is a picture of environment which does not include distance information [1]. We can do a 1D image to world mapping. This method allows the inference of distance information along the projection of the optical axis onto the ground plane [4]. This method can acquire distance information. But these distance information has doubtful accuracy.

In this paper, the role of the vision sensor is finding landmark which consists of two circles. The landmark's color is blue. The information of the vision sensor is the position of each blue circle. We don't use the distance information using the landmark. In this paper, the method for finding the landmark is Connected Component Labeling algorithm. Once a gray level image has been processed to remove noise and thresholded to produce a binary image, a connected components labeling operator can be employed to group the binary-1 pixels into maximal connected regions. The landmark is detected through this process [5, 10]. Fig. 2 shows the result of this process. In this paper, we use 320×240 size image.

The method using Connected Component Labeling algorithm can't find the landmark. In general environment, there are many objects of the blue color. Moreover, objects are various shapes. We must detect only two circle of the blue color. Because these reasons, it must be added rules for detecting the landmark. Firstly, it is a search in region for detection the landmark. It processes an image part of the restricted region on the basis of the vanishing point. Secondly, the landmark's shape is a circle. The ratio of the length and the width of a circle is the same. If the clustered area is not the same the ration of the length and the width, the area is not the landmark. The equation in this rule is given by

$$Shape\ Ratio = \frac{L_y}{L_x} > 0.7 \ \ (L_x \geq L_y) \ \ or$$
$$Shape\ Ratio = \frac{L_x}{L_y} > 0.7 \ \ (L_y > L_x) \qquad (1)$$

Thirdly, the number of pixels of the landmark is restricted. If the number of pixels of the clustered area is between the minimum number of pixels and the maximum number of pixels then it is the landmark.

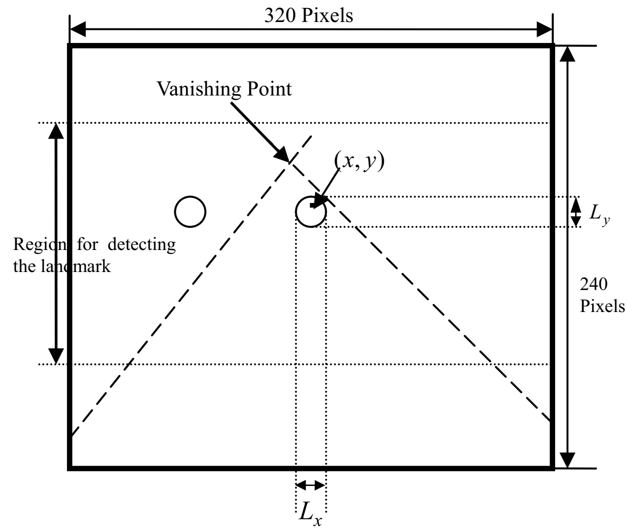


Fig. 3 Acquired information through the image processing

2.3 Fusion

2.3.1 The controller for fusion

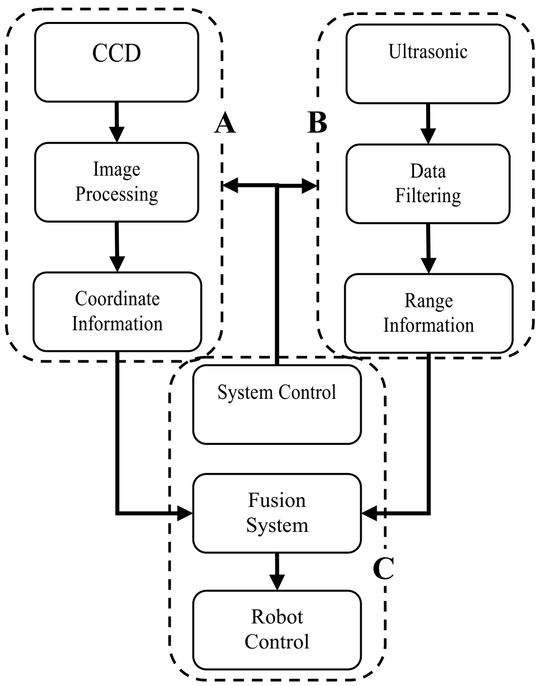


Fig. 4 The overall block diagram of the system for fusion of data.

Fig. 4 shows the overall system for fusion. The controller consists of three parts. The A-processor and the B-processor are independent system each other. The A-processor is inputted information from the vision sensor then provides C-processor with the position information of the landmark. The B-processor is inputted information from the ultrasonic sensor then provides C-processor with the range information of each ultrasonic sensor. C-processor makes the information for robot control of provided information from the A-processor and the B-processor. And the C-processor controls the A-processor and the B-processor. The reason for consisting three parts removes that the time-delay is occurred for providing to C-processor the information of each sensor. The required time for processing information of image is longer than the required time for processing information of ultrasonic sensor. This time difference cause the time-delay. The occurring time-delay is the cause of the result which is involving error.

2.3.2 The Data Fusion using MLP

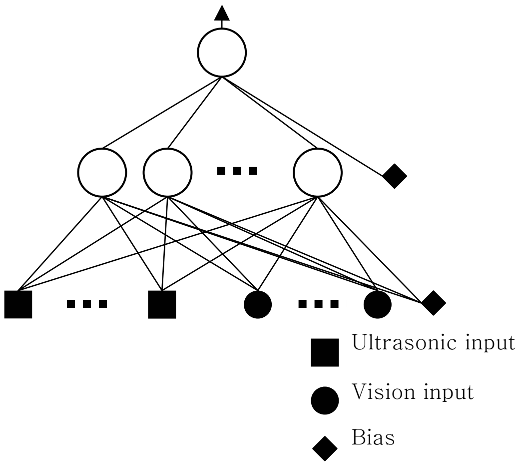


Fig. 5. Structure of MLP

In this paper, we use MLP(Multi Layer Perceptron) for the data fusion of two data which are holding the differential character. Fig. 5 shows a structure of MLP used in this paper. Input data are consisted position information of the landmark and range information of the ultrasonic sensor. And we use history data for a fast convergence. So, the input layer has 37 input nodes (involving 1 bias node). The equation of update is given by

$$E = \frac{1}{2}\sum_{k}(t_k - y_k)^2 \tag{2}$$

$$\Delta w_{jk}(t+1) = -\varepsilon \frac{\partial E}{\partial w_{jk}}(t+1) + \alpha \Delta w_{jk}(t) \tag{3}$$

t_k is k-th target value of neuron in the output layer. y_k is k-th output value of neuron in the output layer. ε is the learning ratio. α is the momentum [6].

3. EXPERIMENT

In this paper, we make the fusion system of using thread form. The structure of the used neural network consists of thirty-six neurons in the input layer, fifteen neurons in the hidden layer and one neuron in the output layer. The reason that the input layer consists of thirty-seven neurons is that we use history data. History data consist of data at current time t, data at the past time t-1 and t-2. Data at each time consist of twelve data. The sum of thirty-six neurons and one bias neuron makes thirty-seven input neurons.

3.1 Result of Using Simulator

The ratio of the real environment and the environment in the simulator is 1 to 1. Fig. 6 shows modeling of the robot used in the simulator. Fig. 7 shows the shape of landmark in the simulator. The filled circle with the blue color is shown by robot and the empty circle is not shown by robot because the obstacle hides the landmark.

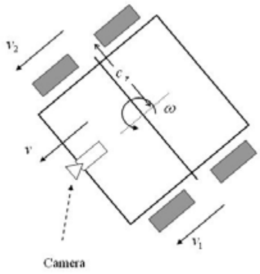


Fig. 6 Modeling of the robot in the simulator

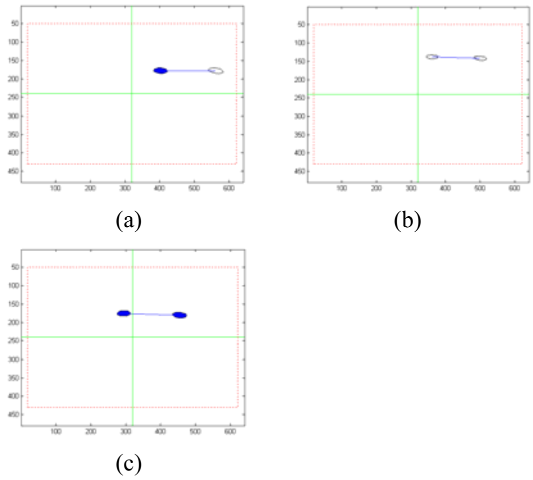


Fig. 7 The ultrasonic sensor arrangement for the Pionner-AT: (a) the right circle is hided by an obstacle, (b) two circles are hided by an obstacle, (c) two circles is showed by the camera of the robot.

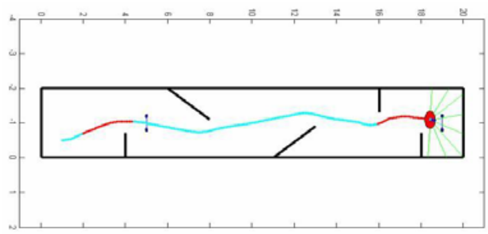


Fig. 8 Simulation of the robot trajectory

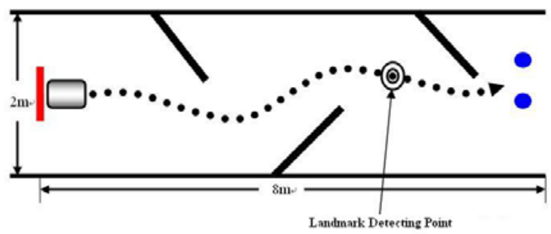
Fig. 8 shows the result of experiment in simulator. Eight lines surrounding robot express the ultrasonic sensor. The landmark is located two places. The location of detected the landmark is approximately 3m~4m before the landmark. In the other place which is not detected the landmark, the robot moves and avoids obstacles like as the robot doesn't have the vision sensor. If the landmark is attached on a narrow place as a door, the robot passes the door without a collision between the robot and the wall.

3.2 Result of Real Experiment

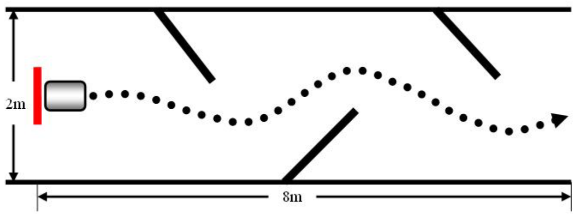
The figure 9 shows Pioneer3-AT, the four-wheel driven mobile robot system that is manufactured by Activemedia Inc. The robot weighs approximately 14kgs standing approximately 24cm in height. A four wheel differential dc motor drive system allows the robot to perform pivotal movement. The control system of Pionner3-AT is comprised of two layers. The high level control is considered to be derived from the host system, a Pentium-IV labtop PC, from only a programming perspective utilizing the Visual C++ language. The PC host system interacts with the robot's various subsystems through RS232 communications.

Fig. 9 Pionner3-AT.

The figure 10 shows the result of experiment in a real environment. The place detected the landmark is 3m before the landmark. The trajectories of the robot are different from this place to the place of the landmark. In case there is the landmark, the trajectory of the robot toward the landmark. But in case there is no the landmark, the robot moves to center of a corridor.



(a)



(b)

Fig. 10 Experimental trajectory of the robot: (a) in case of using the artificial landmark, (b) in case of no using the artificial landmark.

4. SUMMARY AND CONCLUSION

In this paper, the proposed method is easy embodied needed the processing method of each sensor for using various sensors at the same time. It can increase the utility factor because the data fusion method for obtaining the ultimate result is easily embodied. In case of using only the ultrasonic sensor the robot can't pass the narrow space of the similar width of the robot. Using the vision sensor is acquired the better result than using the ultrasonic sensor. If you use only the vision sensor for passing narrow space, the robot can't avoid obstacles that suddenly appear on the front of the robot. In this case, the robot needs a system for avoiding obstacles. At this time, we can use a cheap ultrasonic sensor. As shown by sample trajectories from simulations and experiments, the algorithm is effective in escaping local minima in reasonably complex environments.

A simple and practical algorithm for sensor fusion is described and demonstrated with simulation and experimental result. And a neural architecture has been proposed for the control of a mobile robot that has the ultrasonic sensor and the vision sensor. The main contribution of the algorithm lies in its implementation of a simple fusion of sensory data. The system permits the robot to navigate with the real time method in the environment.

The proposed method can't predict the position of the landmark. If the robot miss the landmark, again, detecting the landmark will be difficult. Future work will be focused on the system for prediction about the position of the landmark.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of Korea toward the ECE Division at POSTECH through the Brain Korea 21 program and also in part by the Ministry of Science and Technology for its Brain Science and Engineering Research Program.

REFERENCES

[1] Teruko YATA, Akihisa OHYA, Shinichi YUTA, "Fusion of Omni-directional Sonar and Omni-directional Vision for Environment Recognition of Mobile Robots," *Proceedings of the 2000 IEEE International Conference on Robotics & Automation San Francisco, CA*, April 2000.

[2] Ren C. Luo and Michale G. Kay, "A Tutorial on Multisensor Integration and Fusion," *Industrial Electronics Society, 1990. IECON '90., 16th Annual Conference of IEEE* , 27-30 Nov. 1990 Pages:707 - 722 vol.1, 1990.

[3] M.A. Abidi and R. C. Gonzales, "Data Fusion in Robotics and Machine Intelligence," *Eds. San Diego, CA: Academic*,1992.

[4] Thomas Bücher, Frank Joublin. "A practicalble approach to 1D Image to World Mapping," *Internal Report 99-01. 1999 Institut the Neuroinfomation. Ruhr-University Bochum,* 1999.

[5] Di Stefano, L.; Bulgarelli, A. "A simple and efficient connected components labeling algorithm," *Image Analysis and Processing, 1999. Proceedings. International Conference on , Page(s): 322 -327, 27-29* Sept. 1999.

[6] Simon Haykin. "Neural Networks," *Prentice Hall Press*, 1999.

[7] Terence chek hion Heng, Yoshinori Kuno, and Yoshiaki Shirai. "Active Sensor Fusion for Collision Avoidance" *Proceedings of the 1997 IEEE IROS 97'*, 1997.

[8] S. G. Goodridge, "Multimedia Sensor Fusion for Intelligent Camera Control and Human-Computer Interaction." *Ph.D thesis, North Carolina State University, Raleigh,* 1997.

[9] Tse Min Chen and Ren C. Luo. "Multilevel Multiagent Based Team Decision Fusion for Mobile Robot Behavior Control". *Proceedings of the 3rd World Congress on Intelligent Control and Automation*, June28-July2, 2000.

[10] Rafael C. Gonzalez, Richard E. Woods. "Digital Image Processing" *Prentice Hall Press*, 2002