A Human-Robot Interface Using Eye-Gaze Tracking System for People with Motor Disabilities

Do Hyoung Kim, Jae Hean Kim, Dong Hyun Yoo, Young Jin Lee, and Myung Jin Chung

Abstract: Recently, service area has been emerging field of robotic applications. Even though assistant robots play an important role for the disabled and the elderly, they still suffer from operating the robots using conventional interface devices such as joysticks or keyboards. In this paper we propose an efficient computer interface using real-time eye-gaze tracking system. The inputs to the proposed system are images taken by a camera and data from a magnetic sensor. The measured data is sufficient to describe the eye and head movement because the camera and the receiver of a magnetic sensor are stationary with respect to the head. So the proposed system can obtain the eye-gaze direction in spite of head movement as long as the distance between the system and the transmitter of a magnetic position sensor is within 2m. Experimental results show the validity of the proposed system in practical aspect and also verify the feasibility of the system as a new computer interface for the disabled.

Keywords: real-time, eye-gaze tracking, computer interface, image-based method.

I. Introduction

The proportion of the population with disabilities has risen markedly during the past century. As the data come from the National Health Interview Survey (NHIS), two distinct trends have contributed to the increasing overall prevalence of disability: a gradual rise, due largely to demographic shifts associated with an aging population, as well as a rapid increase that is due to health impairments and accidents [1].

According to an annual report of the Ministry of Public Health and Welfare, 0.73 million people have a motor disability that is on the legs and arms. Their report showed that the causes of a motor disability are a cerebral palsy, an infantile paralysis, a cerebral accident, a cerebral injury, a spinal injury, a muscular injury and an inflammation of a joint. This report also classifies a motor disability into 9 types that is based on the region of a disability (Fig. 1).

The activities of daily living are a set of basic self-maintenance activities considered essential for everyday functioning: bathing, dressing, eating, using the toilet, and transferring (getting into or out of a bed or chair). People with those activity limitations require the assistance of another person to perform them.

Recently, research of assistant robots is also an emerging field of robotic applications. Even though it takes possession of an important role for the disabled and the elderly, they still suffer from operating the robots using conventional interface devices such as joysticks or keyboards. In this paper we propose an efficient computer interface by using real-time eye-gaze tracking system.

In fact, people use their eyes intensively for a large variety of purpose in everyday life. Eyes are regarded as output organs as well as input ones that observe the surroundings. The information from eyes is eye-gaze direction because people can precisely observe an interesting object when the image of the object is located on the middle of their retinas [28].

Much research has been carried out in the area of human-computer interface. Many techniques have been proposed to

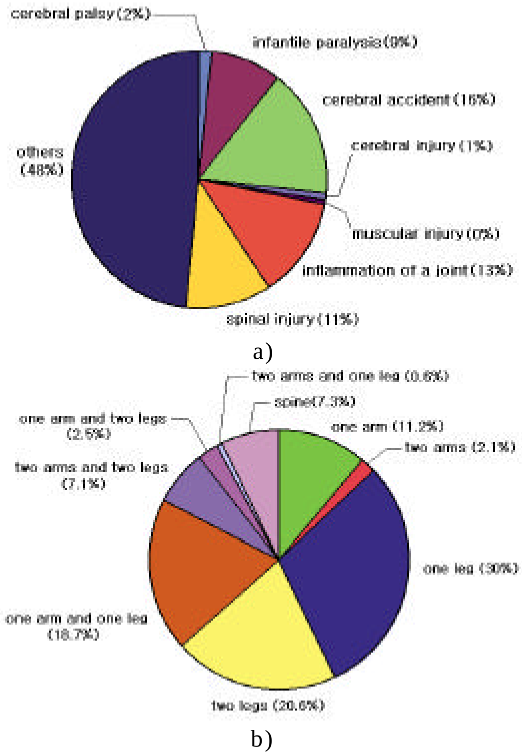

Fig. 1. Distribution chart for the causes and the region of a motor disability: a) the causes, which are a cerebral palsy (2%), an infantile paralysis (9%), a cerebral accident (16%), a cerebral injury (1%), a muscular injury (0%), an inflammation of a joint (13%), a spinal injury (11%) and others (48%), b) the region of disability, which can be categorized like these: one arm (11.2%), two arms (2.1%), one leg (30%), two legs (20.6%), one arm and one leg (18.7%), two arms and two legs (7.1%), one arm and two legs (2.5%), two arms and one leg (0.8%) and spine(7.3%).

obtain the eye-gaze direction [2][3]. Current techniques can be classified into three types on the basis of tracking methods. The first type is based on the measurement of the reflection of some light that is injected onto the eye. Typically, infrared light is used to decrease the distraction of the user, and to avoid interference from other light sources such as lamps. However the systems using this method are very expensive

Manuscript received: Aug. 29, 2001., Accepted: Nov. 30, 2001.
Do Hyoung Kim, Jae Hean Kim, Dong Hyun Yoo, Young Jin Lee, Myung Jin Chung: Department of Electrical Engineering and Computer Science, KAIST.(bborog@cheonji.kaist.ac.kr/ mjchung@ee.kaist.ac.kr)

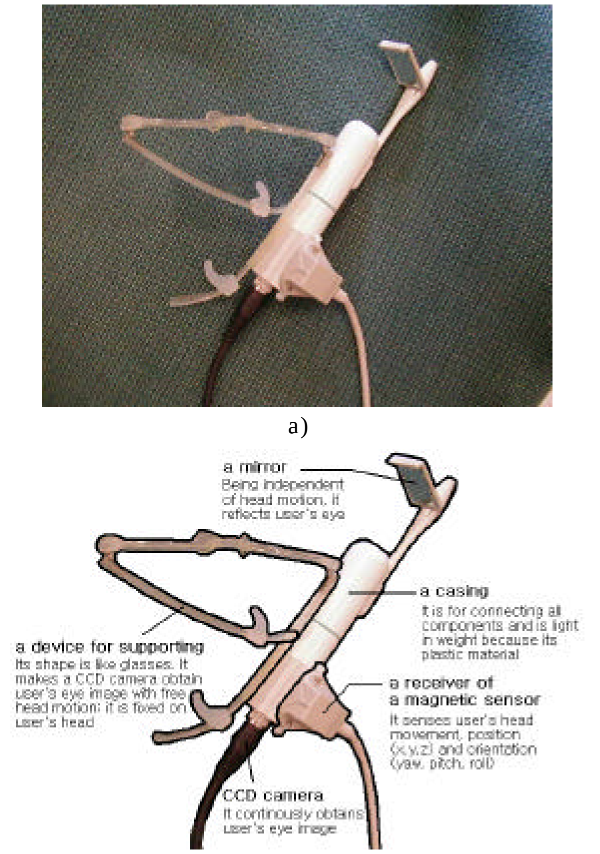

Fig. 2. An eye-gaze tracking system: a) a real image of this tracking system, b) a system structure; the system is composed of a mirror, a casting, a receiver of a magnetic sensor, a device for supporting and a CCD camera.

[14]. The second is the measurement of the electric potential of the skin around the eyes by attaching patches on the skin [15]. The third type uses a special contact lens to detect eye gaze direction. The inconvenience to a user is a disadvantage of this method.

We propose an image-based method that does not use any special light source such as infrared light. The proposed system extracts the eye features in real time, tracks the feature points and calculates the eye-gaze direction. Experimental results show that the proposed system is feasible as a new computer interface for the disabled.

II. Eye-Gaze tracking system

To determine the eye-gaze direction, the head and eye movement data should be given. Therefore, a head mounted eye-gaze tracking system is composed of two subsystems; a head tracking subsystem and an eye movement tracking subsystem. So the eye-gaze tracking system ascertains the eye-gaze direction combing head tracking data and eye movement data obtained from the head tracking subsystem and the eye tracking subsystem, respectively. Fig.2 shows the designed system to track eye-gaze direction. In Fig. 2, a casting attached to spectacles is 17cm long and has a mirror in front of it. Inside the casting, there is a CCD camera. A cubic sensor, which adheres to the upper side of the casting, is a receiver of a magnetic position sensor. This magnetic position sensor detects head motion with respect to the inertial frame.

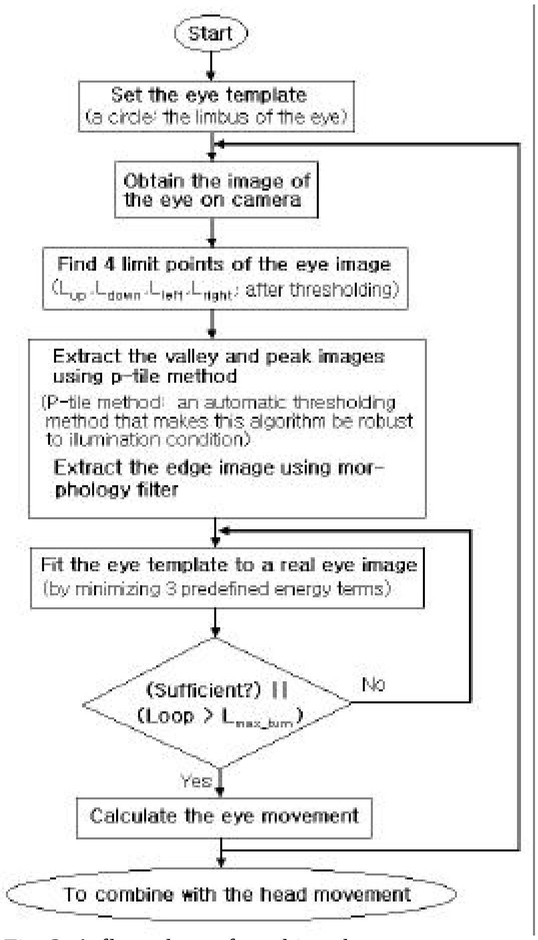

Fig. 3. A flow chart of tracking the eye movement.

III. Eye movement tracking

In this section the proposed algorithm to extract and track the eye movement is described. The eye movement tracking system that was previously addressed is composed of a CCD camera and a mirror and is a head mounted system. Two components are fixed to a plastic case so that the eye movement tracking system continuously can obtain images of user's eye and process them and draw out the information of the eye movement sequentially while user's head is moving.

In this paper, using an iris boundary model we propose an algorithm that overcomes the problems of the existing limbus tracking methods [7]. Limbus is the boundary between the white sclera and the dark iris of the eye. One problem of the existing methods is that it is difficult to detect the vertical eye movement precisely because eyelids partially cover the limbus. The other is that they use infrared source to be robust to the change of illumination and so, it is expensive.

Fig. 3 shows a flow chart of the proposed algorithm. The proposed algorithm is as follows.

1) Define matching functions which indicate the correspondence between the iris boundary model and basis images.

2) Continuously extract three basis images from sequential

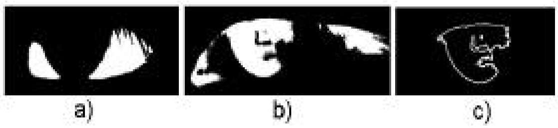

Fig. 4. Three basis images of an eye: a) a binarized peak image, b) a binarized valley image and c) a binarized edge image.

images that are obtained by a camera.

3) Search iteratively for the optimal position of the iris boundary model that maximizes the matching function.

4) Calculate the rotation of eyeusing the position of the iris boundary model obtained from 3).

5) Repeat Step 2) to 4) until eye movement tracking is terminated

1. Extraction of three basis images

Three basis images from the original images acquired from the CCD camera are extracted. Thus, three basis images have sufficient information about the iris boundary. The iris boundary that we want to know here is the boundary that divides the bright and dark regions of the eye. First, we select two basis images to express the bright region and dark region. We define the valley image with the high value at the dark pixel and the peak image with the high value at the bright pixel. These two basis images are binarized with a proper threshold value to divide the bright and dark regions. To select the threshold value robustly against the change of illumination, p-tile method which decides the threshold automatically is used [13]. The p-tile method is widely used when the size of region is known. Secondly, we select third basis image with the high value at the pixel, which is the edge of the iris boundary. We denote this basis image as an edge image. To extract the information of the iris boundary after the iris region is obtained, a morphology filter is used. The blob coloring algorithm, which is generally used in image processing to classify the regions, is used to remove the unnecessary region while obtaining the basis image. Three basis images obtained by the above image processing are shown in Fig. 4. Fig. 4a) is a binarized peak image which presents the bright region of the eye and Fig. 4b) is a binarized valley image which presents the dark region of the eye. An edge image, which presents the boundary of the iris, is shown in Fig. 4c).

2. Iris boundary model

The boundary of the iris can be modeled as a circle. The circle is well fitted and its modeling error is very small [5]. Hence, we need three parameters for representing the circle: the center of circle (Xc,Yc) and radius r.

3. Model matching

In this research, we define three matching functions that indicate the degree of agreement between the iris boundary model and three basis images. Three matching functions are as follows.

1) The matching function for a valley image M_v is used for expanding a radius of the limbus model and is described as follows.

$$M_v = \sum_{i \in circle} valley\_image_i \tag{1}$$

where $valley\_image_i$ is the pixel value of a binarized valley image and *circle* describes the inner area of the iris boundary model. The summation is made over the inner area of the iris boundary model.

2) The matching function for a peak image M_p and is described as follows.

$$M_p = -\sum_{i \in circle} peak\_image_i \tag{2}$$

where $peak\_image_i$ is the pixel value of the binarized peak image .

3) The matching function for an edge image M_e is described as follows.

$$M_e = \sum_{i \in boundary} edge\_image_i \tag{3}$$

where $edge\_image_i$ is the pixel value of a binarized edge image and *boundary* describes the iris boundary. The summation is made over the circular trajectory of the iris boundary model. We construct the combining matching function by properly weighting the above three matching functions. We indicate this matching function as M_{total}.

$$M_{total} = W_v M_v + W_p M_p + W_e M_e \tag{4}$$

Due to this matching function, the iris boundary model moves to the best matching position where the iris boundary model, *i.e.*, circle, matches to the iris image region optimally.

IV. Head movement tracking

In order to measure the 6DOF pose (position, orientation) of head movement, we add a magnetic position sensor to the eye movement tracking subsystem [7]. A receiver of the magnetic position sensor is small and light and is mounted on the eye movement tracking subsystem. The magnetic position sensor consists of two parts: a transmitter, which generates a magnetic field, and a receiver, which senses the position and orientation [4]. This information is transmitted to the electronic unit. And the electronic unit processes analog signal of sensed data. Processed data provide 3 dimensional position and orientation (roll, azimuth, elevation) of a head. Measurement quality is influenced by low frequency electro-magnetic interference (EMI) and the presence of large metal objects in the field. To ensure performance, the transmitter and receiver are located away from sources of EMI such as monitors and power supplies [4]. But the required distance located away from them is not far.

V. Calibration process for calculating eye-gaze direction

The eye gaze direction is calculated from the data obtained in the previous section. The procedure for calculating eye gaze direction is comprised of the following four steps.

1) The center of eye rotation with respect to the coordinates attached to the receiver is obtained from the structure of the system.

2) The angle of eye rotation (pitch, yaw) is obtained from the CCD camera. For this process, the relation between the

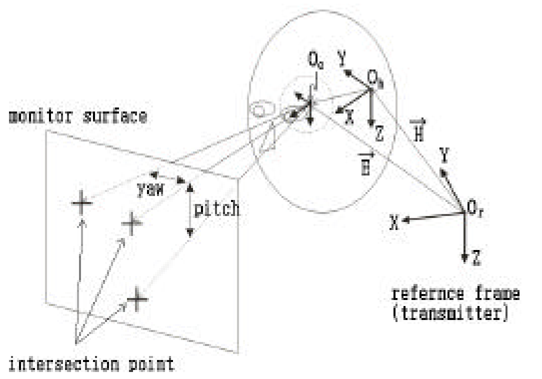

Fig. 5. An outline how to calculate of the eye-gaze direction of user.; It roughly presents the way how to combine head and eye movements.

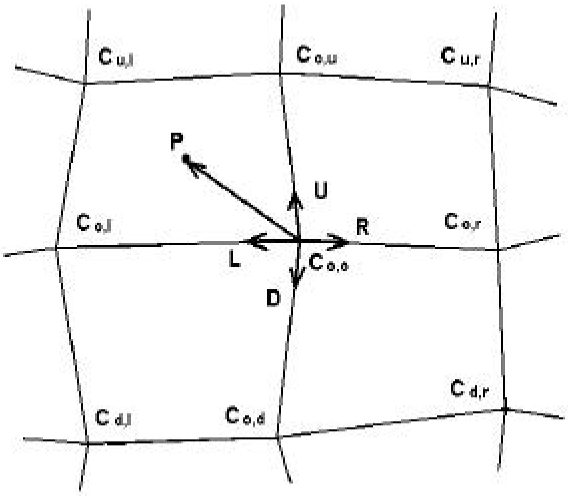

Fig. 6. A calibration map for calculating the eye movement

image coordinates and the eye coordinates must be calibrated.

3) The vector that indicates the eye gaze direction with respect to the monitor coordinates as inertial coordinates is calculated. For this process, the relation between the monitor coordinates and the coordinates attached to the transmitter must be calibrated.

4) The intersection between the vector obtained from 3) and the monitor plane is obtained. In this process, the starting point of the vector is the center of eye rotation (Fig. 5).

1. Calculation method for the eye rotation

The method of calculating the angle of eye will be presented here. As mentioned before, the calibration between the image coordinates and the eye coordinates is needed in order to find transformation matrix between two coordinates [6]. The eye coordinates have its origin at the center of eye rotation. The calibration method and the calculation of the angle of eye rotation are as follows:

1) Let a point P be the center of the iris obtained by the proposed algorithm.

2) Assume that an eye does not rotate. Let the position vector P be the origin of the image coordinates. Search four direction unit vectors of the image coordinates that correspond to the eye rotation (pitch, yaw). Obtain the boundary of the image coordinate that corresponds to the limit of eye rotation (Fig. 6).

3) If the point P is outside of the boundary, set this point as the nearest point from the boundary. If not, obtain the position vector by summing four unit vectors obtained above. Calculate the angles (pitch, yaw) of the eye rotation from this vector representation.

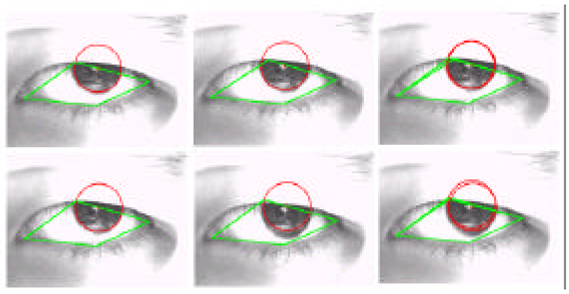

Fig. 7. The experimental results of the iris boundary (limbus) model matching in case of the sequential images (bright environment).

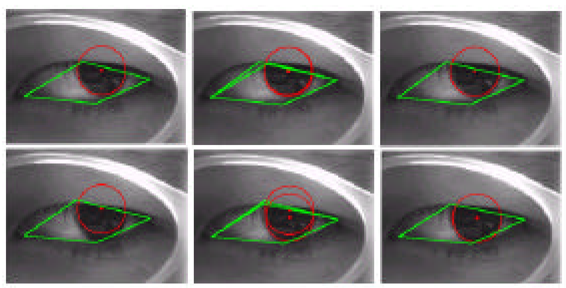

Fig. 8. The experimental results of the iris boundary (limbus) model matching in case of the sequential images (dark environment).

VI. Experimental results

The proposed algorithm is implemented on Pentium III 600 IBM PC and the sequential eye images are captured by Meteor image grabber. The algorithm is programmed by Visual C++ 6.0.

1. Experimental results for eye movement

Figs. 7 and 8 show the results of the dynamic images. Fig. 7 shows the experimental results performed in a bright environment and Fig. 8 shows the experimental results in a dark environment where illumination is removed. It is observed that the proposed algorithm is robust to a change of illumination and applicable to the dynamic images. Real time implementation of this system is important in practical aspects. Elapsed time to perform the proposed algorithm in the experiment is 110msec for the 320 by 240 image. This indicates that the proposed system produces about 10 eye gaze directions per second and operates almost in real time.

2. Experimental results for eye-gaze tracking

The experimental results of our eye gaze system are presented here. To analyze the accuracy of the system, a

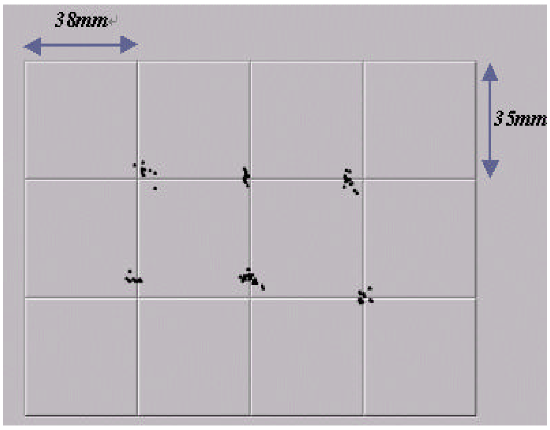

Fig. 9. Experimental result for the resolution of an eye-gaze tracking system.

subject gazed exactly at the intersections of the lattice on the monitor (Fig. 9). Then we calculated the subject's eye-gaze direction and the eye-gaze point on the monitor plane by using the proposed algorithm We compared the calculated points with the intersections of the lattice on the monitor. In this experiment, the distance between the subject's head and the monitor is approximately 0.5m.

Fig. 9 shows the results of the experiment, and error analysis of this experiment is shown in Table 1.

Table 1. Experimental results for the accuracy of the eye-gaze tracking system.

mm	Min error	Max error	Mean	Deviation
u	0.28	6.75	0.68	8.1
v	0.27	8.12	2.45	9.3

3. Application to human computer interface

To verify the feasibility of the proposed system as a human-robot interface, the following experiments were performed. We constructed a program for controlling a robotic arm. The program has a pull-down command menu and an additional 'OK' menu. The command menu is for transferring the command of the joint movement. When a user moves the eye mouse pointer to the menu on the monitor, the menu automatically is pulled down. If the user wants to confirm to select the menu, the user moves the eye mouse pointer to the additional 'OK' menu instead of clicking the menu. The user controlled the mouse pointer by using the proposed system to operate the robotic arm. Fig. 10 shows the program used in this experiment. In this figure, an additional 'OK' menu, which is for the confirmation of menu selection, is shown.

Another experiment, wherein a user played a computer game, "freecell", was performed. The experimental environment is shown in Fig. 11. The user can play "freecell" by using the proposed eye-gaze tracking system and manual click operation. The click operation can be performed by another interface method such as voice command. Experimental results showed that the proposed system is applicable as a human-robot interface and good interface device for the disabled.

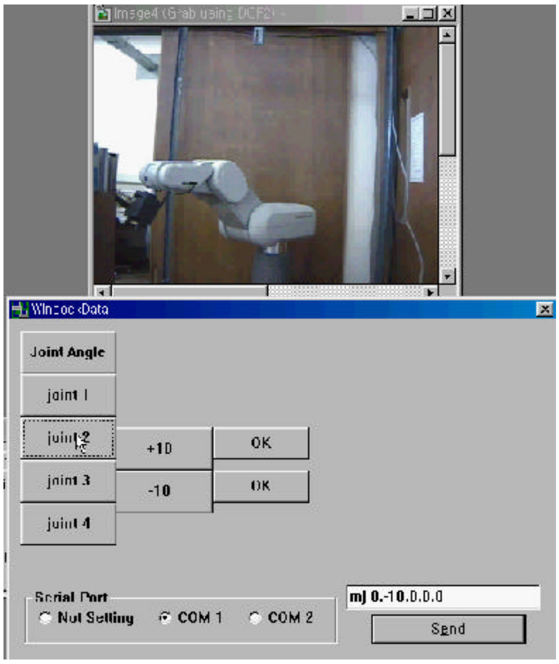

Fig. 10. Experiment of controlling robotic arm by using the eye-gaze tracking system.

Fig. 11. A scene of playing a computer game, "freecell", by using the proposed eye-gaze tracking system.

VII. Conclusions

We presented a real time eye gaze tracking system and proposed an eye tracking algorithm using images captured by a CCD camera. To compensate for head movement, a magnetic position sensor attached to the camera gives the position and orientation information of the head with respect to the monitor coordinates (inertial frame). The user's eye gaze direction is calculated very accurately in spite of head movement as long as the distance between the system and the transmitter of the magnetic position sensor is within 2m. We verified the feasibility of the proposed system as a human-computer interface for the disabled by performing experiments.

Our future work will concentrate on improving the accuracy of the proposed system, making the system more robust and faster and considering other eye movement like a blink motion of a user. To improve the accuracy of the proposed system, the calibration of eye-rotation should be considered. One method is to make a calibration map be denser [6]. It must made the angle of eye-rotation be more accurate. The proposed system will be verified by numerous experiments for people with motor disabilities.

References

[1] S. Stoddard, L. Jans, J. M. Ripple and L. Kraus, "Chartbook on work and disability in the united states, 1998," An InfoUse Report, U. S. National Institute on Disability and Rehabilitation Research, 1998.

[2] A. J. Glenstrup, T. Engell-Nielsen, "Eye controlled media: present and future state," *B. S. Dissertation, Copenhagen* University, 1995.

[3] K. F. Arrington, "Viewpoint eye tracker," *Arrington Research*, 1997.

[4] *Flock of Birds: Installation and Operations Guide*. Burlington, VT: Ascension Technology, 1999.

[5] T. Haslwanter, S. T. Moore, "A theoretical analysis of three-dimensional eye position measurement using polar cross- correlation," *IEEE Trans. on Biomed. Eng.*, vol. 42, no. 11, pp. 1053-1061, November 1995.

[6] K. T. Lee, "Eye-controlled human/ computer interface using the line-of-sight and the intentional blink," *Ph. D. Dissertation*, Korea Advanced Institute of Science and Technology, 1995.

[7] D. H. Kim, "Development of an eye-gaze tracking system using video image sequence and magnetic sensory Information," *M. S. Dissertation*, Korea Advanced Institute of Science and Technology, 2000.

[8] A. L. Yuille, P. W. Hallinan and D. S. Cohen, "Feature extraction from faces using deformable templates," *Int. J. Comp. Vision*, vol. 8, no. 2, pp. 99-111, February 1992.

[9] X. Xie, R. Sudhakar and H. Zhuang, "On improving eye feature extraction using deformable templates," *Pattern Recogn.*, vol. 27, no. 6, pp. 791-799, January 1994.

[10] R. S. Allison, M. Eizenman and B. S. K. Cheung, "Combined head and eye tracking system for dynamic testing of the Vvstibular system," *IEEE Trans. on Biomedical Engineering,* vol. 43, no. 11, pp. 1073-1082, November 1996.

[11] X. Xie, R. Sudhakar and H. Zhuang, "Real-time eye feature tracking from a video image sequence using kalman filter," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 25, no. 12, pp. 1568-1577, December 1995.

[12] 1998 Annual Report for People with Disability in Korea, *Ministry of Public Health and Welfare*, 1998.

[13] R. Jain, R. Kasturi, and B. G. Schunck, "Machine vision," McGraw-Hill, 1995.

[14] (2001, Jul. 13). [Online], Applied Science Laboratories, Available: http://www.a-s-l.com/

[15] (2001, Jul. 13). [Online], "The eagleeyes project," Available: http://www.cs.bc.edu/~eagleeye/

Do Hyoung Kim

received the B.S. degree in control engineering from Chonbuk National University in 1998, the M.S. degree from Korea Advanced Institute of Science and Technology in 2000. He is currently studying toward the Ph. D. degree in electrical engineering at Korea Advanced Institute of Science and Technology. His research interests include mobile robot, human interface and computer vision.

Dong Hyun Yoo

rceived the B.S. degree in control and instrument engineering from Ajou University in 1997 and the M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology in 1999. He is currently studying toward the Ph.D. degree in electrical engineering and computer science at Korea Advanced Institute of Science and Technology. His research interests include computer vision and face tracking.

Jae Hean Kim

rceived the B.S. degree in electrical engineering from Yonsei University in 1997 and the M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology in 1999. He is currently studying toward the Ph.D. degree in electrical engineering and computer science at Korea Advanced Institute of Science and Technology. His research interests include mobile robot and computer vision.

Young Jin Lee

rceived the B.S. and the M.S. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology in 1994 and 1996 respectively. He is currently working as a Ph.D. candidate in the department of electrical engineering and computer science, Korea Advanced Institute of Science and Technology. His research interests include mobile robots, machine vision.

Myung Jin Chung

received the B.S. degree in electrical engineering from Seoul National University in 1973, the M.S. degree in electrical and computer engineering in 1977 and the Ph.D. degree in computer, information and control engineering in 1983 both from the University of Michigan. Since 1983, he has been with the Department of Electrical Engineering and Computer Science at Korea advanced Institute of Science and Technology, where he is currently a Professor. In 1976, he worked for the Agency for Defense Development, Seoul, as a member of technical staff and from 1981 to 1983 he worked for the Center for Robotics and Integrated Manufacturing, the University of Michigan. His research interests include robust control, intelligent control, robot path planning and control, telerobot control and multisensor fusion.