3rd International Conference on Electrical & Computer Engineering
ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh

VISUAL SENSING APPLICATIONS IN INTELLIGENT ROBOTIC SYSTEMS

Seiji Ishikawa

Department of Mechanical & Control Engineering, Kyushu Institute of Technology
Sensuicho 1-1, Tobata, Kitakyushu 804-8550, Japan
E-mail: ishikawa@cntl.kyutech.ac.jp

ABSTRACT

Intelligent machines or robots have been developing continually these days. In this paper, two novel robotic systems are presented that may contribute to supporting a human life in respective ways. A multiple robots system performs understanding a user's request by visually sensing his/her action and carrying the requested object by cooperation of an independent camera, a manipulator and a mobile robot. This system is intended for assisting a patient in bed or a handicapped person. An aerial robot system, on the other hand, finds a specified human on the ground by visual sensing from the air and tracks the person or it comes close to a human who gives a signal by hands to the aerial robot. The system can be employed for the purpose of security or rescue. Overviews of the systems are given with some experimental results.

1. A MULTIPLE ROBOT SYSTEM ASSISTING A PATIENT

1.1 Introduction

Mobile robots have been developed initially for industrial use. Recently they have become much more popular than ever by the request of possible employment in various fields. For example, a walking robot with legs like a human or an animal, a rescue robot which can work in dangerous zones such as disaster areas, an exploring robot in space, a cleaning robot in an office, *etc*., are eagerly requested robots in respective areas. A mobile robot that supports people's life is also requested increasingly due to the forthcoming aging and mature society.

The necessary functions such a mobile robot should be equipped with include an autonomous nature by which it can recognize its environments, judge what to do next, and make an action by itself. Moreover, if one intends to employ it for the purpose of welfare, communication is indispensable between a human and a robot. For this purpose, the mobile robot should act based on simple instructions by a human. It is therefore important to design a man-machine interface that realizes such simple and easy instructions to robots [1].

In this paper, we propose a multiple robots system for assisting those who need help in taking objects around him/her. Man-machine communication is realized by the employment of visual information processing.

1.2 Overview of the system

Suppose that a person is lying in bed and wishes to take an object on a distant table without getting up. A robotic function expected in this situation is transport of the requested object. Therefore, a mobile robot is needed. One may think of making a single mobile robot having high performance to achieve this kind of service. It should include a function of finding out a requested object (FIND), a function of transporting the object (TRANSPORT), and passing the object to the requested person (PASS). If these three functions are integrated into a single mobile robot, it is not efficient, however, since, when one function is employed, the other functions are not used. When FIND runs, TRANSPORT and PASS are on standby, for example. This fact leads to the employment of a multiple robots system. In the system, the services are done through independent work by multiple robots. In this situation, when TRANSPORT runs, FIND can devote itself into another service for another person.

Since each robot in the proposed multiple robots system has a sensor and a PC so as to behave autonomously, it is expressed as an agent in the following contexts.

The proposed system is composed of three main agents. The first agent seizes information in the work region. This agent is called a Vision Agent

(VA). The second agent captures and transports the object in the distant place. It is called a Robot Agent (RA). The third agent passes the requested object to the user. This is a Pass Agent (PA). In addition, the user is considered to be an agent. We call this agent a Human Agent (HA). Information is exchanged according to the TCP/IP communication among these agents.

A VA and a RA can proceed to the next service while a PA concentrates on its present job in the proposed system. In this way, the agents realize mutual independ- ence [2]. **Figure 1** shows the system configuration.

A VA is composed of a video camera and a PC. The VA has the function to seize information in the work space. VA's role is object recognition, detection of direction of the tip of a finger, and selection of a requested object. A RA is composed of a video camera, a mobile part and a PC. The RA has the function to capture and transport the object. The mobile robot is shown in **Fig.2a**. A PA is composed of a video camera, a manipulator and a PC. The PA has the function to capture and pass the object to a HA. **Figure 2b** shows the PA containing a manipulator.

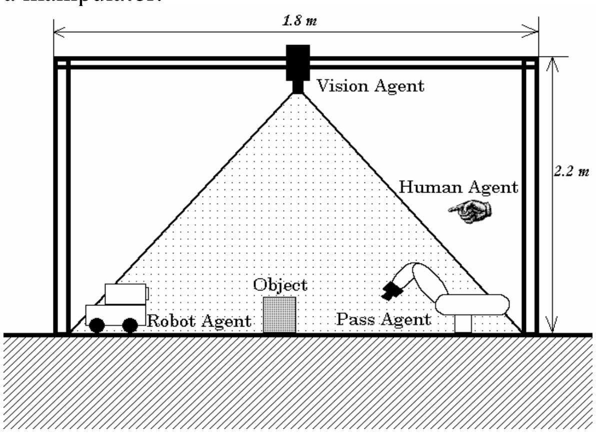
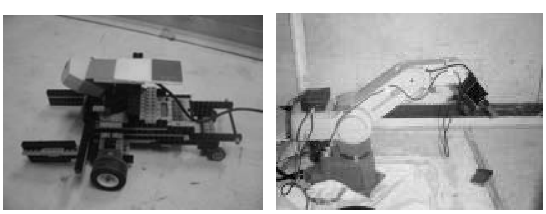


Fig. 1 System configuration.



(a) **(b)**

Fig. 2 A mobile robot and a PA containing a manipulator.

1.3 Procedure of the system

The system is meant to assist those people who find difficulty in taking a distant object, as they lie in bed, for example. Suppose a patient (a HA) in bed wishes to take an object on the table which is a little distant from him/her. If HA points to the object by his arm, hand and finger, a VA on the ceiling observes the direction of the finger and regards the object which is the closest to the line of the direction as the one HA wishes to take. This is of course not always exact and the system must make certain if it has found a correct one.

The VA informs RA the object's position and RA moves and comes to the object. If VA still finds HA's hand within its sight, as it means Yes, the RA gets the object and carries it to PA. If VA doesn't find HA's hand in the sight, as it signifies No, VA informs RA the position of the second candidate and RA comes to it. VA again looks for HA's hand. This is repeated until the system finally gets the right object.

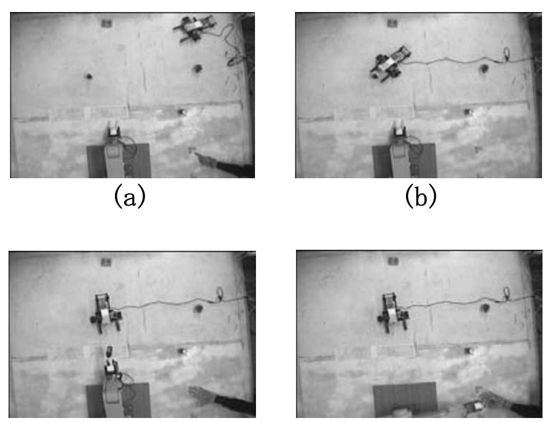
PA receives the object and passes it to HA. When this is performed, the system can proceed to accepting the next request from other patient. In this way, independent processing is realized in the proposed system.

1.4 Experimental results

The experiment was performed in a laboratory environment. The area VA covers is about a 1.8m×1.8m floor on which some objects are placed and RA moves. In the experiment, four objects are used:

They are two small bottles, a cigarette box and a doll. Performance of the system is shown in **Fig.3** by 4 successive photos.

In the figure, HA decides a particular object from



(a) (b) (c) (d)

Fig. 3 Photos of the results.

the four objects. (a) When the instruction enters from HA, VA detects the direction of the tip of a finger of HA, and selects the requested object. Afterwards, RA moves to the requested object. (b) Because the judgment result is "Yes" in this case, RA captures the object, and transports it to the vicinity of PA. (c) Because PA has received the arrival notification from RA, the transported object is captured. (d) PA passes the object to HA.

2. AN OBJECT TRACKING AERIAL ROBOT

2.1 Introduction

A vision based tracking of a person is one of the popular as well as important research topics. Especially, automatic tracking of a specified person is an indispensable technique in surveillance. But it is also a difficult technique, since the target is often occluded by other objects. Using an aerial image obtained from the air, we can exclude the occlusion at least among the objects on the ground. Moreover, we can also obtain a large view according to the height of the camera. Therefore, a vision based aerial robot, that captures an aerial image and moves in the air freely, will show an excellent performance in human tracking. But no studies have ever tried to track a specified person employing an aerial robot.

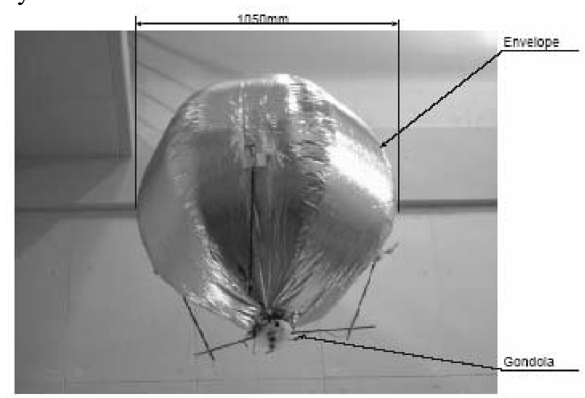
A blimp has a lot of advantages such as flying stably in the air, moving silently and slowly. These advantages are effective for tracking a person. Over the past few years, a number of studies have been reported about the control of a blimp, and most reports employ a vision sensor (e.g., a CCD camera), since it is able to obtain various kinds of information needed for controlling the blimp itself. But most of the researches use only simple visual information such as position, area, and so on [3]. So, we attempt to use human motion information, for tracking a person. Although various techniques for tracking a person employing a robot system are proposed, nobody attempts to recognize the motion of the tracked person yet. If a robot can recognize the motion of the tracked person, the robot will be able to realize a more efficient tracking and to perform various tasks.

Human motion recognition has been attracting much attention in the computer vision field [4,5]. In a robot system, it is important that an entire system should realize real-time operation. Therefore, if we develop an aerial robot which can recognize a human motion, the motion recognition itself should be done in real-time. Although a large number of studies have been performed on human motion recognition based on visual information, only a few attempts have so far been done on real-time motion recognition.

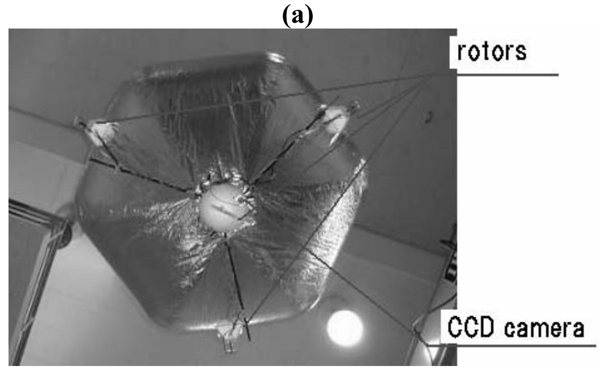
The purpose of this study is to develop a new tracking system employing a visually controlled aerial robot which recognizes a motion of the specified person in real-time. Here we propose the developed vision based aerial robot and an effective real-time motion recognition technique. We use a motion history image (MHI), a superposed motion image (SMI) and an eigenspace technique [6] in order to realize real-time motion recognition.

2.2 Construction of an aerial robot

Figure 4 shows the developed aerial robot. We use a spherical blimp as an aerial robot in the system.



(a)



(b)

Fig. 4 Designed aerial robot: (a) A side view, and (b) a bottom view.

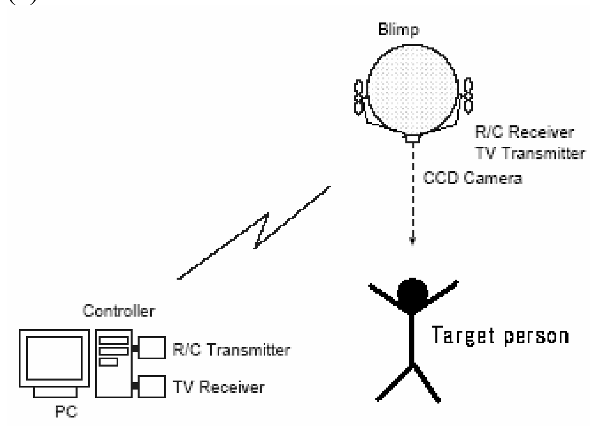


Fig. 5 Configuration of the entire system.

Normal blimp is an oval shape, whereas our designed blimp is a spherical shape. Advantages of the spherical shape are that it is not only smaller than an oval type but also movable to any direction.

The developed robot consists of an envelope and a gondola as shown in Fig.4a. The envelope is a balloon filled with helium gas, and its shape is sphere whose diameter is 1050mm. Moreover, it has 500g buoyancies. As is seen in Fig.4b, the gondola has four rotors and a single CCD camera. The three rotors are used for horizontal movement, and the other for the vertical movement. The CCD camera is set downward at the gondola for capturing downward images.

Configuration of the entire system is shown in Fig.5. The entire system consists of an aerial robot and a remote computer for visually controlling movement of the aerial robot. The remote computer processes downward images provided by the CCD camera on the aerial robot, and control signals obtained from the processed images are transmitted to the aerial robot by the remote computer. The aerial robot and the remote computer mutually communicate through a wireless communication system.

The images obtained from the CCD camera are transmitted to the remote computer on the ground through a TV-transmitter. Employing the information in the images, radio control signals composed of some pulses are transmitted to the blimp through the Radio Control Unit. Stable state of the blimp is realized by controlling it referring to the image information obtained from the camera. Once the position (x, y) of the target in the downward image is obtained by image processing, torques of the rotors are calculated automatically to keep the target in the center of the camera.

2.3 Tracking algorithm

The developed tracking algorithm is divided into two parts. One is a target extracting part, and the other a motion recognition part. The target extraction part works to find a target person from obtained downward images. The motion recognition part, on the other hand, performs recognizing motions of the target person. The aerial robot is tracking a target person according to the result of these processing.

2.4 A motion recognition part

A motion recognition part works to recognize motion of a tracked target. In this part, we use two feature images called a motion history image and a superposed motion image, and an eigenspace technique is employed for realizing real-time motion recognition.

One of the feature images we use in this study is *a motion history image* (MHI). A MHI is the sum of recent successive images with some weights decreasing as time lapses. The latest image is shown clearer than past ones in the MHI. In other words, past images vanish little by little in the MHI. This is a desirable character for realizing real-time recognition.

The other one is a *superposed motion image* (SMI). It is a superposed image generated from summing past successive images with an equal weight. A SMI shows the features of motion well by only one image, and it is also a desirable character for high-speed recognition. A number of MHIs generated from successive images expressing a single motion are included in the SMI generated from the same images. Therefore, MHIs and a SMI have strong correlation between them.

The motion recognition method employed is divided into two steps. One is called the learning step, and the other the recognition step. In the learning step, it performs creating an eigenspace for real-time recognition by the following procedures; generating MHIs and SMIs from learning image data, creating an eigenspace, and calculating reference points. On the other hand, the recognition step generates MHIs from the target forecasted in the target extraction part, maps them into the eigenspace, and recognizes the motion using the distance to the reference points representing each motion.

The reference point expressing motion M is an average point of only SMIs generated from the learning motion M. However, we use both MHIs and SMIs for creating an eigenspace. It is an important point of the developed technique. Usually, images that have strong correlation with each other are mapped on mutually close positions in the eigenspace. So, MHIs and a SMI expressing a single motion are mapped in a narrower region around the SMI. Therefore, an unknown motion can be recognized by a distance measure between the reference points and the point to which the MHI of an unknown motion has been mapped.

2.5 Experimental results

In the experiment, the developed motion recognition system was implemented on the aerial robot. The aerial robot was controlled to capture images of a target person. The image information obtained by the mounted camera was used for controlling the aerial robot. **Figure 6** shows configuration of the experimental environment.

In the experiment, we have achieved the task that, when a person called the aerial robot by a ‘call’ motion, the motion of waving the both hands widely, the aerial robot recognized it and approached to the person. A shot of the experiment is depicted in **Fig.7**.

3. CONCLUSIONS

In the present paper, two intelligent robotic systems were explained as the application of visual sensing employing cameras and computers.

A multiple robots system was proposed for assisting human intention. When a vision agent and a robot agent find out the right object a human agent requests quickly by robot agent's first move, the entire process time is fast. On the other hand, if

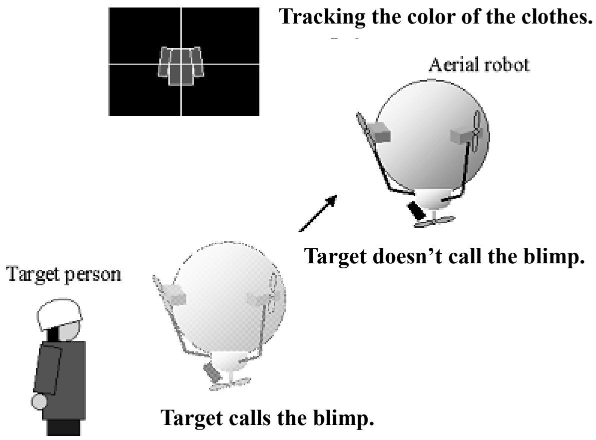


Fig. 6 Configuration of the motion recognition system employing an aerial robot.

Fig. 7 A blimp and a target person in the experiment.

the system selects the right object after the second or more move of robot agent, system performance becomes worse in the system. Further refinement of the communication between a human agent and the system is necessary.

One way of realizing natural communication between a human and a machine system was also presented in the paper. The idea is to understand human sign language using a hand by image processing. Alternatively one may employ voice/speech recognition into the present system. Both techniques are going to be employed in our future version in order to achieve more natural man-machines communication.

In the second part of the present paper, a new human tracking system employing a visually controlled aerial robot was proposed and its performance was shown by experiments. The achievements are that (i) we have developed an aerial robot system, (ii) we have proposed a motion recognition technique employing motion history images and eigenspaces, and (iii) the motion recognition technique was implemented on the developed aerial robot and the robot recognized 'call' motion of a person on the ground in real-time and accessed him successfully.

Extracting a human region without color information remains for further study in order for the proposed technique to achieve better performance. The present technique will be applied to tracking a person who behaves in a suspicious manner, to searching for a person sitting down on the road as he/she feels bad, and so on.

ACKNOWLEDGEMENTS

Contribution to this paper by Mr. Hajime Tawaki, a master student, and Mr. Takehito Ogata, a Ph.D. student of Kyushu Institute of Technology is greatly acknowledged. The author is also thankful to Dr. Joo Kooi Tan and Dr. M. Masudur Rahman of Kyushu Institute of Technology for their continuous support to this study.

REFERENCES

[1] S. Ishikawa, H. S. Kim, F. Inoue, S. Shiramura: "Man-machine collaborative work based on visual communication," *Proc. TENCON2000,* II -321-325, 2000.

[2] S. Igawa, H. S. Kim, S. Ishikawa: "Development of a mobile robot for handicapped people", *Proc. of the Int. Conf. on Control, Automation and Systems*, 59-62, Oct., 2001.

[3] T. Ogata, J. K. Tan, S. Ishikawa: "An aerial robot system tracking a moving object", *Proc. of ICCAS2003*, pp.1917-1920, Oct., 2003.

[4] M. M. Rahman, S. Ishikawa: "Representing human postures/motions using an eigenspace technique", *Proc. of the Int'l Conf. on Artificial Intelligence in Engineering & Technology*, pp.232-236, June, 2002.

[5] D. Cao, O. Masoud, D. Boley, N. Papanikolopoulos: "Online motion classification using support vector machines", *Proc. of IEEE International Conference on Robotics and Automation*, Apr., 2004.

[6] H. Murase, R. Sakai: "Moving object recognition in eigenspace representation: Gait analysis and lip reading", *Pattern Recognition Letters*, Vol.17, No.2, pp.155-162, Feb., 1996.