Proceedings of 2008 CACS International Automatic Control Conference
National Cheng Kung University, Tainan, Taiwan, Nov. 21-23, 2008

Design and Implementation of an Intelligent Interactive Home Robot

Ming-Hung Tsai, Ming-Yuan Shieh*, Cheng-Ming Lu and Jian-Sheng Chen

Department of Electrical Engineering,
Southern Taiwan University
Tainan County, Taiwan
**e-mail: myshieh@mail.stut.edu.tw*

***Abstract*—This paper presents the design and implementation of an intelligent interactive home robot based on multi-sensory system. In order to accomplish basic home services, the robot is designed to have the functions of environmental map construction, autonomous navigation, collision-free control, and human-robot interaction. For environmental sensing and object tracking, a series of image features recognition procedures are used to establish the database of objective images. By integrating image information of objectives and ground lines, the robot can determine the relative localities of specific objectives and then perform environmental map construction and robot localization. Moreover, the user can also control the robot remotely to navigate and construct the environmental map through the interaction website. Besides, in order to let the robot adapt to various environments, there are fuzzy long and short range obstacle avoidance control systems designed in the robot to avoid collision and/or downfall whenever it performs navigation or tracking. By combining these sensory and image recognition functions, the proposed robot really has the rudiments of vision, multi-sensing and interactive integration. The experimental results demonstrate the environmental adaptability of the robot; these also show such kind of robot can provide plentiful home entertainments.**

***Keywords*—Multi-sensory system, Navigation, Image recognition, Home robot, Human-robot interaction.**

I. INTRODUCTION

Following rapid advances in human life, to develop suitable and powerful robots and related products becomes more essential and emerging. Generally speaking, assistant or service robots are the most conspicuous of autonomous robots. Such robots are unlike industrial robots and usually do not emphasize the accuracy and performance of operation but always focus on giving assistance and/or providing comfortable services. Especially for children and elders, there are many real robots developed for home nursing during these years. For example, the robot "ifbot" [1, 2] which developed by Japan industry-university joint research project, can communicate with humans by joyful conversation and emotional facial expression. Such kind of robots, which always are designed in automatic or semi-automatic way, have their cute appearance, can follow, guide, perceive, recognize, avoid obstacles, and interact with people or environment. Hashima *et al.* who tried to design an assistant robot replacing manpower for elder care in hospital or the elder's home [3].

To design a robotic system to realize correctly what meanings about a series of voice sounds is a hard task. It is because that any language has its distinguishing characteristics, moreover, it is not algorithmic in comparsion with computer program. Especially in Mandarin, there are vague meanings existed in daily dialogue. Yang and Wu proposed a novel understanding approach named by Semantic Dependency Analysis (SDA) [4] to find the implicit semantic dependency between the concepts. Instead of semantic frame/slot, SDA can keep more information when the system cannot clearly identify speech act or intention.

In the relevant research of the autonomous robot, how to detect obstacles is also an important key. In general, there are several kinds of methods used to detect environmental objects such as ultrasonic wave sensing [5], laser scanning [6] and infrared ray sensing [7], etc. However, these detecting techniques all have their drawbacks, for example, the ultrasonic wave sensing is limited to repeated reflection in facing a corner; the laser scanning needs to scan all aspects recurrently; and the infrared ray sensing has a short and narrow searching range. In order to solve these problems, most researches combine the visual servo method with the above techniques to perform navigation controls.

Ohya *et al.* fused the continuous image array captured from a single camera with ultrasonic wave sensing signals to detect obstacles [8]. Meng *et al.* proposed a neural network based visual navigation control method [9], in which, the distinct identities of the building were adopted as the guiding signs to set up nonmetrical environmental models. They had shown that the robot not only could follow the road or the corridor bust also could distinguish what intersection it is staying at. In [10], Han *et al.* designed a tracking system based on the three-dimensional vision of dual cameras. In this paper, a fusion of ultrasonic wave sensors, a laser scanner, infrared ray sensors and a vision system is proposed to make sure the robot be autonomous, guidable, collision-free, and safe whenever and wherever it works in home.

In this paper, the authors also focus on the problem of the human-robot interaction. It needs to consider those factors such as environmental data, speech recognition, perceptive and understanding. How to design auditory system to determine human speech from environmental sounds in the implementation is the first problem. How to adopt suitable algorithms to recognize the speech and then to conclude the reflective actions will be the next problem. Although, such schemes combined with those algorithms have compensatory effects, the understanding system is still obscure because of the complex of human language. To solve such problem, the authors only design the robot to focus on those cases that are used frequently in daily conversations.

II. THE SYSTEM CONFIGURATION

The proposed interactive home robot is named as "Fairy II" whose realized appearance are shown in Fig. 1. The fundamental system structure is shown in Fig. 2. The mechanism consists of the robotic skeleton, a personal computer, visual devices, acoustic modules, wireless networks, sensory modules, DC motors, drivers, DC/AC converters, batteries, FPGA modules and related modules of feature display. The vision system of the robot is based on a webcamear and an image proceeding interface set up by Microsoft windows programs. Fig. 3 shows the procedures of image processing and visual servoing. The obstacles and objectives will be recognized and determined for robotic controls of obstacle avoidance, interactive following, or interactive feature appearance by these visual processing procedures from captured images.

The control system of DC servo motors and drivers is the FPGA starter kit developed by Altera Corp. There are powerful application interface softwares provided by Altera Corp. for users performing VHDL programming and debug. It is easy to construct a MAX II embedded system for controlling the servo motors fastly by the develop tool -- Quartus II. Remote monitoring functions of Fairy are realized by internet, wireless transmitted system, and PC based interface programs developed in Visual Basic. The robot Fairy is allowed to be remote controlled by predefined users via internet. It means that the master can not only operate the robot but also observe the eyesight of the robot from the remote control system via internet.

In this paper, as shown in Fig. 4, a fuzzy speaker localization system based on a semi-spherical microphone array is mounted on the robotic head. In the system, there consist of 12 capacitor microphones disposed in two layers on the semi-sphere of 19 cm diameter. In which, the upper layer disposes four microphones at every 90 degrees, and the lower layer disposes eight ones at every 45 degrees. The hardware can be mounted upon the head of the robot to determine the 3D relative locality relationship between the sound and the robot. Since the disposal could receive all the sounds come from any direction, it provides omni-bearing detection to determine the exact localization of the signal. Based on this data, the robot can perform face-to-face conversations, direct interaction, or/and objective searching.

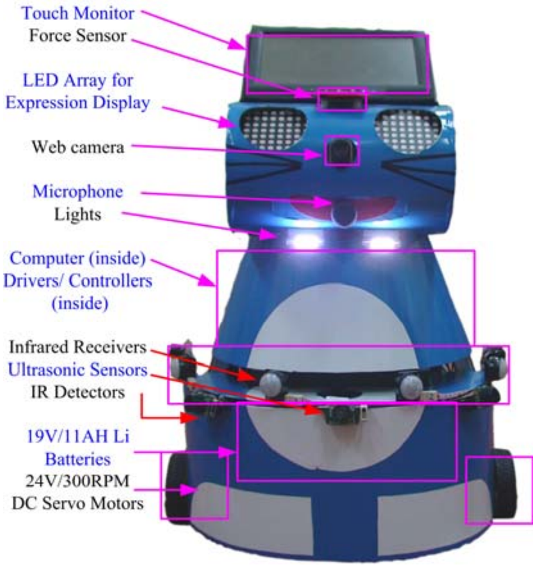


Fig. 1 The appearances of the "Fairy-II" Robot

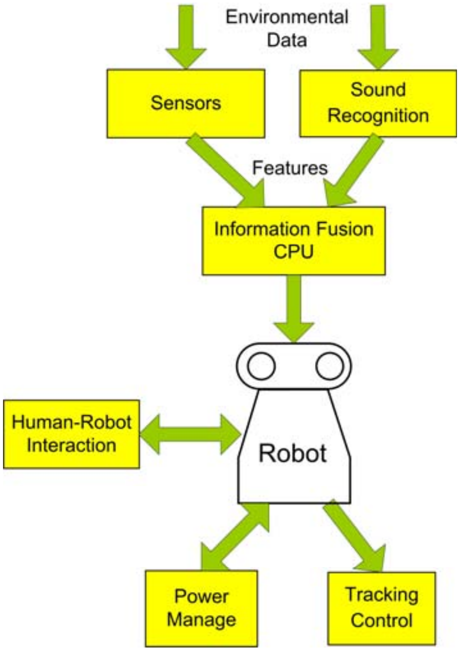


Fig. 2 The system structure of the "Fairy-II" Robot

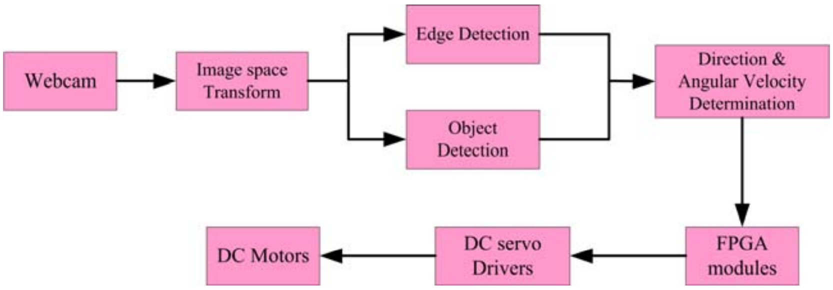


Fig. 3 The procedures of image processing

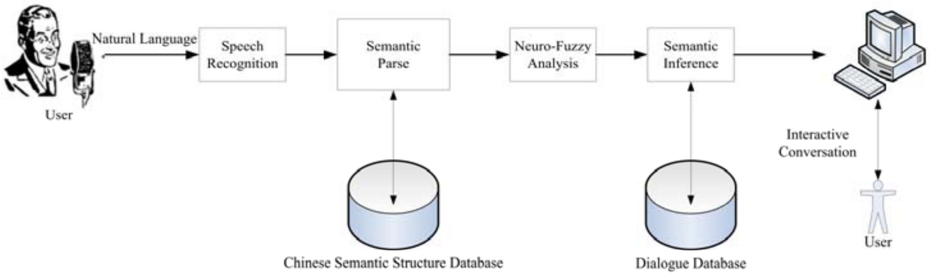


Fig. 5 The diagram of natural language and semantic analysis

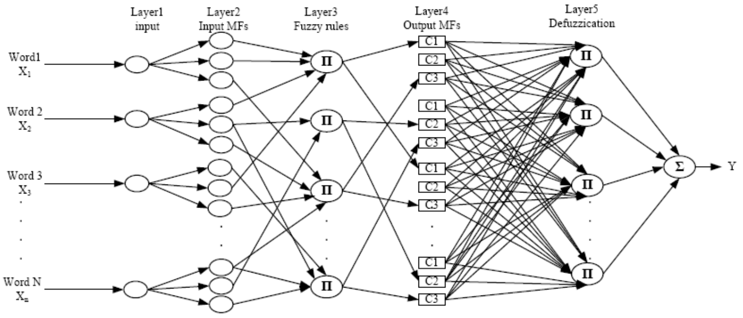
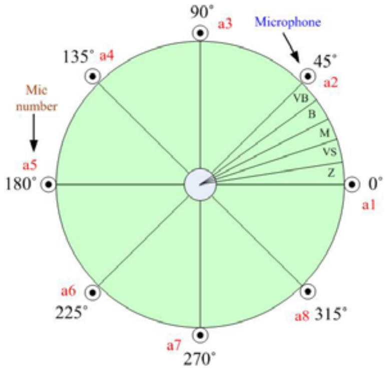


Fig. 6 The neuro-fuzzy network of semantic analysis

(a) Side view



(b) the disposal of microphones in the lower layer

Fig. 4 The configuration of the semi-spherical microphone array

III. SIGNAL PROCESSING PROCEDURES

A. Natural language processing

Natural language is the most important medium for communication in human life. During the high information development years, how to design a natural language based human-computer conversation system has become the main objectives for many researchers. This paper proposed a human-robot conversation system based on Neuro-Fuzzy Networks which allows people talking with the robot in natural language.

The VCMM phonic control module provides high speed data acquisition for speech recognition in the robot Fairy. The robot can display appropriate features or motions to reply what it realized from the dialogue. Fig. 5 illustrates how the whole acoustic system works. The system firstly performs sentence breaks and analyses for natural speech, which separates a sentence into few categories, and these are then used to be the inputs of the Neuro-Fuzzy Network. After the Neuro-Fuzzy Network inferring and learning, the nearest meanings of words will be concluded by comparing the keywords with the vocabulary and semantic database.

Fig. 6 shows the structure of the Neuro-Fuzzy Network (NFN) which consists of five layers. The inputs of the first layer are the characteristic values of semantic words. Each node only transmits one input to the next layer. Note that there is no computation of values or weights in this layer. The next layer is fuzzification layer, its jth input will be transferred to the ith membership value of V signal, in fuzzy domain by Gaussian membership functions. The third layer performs fuzzy rules of the proposed NFN, in which, the “Fuzzy AND” operation is chosen as the algebraic multiply.

The fourth layer aims to accomplish the linear determination of fuzzy numbers by using the technologies of α-cuts. The next (fifth) layer will produce the output of semantic number, which mainly performs related defuzzification operations.

B. Vision-based navigation control

The proposed navigation control system basically contains two kinds of controllers. The first is the controller of objective following. In some statues, a family member (child or elder) might need an assistant or secretary following with him to give some helps or accompanying. Based on this reason, a fuzzy following controller is proposed and aims to determine how to track some characteristics of the objective customer like shoes, legs, face, or specified color blocks on clothes. In this paper, the authors call the image component of a characteristic as an “eigenblock.” The another is the controller of obstacle avoidance. As well known, to track objects and to avoid obstacle both need to recognize the objective blocks, but they perform the opposite actions. The former aims to keep the objective in sight as possible; however, the latter always tries to let the object (the obstacle) out of sight.

In the control mode, the guiding control is generated after the obstacle avoidance controls. Through the image processing procedure of detecting obstacles in warning area, the system will produce collision-free steering controls of the robotic movement. As the robot has successfully evaded those obstacles appeared in the warning area, it just could perform the next operation–guiding or following, otherwise it must focus on obstacle avoidance.

The fuzzy guiding controller takes consideration to two input signals: one is the error between the slope of the left sideline and the reference one, *ΔS*; and another is the difference of x-coordinate value between the vanishing point and the image center, *ΔX*. These values can be defined as

$$\begin{cases} \Delta S = S - S_{ref} \\ \Delta X = X_{V_P} - X_{center} \end{cases} \tag{1}$$

where S denotes the slope of the left sideline; S_{ref} is the reference slope which is the slope of the left sideline when the robot stays on the center line of the passage; and ΔS represents the error of these two slopes. Further, in Eq. (1), X_{V_P} denotes the x-coordinate value of the vanishing point V_P; X_{center} is the x-coordinate value of the image center; and ΔX represents the difference between them.

Since the goal is to promise a collision-free guiding control, the most simple way is to evade obstacles firstly and then to guide the robot toward the center line of the passage. From the views of vanishing point and vanishing lines, we have when the vanishing point is appeared in the left part of the image (*ΔX is NB or NS*) just because the robot is facing the right side. Moreover, if the robot is also close to the right sideline, the slope of the right sideline will be smaller than the reference value of S_r (that is *ΔS is NB or NS*), and then the steering command should be set as "to turn left large or little" to let the robot return to the center region of the passage. It results in $"\omega_L$ is VS, ω_R is $VB"$ or $"\omega_L$ is S, ω_R is $B."$ Similar inferences can get the fuzzy rules as tabulated in Table 1. Here, it is noted that the reference slope S_r is equal to 1 because that the horizontal angle of sidelines are both 45 degree as the robot stays right at the center line.

Table 1. The Rule Table of the Fuzzy Collision-Free Guiding Controller

ω_L,ω_R		ΔX									
		NB		*NS*		*ZE*		*PS*		*PB*	
ΔS	*PB*	*M*	*M*	*F*	*M*	*F*	*S*	*VF*	*S*	*VF*	*VS*
	PS	*M*	*F*	*M*	*M*	*F*	*M*	*F*	*S*	*VF*	*S*
	ZE	*S*	*F*	*M*	*F*	*M*	*M*	*F*	*M*	*F*	*S*
	NS	*S*	*VF*	*S*	*F*	*M*	*F*	*M*	*M*	*F*	*M*
	NB	*VS*	*VF*	*S*	*VF*	*S*	*F*	*M*	*F*	*M*	*M*

PB = Positive Big, PS = Positive Small, ZE = Zero, NS = Negative Small, NB = Negative Big, M = Medium speed, F = Fast, S = Slow, VF = Very Fast, VS = Very Slow

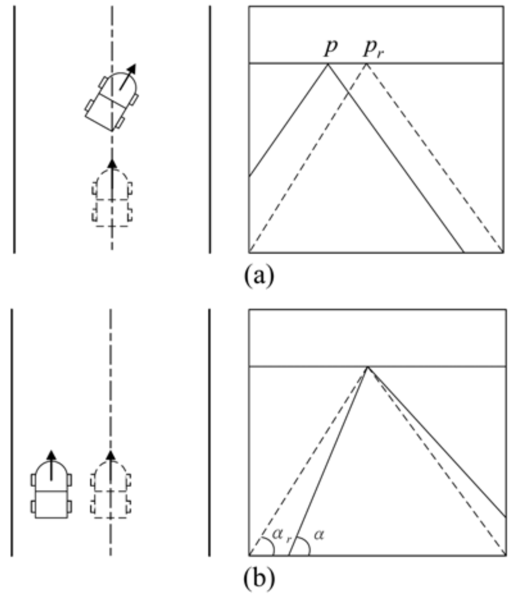


Fig. 7 Geometric relationship of vanishing point and lines

C. Robotic Posture Estimation

From the image of a hallway scene, it is easily seen there are various lines and segments. Since a family of parallel lines on the horizontal plane converges to a vanishing point in the image coordinate under perspective projection. We paint the parallel lines on the sides of a hallway in supermarket as the guiding lines. Fig. 7(a) shows that the vanishing point p_r denotes the image is captured when the robot (the camera) is located on the hallway's center line and towards forward. Another vanishing point p is the convergence of the same parallel lines. Though, the lateral position of robot is still the same (on the center line). Actually, it has turn towards right side.

Fig. 7(b) describes that the slopes of the vanishing lines in the image coordinate vary as the robot stays at different lateral positions. From the figure, it is seen that the slope of the current vanishing line (left solid line) is larger than the slope of the reference vanishing line (left dashed line). It is because that the robot has varied its lateral position from zero (center line, forward) to negative (the left side of center line, forward). It means that the lateral position of the vanishing point can indicate the steering angle of the robot. On the other hand, the slope of the vanishing line is increased with moving left in the lateral position. Since the goal of the guiding control is to let the robot travel straightly forward and keep on the center line of hallway.

IV. EXPERIMENTAL RESULTS

Navigation is always an important topic in robotic mobile controls. Especially, the complexity of the environment where home robots are patrolling will make the robotic design hard. However, if we have solved similar but simpler cases, we maybe can obtain some useful ideas to solve such complex problems. For example, to consider the actual environments such as the passage shown in Fig. 8(a) and the laboratory shown in Fig. 8(b), one can see that both the sides of a passage can be regarded as lines. That is, if the robot can detect the distances from the sides by processing captured images, moreover, also can estimate the space features of the robot with respect to the environment, we believe the robot can perform simple navigations based on algorithms of wall following or side tracking in the passage without obstacles.

Fig. 9 shows the trace of the actual navigation of the experiment as shown in Fig. 10. It indicates that the robot really can pass the obstacles and successfully reach the destination. Besides, Fig. 11(a) and Fig. 11(b) show the actual screens of experiments when the robot stays on two different positions. Fig. 11(a) indicates the robot is at the position 49.68m with respect to its left wall where the distance between the robot and a fire extinguisher is estimated by the vision-based navigation controller as 2.89m because that the pixels of the fire extinguisher's image are 14*15. It also can determine that the robot is at (50,211) with 0 degrees of angle of orientation from the robotic posture estimation.

Fig. 11(b) indicates the robot is at the position 77m with respect to its right wall where the distance between the robot and a fire extinguisher is estimated as 1.47m because that the pixels of the fire extinguisher's image are 24*90. It also can determine that the robot is at (106,144) with 188 degrees of angle of orientation from the robotic posture estimation.

V. CONCLUSIONS

In this paper, a interactive home robot is implemented in an actual home environment. A neuro-fuzzy semantic reasoning system is proposed to deal with interactive conversations between people and robot. To guide, to follow or to avoid obstacles depends on the vision-based navigation controller and the robotic posture estimation system. The experimental results reveal the feasibility of the proposed robotic system.

ACKNOWLEDGMENT

This work was supported by the National Science Council, Taiwan, under grant number NSC96-2221-E-218-039-MY3.

REFERENCES

[1] S. Kato, S. Ohshiro, H. Itoh and K. Kimura, "Development of a communication robot Ifbot," *Proc. of 2004IEEE International Conference on Robotics and Automation, ICRA '04.* 2004, Volume 1, pp.697 - 702.

[2] M. Kanoh, S. Kato and H. Itoh, **"**Facial expressions using emotional space in sensitivity communication robot "Ifbot"**,**" *Proc. of 2004 International Conference on Intelligent Robots and Systems, IROS 2004,* Volume 2, 28 Sept.-2 Oct. 2004, pp.1586 – 1591.

[3] M. Hashima, F. Hasegawa, S. Kanda, T. Maruyama and T. Uchiyama, "Localization and obstacle detection for robots for carrying food Trays," *Proc. of IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Vol. 1, Sept. 1997, pp. 345-351.

[4] M.-Z., Yang and C.-H. Wu, "Semantic dependency based natural language understanding in a medical dialogue system," (in Chinese) the master thesis of NCKU, Taiwan, 2003.

[5] J. Borenstein and Y. Koren, "Obstacle avoidance with ultrasonic sensors," *IEEE Trans. Robotics and Automation*, Vol. 4, Issue 2, Apr. 1988, pp. 213-218.

[6] M. G. Plaza, S. S. Prieto, O. G. Davila, O. G. Poblacion and D. M. Luna, "An obstacle detector system based on laser technology," *IEEE Circuits and Devices Magazine,* Vol 17, Issue 5, Sept. 2001, pp.9-15.

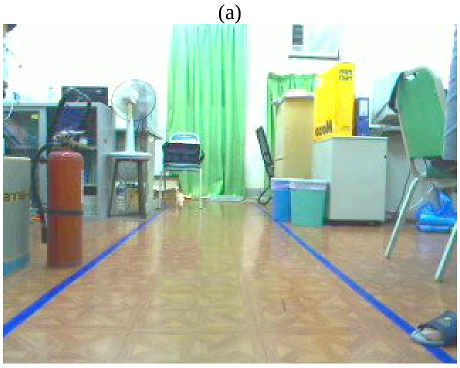
[7] T.-H. S. Li, S.-J. Chang and W. Tong, "Fuzzy target tracking control of autonomous mobile robots by using infrared sensors," *IEEE Trans. Fuzzy Systems,* Vol 12, Issue 4, Aug. 2004, pp. 491-501.

[8] I. Ohya, A. Kosaka and A. Kak, "Vision-based navigation by a mobile robot with obstacle avoidance using single-camera vision and ultrasonic sensing," *IEEE Trans. Robotics and Automation.*, Vol. 14, Issue 6, Dec. 1998, pp. 969-978.

[9] M. Meng and A.C. Kak, "Mobile Robot Navigation Using Neural Networks and Nonmetrical Environment Models," *IEEE Control System Magazine,* vol. 13, Issue 5, Oct. 1993, pp. 30-39.

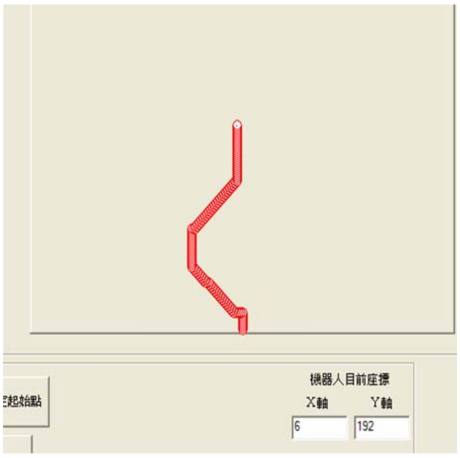
[10] S.-H. Han, W. H. Seo, K. S. Yoon and M.-H. Lee, "Real-Time control of an industrial robot using image-based visual servoing," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Oct. 1999, pp. 1762-1796.

(a)



(b)

Fig. 8 The actual environments of experiments



機器人目前座標

X軸 6

Y軸 192

Fig. 9 The actual navigation results

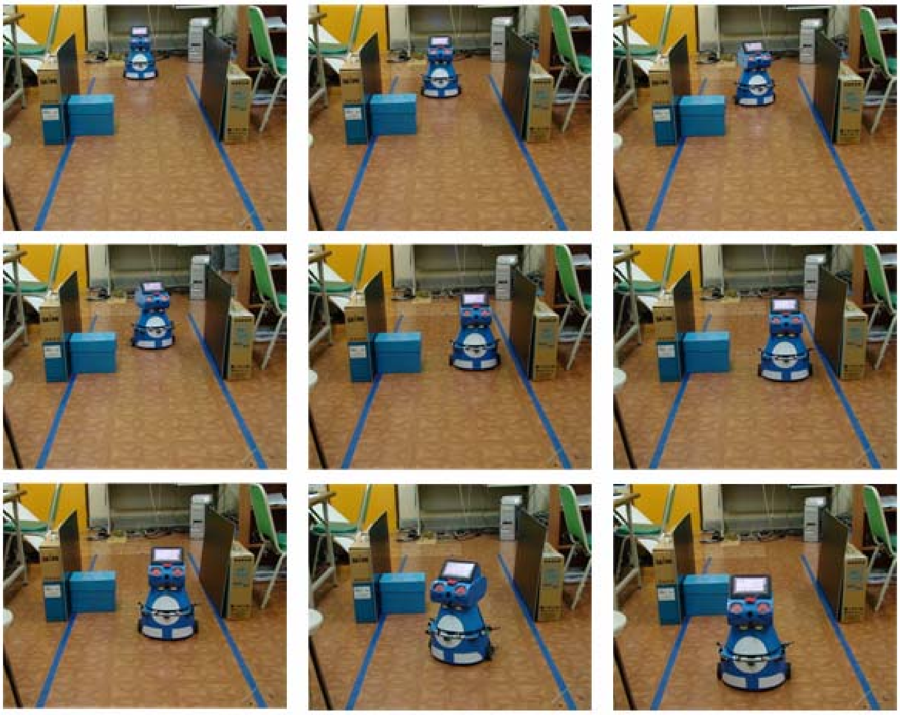
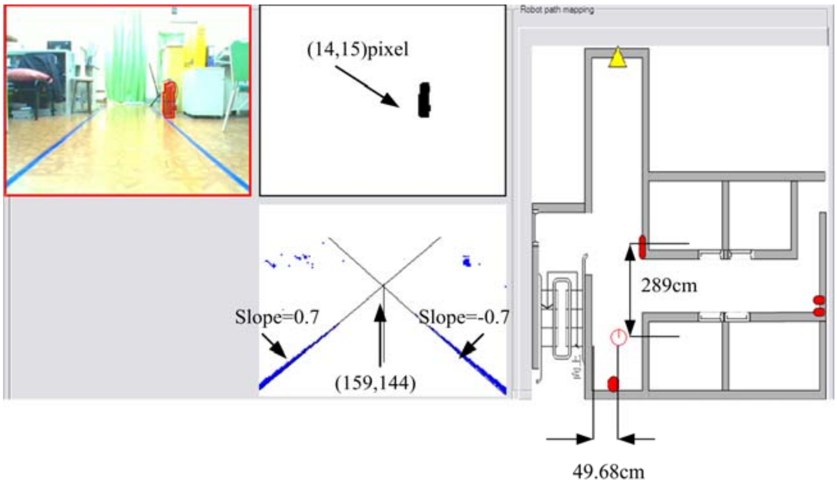
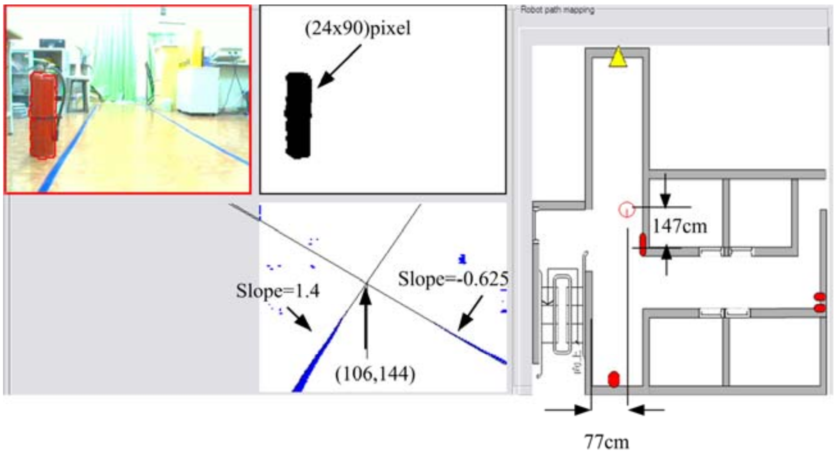


Fig. 10 The successive images when the robot navigates collision-free in the passage.



(a) The robotic monitor interface when the robot faces to the triangle mark.



(b) The robotic monitor interface when the robot is approaching to the fire extinguisher.

Fig. 11 The position estimation results