Research Article

New Service Robot Controlled by Human Eye Which Allows Patients in Hospitals Self-Services Remotely

Kohei Arai*

Department of Information Science, Saga University, Saga City, Japan

*Corresponding author: Kohei Arai, Department of Information Science, Saga University, Saga City, Japan

Abstract

New service robot controlled by human eye which allows patients in hospitals self-service and speaks remotely. Patients' sight can be estimated with camera images which are mounted on their glass. Also, patients may control the service robot by using the estimated sight for choosing some functions in the menu (moving to anywhere they want, talking to the other persons, etc.). Three types of command (move forward, turn left, and turn right) are available for robot control. These commands can be selected by looking to the center, right and left directions. Patients can create sentences by choosing desired keys from the keyboard which is shown in their head mount display. By using text to speech function, they can talk with some other persons. On the other hand, the service robot has another camera at the tip of the robot for capturing the scenery. Between robot and user, they are separated by distance. The robot streams the scenery video and receives command sending by the patients as clients. The patients could control the robot remotely even in the long distance because user could see the scenery in the front of robot. The performance of the robot is evaluated for the routes including obstacles which are designated in the hospital. All the functions are confirmed together with moving performance.

Keywords: Service robot; Computer input just by sight

Introduction

In recent years, service robots are getting much popular in the world. Bill Gates' perspective, "domestic robot (service robot)" is its typical one [1]. Erico Guizzo summarized the recent development of "service robot" technology development [2]. As shown in the paper, steady home of the robot's activity has been given. These are mainly developed by the robot on behalf of part of the household chores such as cleaning and guidance at the facility. The Electrolux Trilobite, Roomba, The SLAM based Neato Robotics vacuum cleaner robot etc. are good examples.

The authors have already realized key input and mouse operation (gaze input system) when using computer by gaze [3-7]. In addition, I propose travel control [8] of electric wheelchair using gaze input system, conversation support system [9], feeding support system [10] etc.

This paper aims to realize a service robot that helps patients, disabled people, long-term care recipients, elderly people (users), etc. in hospitals, group homes, etc. As for the function of the proposed robot:

1)	It controls the traveling by the line of sight, so that users with physically handicapped can freely select and determine destinations while watching the surrounding scenery with only the line of sight.

2)	Enjoy conversation with surrounding people on the way, can communicate via voice.

3)	Recognize the obstacle and automatically perform evasive behavior [11], if it is impossible or depending on the judgment of the user performing the running control of this service robot, depart from the service robot automatically returning to the original point.

As a result of experiments by healthy volunteers, I can confirm the desired function and I decided to withstand practical use enough, so I will present it here.

As research and development related to living support robots, there is a life support robot practical project of the new energy Industry Comprehensive Technology Development Organization. This focuses on human mobility support. There is a proposal by Koren in the robot running navigation [12], and there are methods proposed by Yoshimi [13] for detecting the object from the image captured by the robot. In addition, there are Okada's proposals [14] in the living support by humanoid robots, Ando et al. [15] make proposals on middleware of robots that support them. The service robot proposed in this paper makes the service robot have the function of the line of sight input, and the user conducts conversation and errands necessary for living from a remote place through this function, and the service robot in the past research Are functionally different from each other.

In Section 2, I propose a system to achieve the above object. In particular, with regard to the gaze input method, a method [16] of realizing high precision and high speed input using a newly devised moving keyboard will be described in detail in comparison with the conventional fixed keyboard. Section 3 describes the performance evaluation experiments required to achieve the objectives of the proposed system and summarizes the main results in Chapter 4.

Proposed Method

System configuration and functionality

The Configuration of the proposed system is shown in Figure 1. The service robot acquires images from the camera installed in front
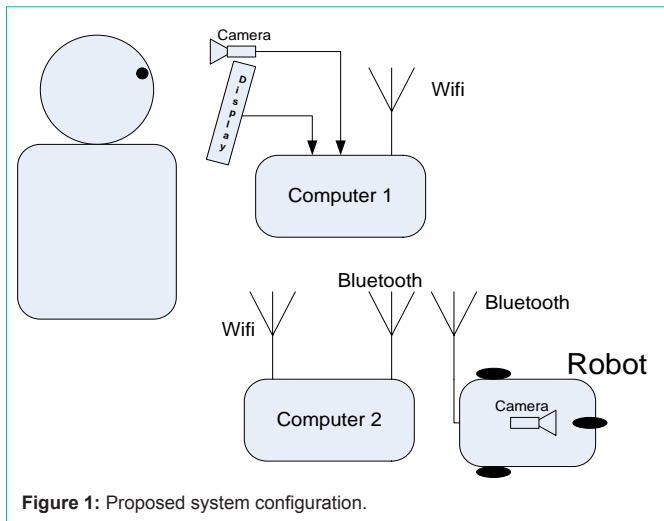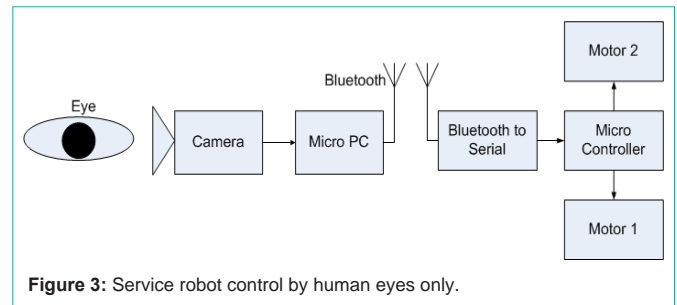
Figure 1: Proposed system configuration.
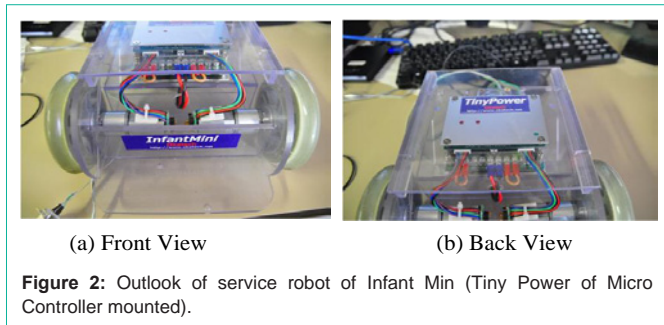

(a) Front View   (b) Back View

Figure 2: Outlook of service robot of Infant Min (Tiny Power of Micro Controller mounted).

Table 1: Major Specification of Service robot (infant mini).

Size	W213xD233xT105mm
weight	1.3kg
Maximum payload	5kg
Speed	18.12cm/s
Rotation	8000rpm
Motor	12V,1.5W
Tire	100mm
Caster	60mm
Li-Polymer Battery	11.1V,2200mAh

of it. The acquired image is transmitted to a Wi-Fi terminal located at the user's hand by way of the Wi-Fi terminal via Bluetooth so that the user can view the image. Bluetooth and Wi-Fi are used properly from the viewpoint of data transmission speed, and Wi-Fi is selected from the viewpoint of effective utilization of existing facilities because the proposed system assumes use in hospitals and group homes.

A user can grasp the line of sight with a camera that images around the eyeball. From this image gaze estimation is performed by the Wi-Fi terminal. The running of the service robot is controlled by the line of sight direction. Further, the user inputs a sentence by a line of sight, and the sentence is outputted as a voice from the speaker provided in the Wi-Fi terminal (computer 2) of the service robot by the reading software. Therefore, the user can experience a pseudo trip within the Wi-Fi coverage and talk with the surrounding people in the place where it goes. (Character input is performed by the line-of-
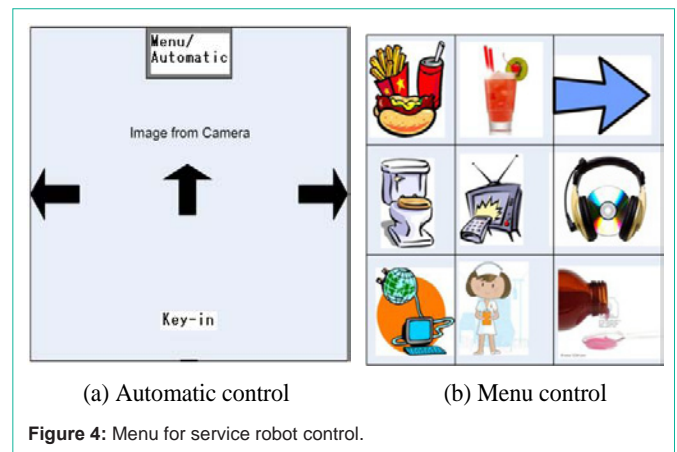

Figure 3: Service robot control by human eyes only.

(a) Automatic control (b) Menu control

Figure 4: Menu for service robot control.

sight input system, voice is output by the reading-out software, User can listen to people's surroundings via the microphone that is being played) and you can also fulfill your business.

The proposed service robot considers the inside of facilities such as hospital, group home for the time being. In addition, in order to avoid collision with static obstacles (patients, doctors, nurses, other service robots, etc.) and to ensure safe driving, multiple service robots only camera images provided in the service robot In addition, I devised to be able to detect obstacles by ultrasonic sensors and to avoid them.

Hardware

As a service robot, Infant Mini manufactured by Okatec Co., Ltd. was used. The main specifications are shown in Table 1, and the appearance is shown in figure. In addition, Figure 3 shows the mechanism of running control of the service robot with sight line. I can grasp the line of sight of the user with the camera so that I can judge which of the screen part (Figure 4) which divided the camera field of view into 3×3.

Moving control

The driving mode includes automatic running and menu driving mode. In the initial setting, the automatic driving mode of Figure 4a is selected. Therefore,

1) When the user is looking at one of the left and right, the running is controlled to turn left and right.

2) If the user is looking at the center, keep traveling forward.

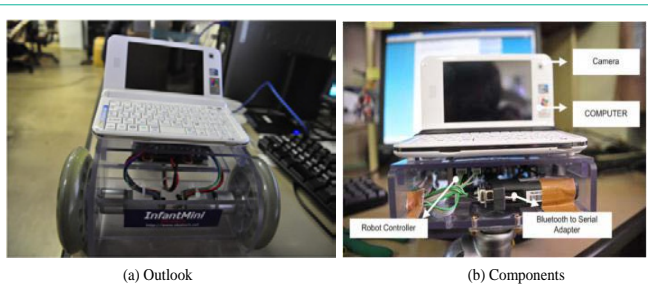3) When the user sees the top, it is possible to switch between menu running / automatic driving.
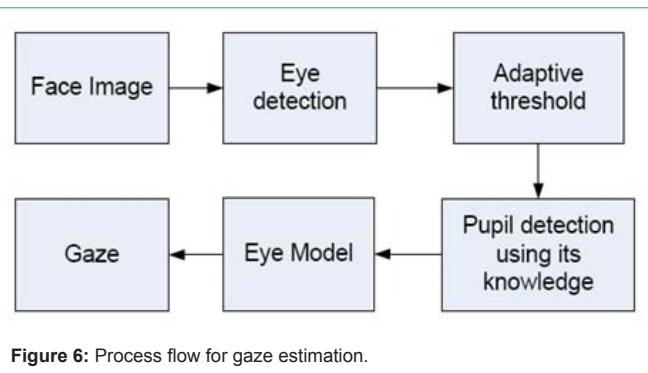
(a) Outlook   (b) Components

Figure 5: Outlook of the service robot.

Figure 6: Process flow for gaze estimation.

Table 2: Specification of the camera used.

Resolution	1,300,000 pixels
Minimum distance	20cm
Frame rate	30fps
Minimum illumination	30lux
Size	52mm (W)x70mm (H)x65mm (D)
Light	105g

Figure 7: Eye detection flow.

Table 3: Specification of the head mount display used.

Pixel size	SVGA+(852x600)
Field_of_view	30degree
Format	SVGA, NTSC, RS-170, PAL
Specification	MIL-STD_810F
Dry Battery	USB_Interface (2_batteries)

4) Stop the service robot when looking at the top left and right.

5) When looking at the bottom left and right, the service robot is designed to return to the home position (initial setting) where the user is located.

When switching to menu traveling in Figure 4b, the service robot is automatically driven to move to a destination in a predetermined menu. When the user sees the upper right corner after reaching the destination, the mode switches to the automatic driving mode. The cruise control menu has nine sub-menus from the upper left corner to the lower right corner, to eat, to have a drink, to have a drink, return to automatic driving mode, want to go to the bathroom, want to watch TV, listen to music Want to search for information, want me to come to a nurse, because medicine is gone, I want you to deliver but you can choose.

Signals transmitted via Bluetooth are serial converted, and microcontrollers control motors of both wheels of the left and right service robots. The appearance of the service robot is shown in Figure 5.

As the payload on the service robot, I chose the ultra-compact PC "KOHJINSHA PM" (4.8-inch wide LCD, CPU: Atom, OS: Windows XP). Because this is as light as 345g, it does not put a big burden on the service robot.

Kohjinsha have announced their latest MID, the PM-series, and if users have followed the scene in ultra compact PCs over the past few months then user will be quick to recognize it as the UMID mbook M1. The Kohjinsha PM keeps the mbook's Atom Z510 1.1GHz processor and 512MB of RAM but seemingly lacks an integrated 3G option. There is also a 16GB SSD, WiFi and Bluetooth, together with a 4.8-inch 1024 x 600 touch screen and a battery which Kohjinsha are saying is good for up to 7hrs; that is at odds with what mbook

M1 users have experienced, certainly. The palm-sized ultraportable measures a scant 158 x 94.2 x 22 mm.

Line of sight (gaze) estimation

The gaze estimation flow is shown in Figure 6. First, an eyeball peripheral image is extracted from the face image acquired by the Net Cow Boy DC-NCR 131: NIR camera shown in Table 2 whiles the major specification of Head Mount Display: HMD is shown in Table 3. For the eyeball detection, as shown in Figure 7, the method of Viola-Jones is adopted [17]. It scans the search window (for example 8 pixels × 8 pixels judgment area) from the upper left and moves it sequentially, and for each area of the search window, the identification of the Cascade structure previously prepared by the machine learning algorithm called Boosting And determines whether the image is an eyeball image by using a device. This image is binarized by the adaptive threshold method. This is to decide an adaptive threshold according to the surrounding gray scale state with respect to the gray scale image and convert it into a binary image. Therefore, the target pixel is determined not to be the entire image, but to be either white or black depending on the brightness of the neighboring region.

Approximate the contour of the binary image by ellipses. Estimate pupil center based on knowledge base assuming ellipse [4].
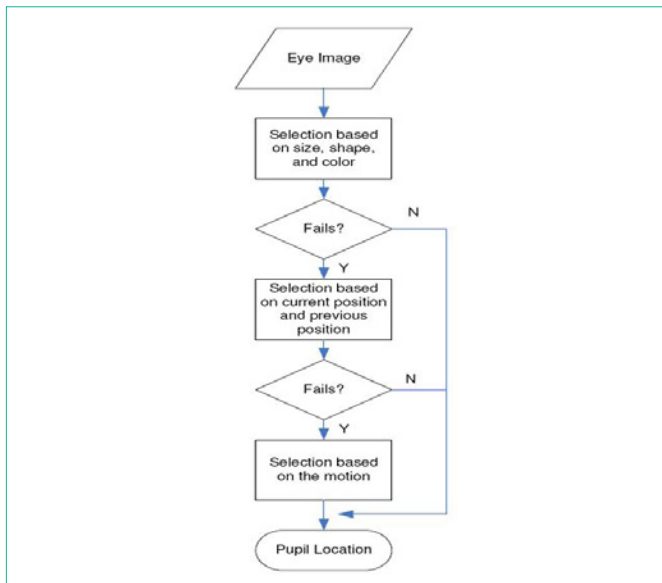
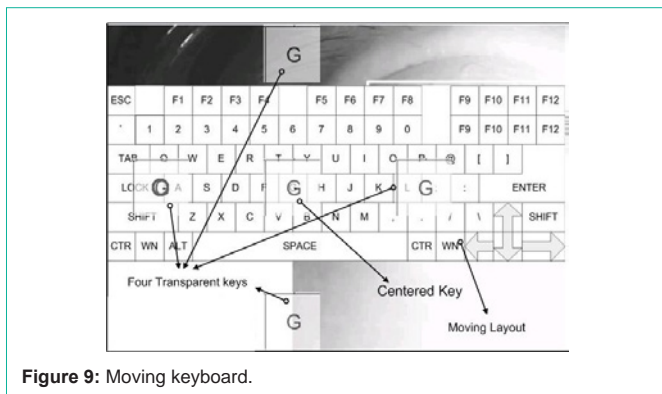**Figure 8:** Pupil center estimation based on knowledge based system.

Figure 9: Moving keyboard.

A pupil is detected by the bright pupil method based on a previously-established near-infrared camera image [8], a pupil shape is elliptically approximated, and the intersection point of the major axis and minor axis is taken as the pupil center [7].

I also consider the knowledge base on the size, shape and color of eyeballs [8]. In addition, by using the pupil extracted from the immediately-processed eyeball as a template, this is deformed to perform matching with the currently acquired eyeball image, and even if the pupil cannot be seen due to momentary blinking or the like, the pupil center It is devised to be able to extract [6]. This method is called a pupil center estimation method based on a knowledge base. This processing flow is shown in Figure 8.

Based on the eyeball shape model (24mm diameter eyeball), a straight line connecting the center of rotation of the eyeball and the center of the pupil is regarded as a line of sight. This line of sight is taken as the point of view with the intersection with the display seen by the subject.

Sentence creation

When entering a sentence, the user first stops the service robot by looking at the top left and right of the display. After that, a moving keyboard shown in Figure 9 appears by looking at the lower part

Table 4: Comparison of pupil detection performance in unit of percentage St.Dev. denotes standard deviation.

User No.	Nationality	Proposed Method	Adaptive Threshold	Template. Matching
1	Indonesian	97.05	97.05	7.28
2	Indonesian	95.44	78.83	56.57
3	Indonesian	78.83	77.72	58.22
4	Vietnamese	95.5	95.16	38.93
5	Sri Lankan	87.42	74.12	47.88
6	Indonesian	97	95.63	64.17
	Average	91.87	86.42	45.51
	St.Dev.	7.33	10.57	20.7

of the display. When inputting sentences with a moving keyboard, subjects can select all keys by just looking at one of the up, down, left and right, and the translucent enlargement keys at the center. That is, by looking at one of the up, down, left, and right enlargement keys, if the keyboard displayed under that key moves by one key and the desired key appears in the center, blinking or gazing at more than 1 second the key can be determined.

In this way, in the case of a moving keyboard, the user moves until the keyboard itself reaches the candidate key, so all the key inputs are possible by looking at the five magnification keys, but in the case of the conventional fixed keyboard. One key size is small due to the limitation of the size of the screen, the key input precision is low, and it is forced to re-enter. The key input success rate and the fixed key input time and the comparison of the moving keyboard are described in detail in reference [16], so they are omitted here.

Experiment

Pupil center estimation accuracy

Estimation of pupil center in gaze estimation is an important factor determining accuracy. Pupillary center detection rate changes due to individual differences such as eye size, shape, head posture, shadow effect, corneal curvature and the like. Therefore, here I first compare the pupil center detection rates of the proposed method and proposed method, Adaptive Thresholding method and simple template matching method, by six subjects with different nationalities. The results are shown in Table 4.

Compared with the adaptive threshold method, the proposed method adds to the adaptive thresholding method using the knowledge base related to the pupil when obtaining the center of the pupil from the binarized image. Therefore, the effect of based on the knowledge base is linked to improvement of the key input success rate of about 5%.

Blink detection accuracy

The desired key selected by the moving keyboard is determined by either the gaze mode or the blink mode. When deciding by the blinking mode, the blink detection success rate is extremely important. Subjects with the highest key input success rate by the proposed method in Table 4. The blink detection success rate was evaluated for the average subject 5 and the lowest subject 3.

I chose a simple template matching method and motion template matching as comparison objects. The number of blink inputs was set

Table 5: Comparison of blink detection among the proposed method and conventional template matching and motion template matching methods St.Dev. denotes standard deviation.

	Proposed Method		Template Matching		Motion Template[1]	
User	Detected	Accuracy (%)	Detected	Accuracy (%)	Detected	Accuracy (%)
1	30	100	29	96.67	25	83.33
2	30	100	22	73.33	23	76.67
3	30	100	13	43.33	28	93.33
Average	30	100	21.33	71.18	25.33	83.75
St,Dev.	0	0	8.02	26.74	2.52	8.39
	Proposed Method		Template Matching		Motion Template[2]	
User	Detected	Accuracy (%)	Detected	Accuracy (%)	Detected	Accuracy (%)
1	30	100	29	96.67	25	83.33
2	30	100	22	73.33	23	76.67
3	30	100	13	43.33	28	93.33
Average	30	100	21.33	71.18	25.33	83.75
St,Dev.	0	0	8.02	26.74	2.52	8.39

Table 6: Comparison of key-in success rate between fixed and moving keyboard.

User	Expertly	Moving Keyboard (%)	Fixed Keyboard (%)
1	Expert	100	92.86
2	Expert	100	76.19
3	Beginner	82.14	71.43
4	Expert	100	85.71
5	Beginner	71.43	78.57
6	Expert	100	66.67
Average		92.26	78.57

to 30, and the detection success rate for each subject was evaluated. The comparison result is shown in Table 5.

As is clear from Table 5, the blink detection success rate by the proposed method can be reliably detected with 100% regardless of the key input success rate of the subject. This is largely because of applying a morphological filter that eliminates the effects of eyelashes, reflections and shadows using a Gabor filter [6] suitable for circular arc edge detection for blink detection [4].

Comparison of key input success rates betien proposed mobile keyboard and conventional fixed keyboard

The input character was 14 of "Saga University", and this was repeated ten times and the key input success rate was evaluated. The subjects had the same six people as in the evaluation of the pupil center detection rate. Among them, subjects 1, 2, 4, and 6 have trial experiences and the other subjects 3, 5 are beginners without trial experience. The key input success rate and required time at sentence input with proposed moving keyboard and conventional fixed keyboard Ire compared and evaluated. The results are shown in Tables 6 and 7.

Compared with Table 4, subjects 1, 2, 4, and 6 are experienced effects, and almost 100% successfully keyed using the moving

Table 7: Comparison of process time between fixed and moving keyboards.

User	Expertly	Moving Keyboard (s)	Fixed Keyboard (s)
1	Expert	117.5	154
2	Expert	138.67	195.33
3	Beginner	180.5	275
4	Expert	101	197.33
5	Beginner	161.5	213
6	Expert	109	227
Average		134.69	210.28

Table 8: Time required for moving along the track.

Trial No	Required Time (seconds)
1	177
2	202
3	192
4	182
5	198

keyboard. In addition, subjects 3 and 5 Ire never as high as 82.14%, 71.43%. Furthermore, these subjects never have high pupillary center detection rates. On the other hand, it can be seen that the time required for key input is 14 subjects who have experiences can enter 14 keys from 101 seconds to 138.67 seconds, while the beginner subjects take from 161.5 seconds to 180.5 seconds. Also, in the case of the fixed keyboard, since the key input success rate is low, it is forced to re-enter the key, and the required time is about 56.12% more than in the case of the moving keyboard.

Functional and performance tests

Figure 10 shows an example of an image captured by tracking a service robot in the lower right, and an image acquired by a camera attached to the tip of the service robot in the center of the figure.

The latter image is displayed on the display in front of the user. The user controls the running of the service robot by moving the line of sight up, down, left and right while watching this image. The running performance of the service robot also confirmed the running speed of 10.8m at the maximum. It was found that this is almost equivalent to the speed of 18cm/s.

When you find an obstacle, you can take evasive action beforehand. Obstacle detection has confirmed that it can be detected when an obstacle appears in about 1m by ultrasonic waves.

Usually, a user can detect an obstacle by looking at the camera image on the tip of the service robot, but in this case, obstacles such as transparent doors cannot be detected in some cases. Therefore, detection of obstacles by ultrasonic waves was also used [11]. Also at this time, since the service robot is stopped immediately after detection of the obstacle, the user can grasp the surrounding situation and perform the obstacle avoiding operation by the user's line of sight.

When the user looked at the top of the display, I confirmed that the menu mode and the automatic mode switched. Also, in the menu mode, I confirmed that I will run automatically to a predetermined destination. I also confirmed that switching to automatic mode will
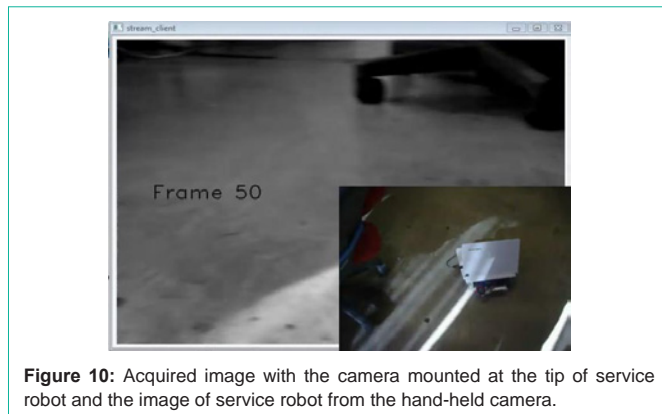
Figure 10: Acquired image with the camera mounted at the tip of service robot and the image of service robot from the hand-held camera.
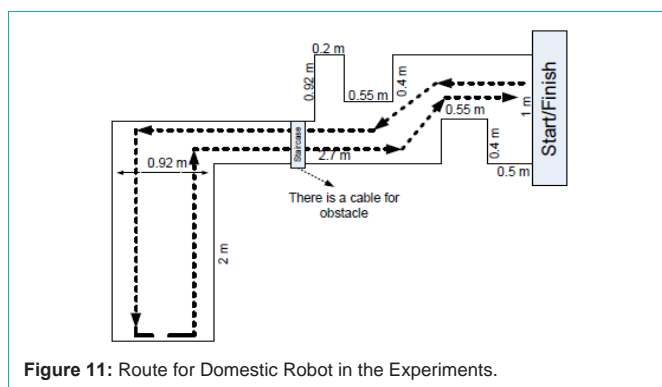

Figure 11: Route for Domestic Robot in the Experiments.

be made when viewing the upper right corner of the display in the menu mode. In this mode, when looking at the top left and right of the display, I confirmed that the service robot stops. Also, I confirmed that I will turn left and right when I see the left and right parts, and go straight forward when I see the center part.

When the user looks at the bottom of the display, it switches to the screen for inputting letters, creates a sentence and confirms that it can output sound by reading software. I also confirmed that the sound collected by the microphone provided in the WiFi terminal (computer 2) can be heard by the computer 1 terminal. Therefore, the user can enjoy a conversation with a person around the service robot.

It also confirmed that by returning to the screen in Figure 4a by selecting / determining the ESC key. Also, when the user looks at the bottom left and right of the display, I confirmed the function to return to the place where the user is (home position) by the floor layout stored in the service robot.

Moving performance tests

I tested the robot control performance by controlling it moves forward about two meters and return while I record the time. Before experiment, first I must connect the robot to local network via WIFI connection. Also, the PC station network must be established. After the connection between them was established, we began to start the server program and video streaming. Also, we start the program on client to receive the video and enable sending robot commands. The Map used for experiment is shown in Figure 11.

Screen shot image of moving picture of image which is acquired with the camera attached at the tip of robot is shown in Figure 11.

Also, the acquired image of hand held camera of which the author traced to the robot. Using this camera acquired image, user can avoid obstacles. Also, the robotics can avoid obstacles using the attached ultrasonic sensor, even if the obstacles are transparent (invisible with the camera).

The robotics control speed is evaluated. The time required for the route stating from the start location to the finished location is measured with five trials. The results are shown in Table 8.

The required time ranges from 177 to 202 seconds for traveling the route of the approximately 16m of travel length. Therefore, robot speed ranges from 7.92 to 9.04 cm/s. It is about half speed in comparison to the specification. This would be good enough for the robot which helps patients to make order, to travel in the hospital virtually, and enjoy conversations with the other persons when the robot meets with them.

I proposed a service robot controlled by sort of domestic robot, gaze. As a result, users such as patients, handicapped people, long-term care givers, elderly people and the like can experience the virtual trip within the facility in hospitals, group homes, etc., and can enjoy conversations one after the other, and doctors, nurses It will be able to tell intent and request business. I evaluated pupil center detection, eye blink detection, key input success rate, etc. for gaze estimation so that these can be reliably performed.

As a result, it was confirmed that key input is 100% surely at a speed of 9.62sec character confirmed. The running performance of the service robot was confirmed to be able to run while avoiding obstacles at a traveling speed of 10.8m/min. Furthermore, I confirmed functions such as stop of service robot, homing to home position, movement to a predetermined place as menu.

The proposed eye-based robot helper has been successfully implemented. The real-time remote controller has been successfully implemented and tested in our laboratory with the average time 182 seconds for robot travelling along the 5.75 meters. Also, our robot has good maneuver to avoid the obstacles and pass it. The user easily controls the robot via the transmitted scenery image. By implementing our robot in the real application, it could help the handicap patient when they are sick in the hospital.

For future work, there are trials by users such as patients, disabled people, long-term care recipients, elderly people, etc. within facilities such as hospitals, group homes. In particular, I am considering trials by people with disabilities who are suffering from visual impairment and who are impaired.

Acknowledgment

References

1. Bill Gates. A Robot in Every Home, The leader of the PC revolution predicts that the next hot field will be robotics. Scientific American Magazine. 2007.

2. Erico Guizzo. "10 stats you should know about robots but never bothered googling up". IEEE Spectrum. 2007.

3. Arai K, Hiromi Kamiaki. Computer input based on eye-gaze estimation by detecting the eye-center of a subject that allows movement of the subject's face. The Institute of Electrical Engineers of Japan. Transaction C. 2007; 127:

1107-1114.

4. Arai K, Makoto Yamaura. Improvement of blink detection success rate in key determination of gaze input system by morphological filter. Journal of the Institute of Image Electronics. 2008; 37: 601-608.

5. Arai K and R Mardiyanto. Camera mouse and keyboard for handicap person with trouble shooting capability, recovery and complete mouse events. International Journal of Human Computer Interaction. 2010; 1: 46-56.

6. Arai K and R Mardiyanto. Real time blinking detection based on Gabor filter. International Journal of Human Computer Interaction. 2010; 1: 33-45.

7. Arai K and M Yamaura. Computer input with human eyes only use two Purkinje images which work in a real time basis without calibration. International Journal of Human Computer Interaction. 2010; 1: 71-82.

8. Djoko P, R Mardiyanto and K Arai. Electric wheel chair control with gaze detection and eye blinking. Artificial Life and Robotics. AROB Journal. 2009; 14: 694,397-400.

9. Arai K, Kenro Yajima. Conversation support system with gaze input. Journal of the Institute of Electrical Engineers of Japan. 2008; 128: 1679-1686.

10. Arai K and K Yajima. Robot arm utilized having meal support system based on computer input by human eyes only. International Journal of Human Computer Interaction. 2011; 2: 1-9.

11. Arai K, R Mardiyanto. Electric wheel chair controlled by human eyes only with obstacle avoidance. International Journal of Research and Reviews on Computer Science. 2011; 2: 1235-1242.

12. Koren Y, Borenstein J. Potential field methods and their inherent limitations for mobile robot navigation. Proceedings of the Robotics and Automation. IEEE International Conference. 1991; 2: 9-11.

13. Yoshimi T, Nishiyama M, Sonoura T, Nakamoto H, Tokura S, Sato H. Development of a Person Following Robot with Vision Based Target Detection. Proceedings of the Intelligent Robots and Systems. IEEE/RSJ International Conference. 2006: 9-15.

14. Okada K, Tokutsu S, Hatao N, Inaba M, Inoue H. A humanoid platform for research on support behavior in daily life environments. Robotics and autonomous systems. 2009; 57: 145-154.

15. Ando N, Suehiro T, Kitagaki K, Kotoku T, Woo-Keun Yoon. RT-middleware: distributed component middleware for RT (robot technology) Proceedings of the Intelligent Robots and Systems. IEEE/RSJ International Conference on. 2005: 27-30.

16. Arai K, R Mardiyanto. Moving Keyboard for Eye-Based Human-Computer Interaction: HCI. Journal of Image Electronics Society of Japan. 2008; 41: 398-405.

17. Viola P, Jones M. Robust real time object detection. Proceedings of the 2nd International Workshop on Statistical and Computational and Theories of Vision-Modeling, Learning. Computation and Sampling. 2002.

18. Ohno Kenji, Takekawa Naoki, Yoshikawa Atsuko. Eye-gaze measurement system based on eye shape model. Information processing research report. 2001; HI-93: 47-54.

19. Kohei Arai, R Mardiyanto. Service robot which is controlled by human eyes only with voice communication capability. Journal of Image Electronics Society of Japan. 2012; 41: 535-542.