Guide Robot's Navigation Based on Attention Estimation Using Gaze Information

Yoshihisa Adachi, Hiroyuki Tsunenari, Yoshio Matsumoto and Tsukasa Ogasawara
Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama-cho, Ikoma city, Nara, 630-0192, Japan
Email: {yoshi-a, hiro-tsu, yoshio, ogasawar}@is.aist-nara.ac.jp

***Abstract*— In this paper, we describe a robotic wheelchair system as guide robot. To build a smart interface for a guide robot, it is important for the system to know the user's attention. This system can detect gaze direction of the user using a real-time stereo vision system, and can recognize its position within the surrounding environment using a range sensor and a map. Based on the detection of both the gaze direction and the position within the map, the robotic wheelchair can estimate the user's fixation, from which the position where the user is paying attention is estimated. We adopt an attention representation based on histogram. The frequency of fixation is represented by an "attention histogram". Experimental results indicate the validity of the control assistance to change the route of the autonomous run using the estimation of user's attention.**

I. Introduction

With the increase in the number of senior citizens, there has been a growing demand for human-friendly robots of boarding type as mobility aids. It is convenient that this kind of robot is used as a guide robot in the place, such as a library and a museum. However, driving conventional vehicles, for example, electric wheelchairs imposes burdens on their users both physically and mentally. They must do continuous steering with the joystick while paying close attention to the surroundings. Thus several intelligent/robotic wheelchairs have been proposed to alleviate these burdens.

Such researches address two major categories. The first one is the autonomous capability. Sensors such as encoders and ultrasonic range sensors are used to localize their own position and to detect obstacles[1][2]. The wheelchairs can navigate to a specified goal and/or avoid obstacles autonomously. However, in the place, such as a library and a museum, there are many attractive things during the run to the original goal. But the change of the user's interest is disregarded and the only navigation to the original goal will be performed. The second category is the interface for easy operation. To build simple and intuitive interfaces, multi-modal information can be utilized. In some researches, voice information[3][4] and electroencephalogram[5] are utilized to give commands to wheelchairs. Other researches utilized visual information from the human face[6][7][8][9]. The detected direction of the face is directly used as the direction to which the wheelchair moves. In a previous work, we utilized both the direction of the face and that of the gaze[10]. It showed that they almost coincided with each other when the user was concentrated in the operation. In the current operation using the face, the wheelchair slows down when the user is looking around. However, the user needs to turn his/her face exactly to the desired direction and burdens for the operation are imposed on the user.

Since the goal of our research is to build practical robotic wheelchairs, both aspects of the research are inevitable. To build a smart robotic wheelchair, it is important for the system to know the user's attention. Human pays attention continuously to surrounding environment and has acquired visual information in order to attain the purpose of movement. Since the visual information on human face, especially the motion of the head pose and gaze direction, is highly correlated to his/her attention. Therefore, the detection of such information can be utilized for attention estimation. The robot which performs autonomous movement to the original destination utilizes the estimated attention in order to estimate the user's intention which is different from the original purpose.

In this paper, we describe a robotic wheelchair system as guide robot. This system can detect gaze direction of the user using a real-time stereo vision system, and can recognize its position within the surrounding environment using a range sensor and a map. Based on the detection of both the gaze direction and the position within the map, the robotic wheelchair can estimate the user's fixation, from which the position where the user is paying attention is estimated. We adopt an attention representation based on histogram. Experimental results indicate the validity of the control assistance to change the route of the autonomous run using the estimation of user's attention.

II. System Configuration

In order to realize the intention recognition from the user's natural behavior, we developed our experimental robotic wheelchair system. Fig. 1 and Fig. 2 show the overview of our experiment system and the hardware architecture. We adapted a commercial electric wheelchair to be controlled by a notebook PC. The system has a pair of stereo CCD cameras to capture the facial image of the user and has a laser range finder and an encoder to recognize the position and the surrounding environment.

Fig. 1. *System overview.*

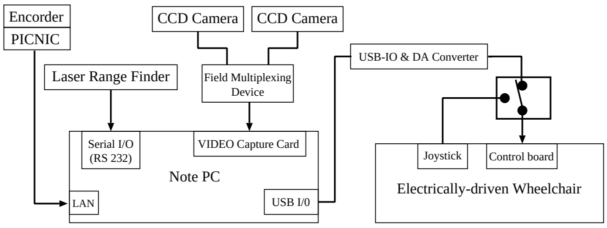

Fig. 2. *Hardware architecture of the control system.*

Fig. 3. *Stereo vision system.*

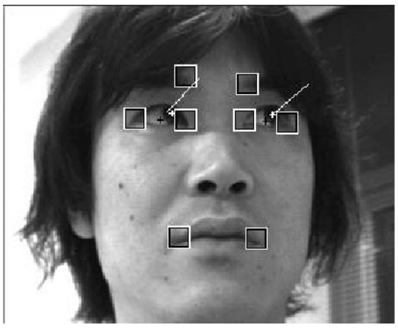

Fig. 4. *Result of face and gaze measurement*.

A. Face and Gaze Measurement System

Our system can detect the head pose and gaze direction using a stereo vision system [11] which has the following advantages: non-contact, passive, real-time, robust, accurate, and compact. It can detect natural behavior with no contact with the user. Fig. 3 and Fig. 4 show an overview of the stereo vision system and an example of the result of the measurement process. The squares indicate the positions of the features and the two lines indicate the gaze direction.

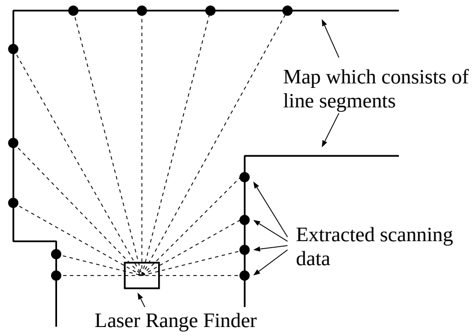

Fig. 5. Extracting a scanning data from the map.

B. Wheelchair Localization

The wheelchair can recognize the position and the surrounding environment using a laser range finder, an encoder and a map. The map consists of sets of line segments. The initial position within the environment is given. The robotic wheelchair observes front scanning data of 180 degrees using a laser range finder. Then, the wheelchair extracts the scanning data about each angle from the map (Fig. 5). Based on the matching of the two data, the wheelchair can estimate its position. The matching process utilizes a histogram-based method [12] which is an extension to the matching method proposed by Weiss et al [13]. The outline of the process is described as follows:

1) By connecting all consecutive points in a scan, the scan can be represented as a set of short lines. Then the direction of each line is calculated for all points in order to create a histogram which is called an “angle histogram”.
2) Calculate angle-histograms of two scans and find the angular displacement $d\theta$ with the highest correlation.
3) Rotate the current scan by $d\theta$.
4) Calculate translation histograms of two scans along x and y directions.
5) Find the highest correlations dx, dy for x and y directions using histograms.

Based on the same matching method, it is shown that the localization error is 5.8cm in the experiment of the robot navigation[14]. The encoder is utilized to decide the position to extract the data from the map in order to cope with the environment with less features because the proposed matching method requires flat walls which have different directions.

C. Estimation of User’s Attention

Based on the detection of both the gaze direction and the position within the map, the robotic wheelchair system can estimate the user’s fixation. The map consists of line segments as shown in Fig. 5. Each line corresponds to a wall etc. in the indoor environment. We assume the environment to be composed of only vertical planes to calculate the fixation point of the user. The fixation point is defined as the intersection of the gaze vector with the vertical planes as shown in Fig. 6. From the estimation of the fixation, the system estimates where the user is paying attention during the run. We adopt an attention representation based on histogram. The frequency of fixation is represented by an "attention histogram" over the vertical plane.

The following facts are used to define the attention histogram's parameters:

- It is known that human being's acuity of vision is about the central two degrees of visual angle, that is called Foveal Vision. The relation between the visual angle and eyesight is shown in Fig. 7. Therefore the attention histogram has a distribution around the gaze direction. In consideration of the accuracy of gaze measurement system (3 degrees), that distribution is assumed as the normal distribution and the standard deviation is defined as 5 degrees. This distribution is added to the amount histogram for every time.
- The attention to the position where the user is not looking attenuates gradually. Therefore, the frequency of the attention histogram in each position is attenuated over time and this is represented by using the rate of attenuation. In human memory, there is a sensory memory for carrying out short-time maintenance of the information. The sensory memory to a vision stimulus is called iconic memory. The duration of the memory is less than about 500ms[15]. Processing time in this experiment system is performed in about 210ms per frame. Therefore, the rate of attenuation is set to 0.5 considering that the frequency of histogram declines to less than 10% after the duration of the memory (3 frames).

The definition formula of the frequency is summarized as follows:

$$w_{ij}(t) = \alpha \cdot w_{ij}(t - \Delta t) + g_{ij}(t) \quad (1)$$

$$g_{ij}(t) = A \cdot \exp(-\frac{\phi^2}{2\sigma^2}) \quad (2)$$

$w_{ij}(t)$:	Frequency of the position P_{ij} at time t
α:	Rate of the attenuation of frequency
ϕ:	Angle between $\overrightarrow{OP_{ij}}$ and the gaze vector
σ:	Standard deviation of normal distribution
A:	A constant

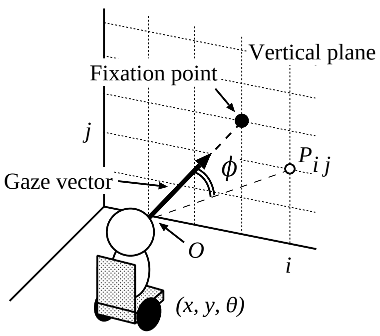

Fig. 6. Fixation point of the user.

At time $t + \Delta t$, the frequency at time t is attenuated and the gaze direction of the user changes. And the new frequency at time $t + \Delta t$ is added. The frequency of the histogram gets larger when the user is looking at the near position. In the case that the user continues looking at the same position, the frequency converges to a constant value as follows:

$$\begin{aligned} w_{ij}(t) &= \alpha \cdot w_{ij}(t - \Delta t) + A \\ &= \alpha \cdot (\alpha \cdot w_{ij}(t - 2\Delta t) + A) + A \\ &= (1 + \alpha + \alpha^2 + \cdots + \alpha^n + \cdots) \cdot A \\ &= \lim_{n\to\infty} \frac{1-\alpha^n}{1-\alpha} \cdot A = \frac{A}{1-\alpha} \\ &(\alpha < 1) \end{aligned} \quad (3)$$

The convergence value is defined as the maximum (100%) of the user's attention. The system regards the user as having started paying attention when the frequency becomes more than 90%, and having stopped paying attention when the frequency attenuates to less than 10%. When the attenuation ratio is set to 0.5, the durations to reach up to 90% from 0% and to reach down to 10% from 100% are both 0.7 seconds. Therefore, even if the user looks off the position for less than 0.7 seconds, the robot keeps regarding the user as paying attention to the same position.

III. Experimental Results

We made experiments in which the attention of the user and the information on the environment are recorded. In the experiments, the user rides on the robotic wheelchair while this one is running autonomously. The robotic wheelchair was equipped with a smart interface that controls the route based on the estimation of user's attention. When the attention of the user is estimated, the wheelchair approaches the position where to pay attention so that the user may have enough time to fully perceive the object. Then, the system selects the route according to the size of the frequency of the histogram. Fig. 8 shows the topological map for the autonomous run in the experimental environment. The nodes and the arcs are set between the start position and the goal position. There are two routes to the goal. The robotic wheelchair basically runs toward the goal by taking the shortest path. Two nodes near mailboxes are set between

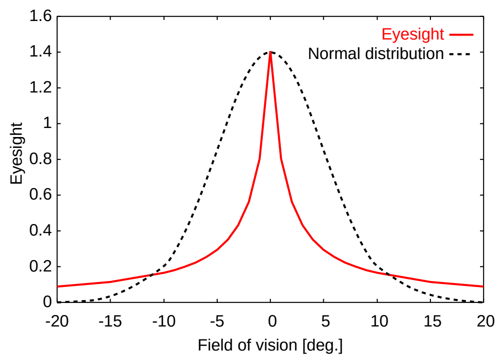

Fig. 7. Normal distribution added to the histogram.

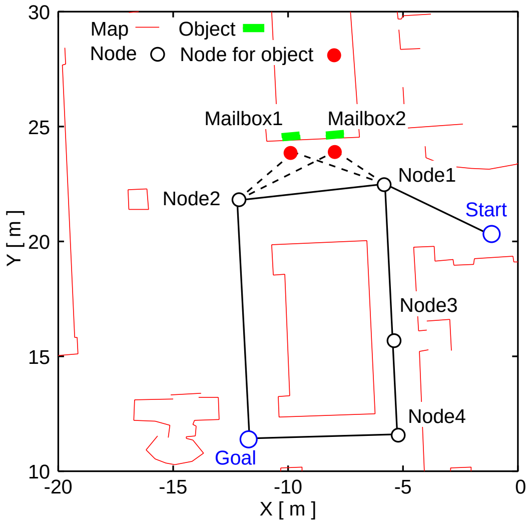

Fig. 8. The topological map for the navigation.

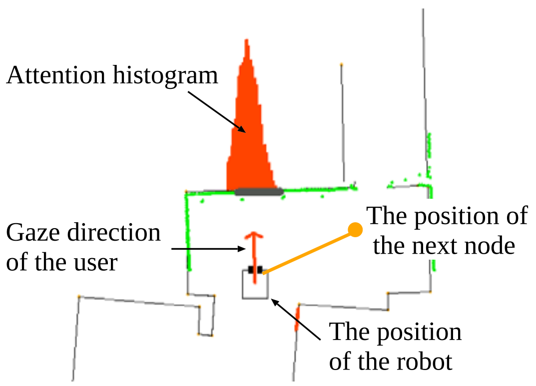

Fig. 9. The state in the experiment.

node 1 and node 2. After the robot arrives at node 1, it estimates whether the user is paying attention to the mailboxes or not and select the route. The normal path is from the start position to the goal position via node 1, node 3 and node 4. However the robot estimates the user is paying attention to the mailboxes, it changes the route and selects the node in front of the mailbox as the next node to go. When the robot estimates that the user stops paying attention to the mailbox, it calculates the shortest path to the goal and selects the next node.

Fig. 9 shows measured gaze direction of the user, the estimated position of the robot and the position of the next node to go. In Fig. 9, a rectangle indicates the position of the robot, and the dots indicate the scans obtained from the laser range finder. The arrow indicates detected gaze direction. The attention histogram is obtained by projecting the maximum frequency at the vertical plane on the map. The frequency of the histogram on the map is defined as follows:

$$w_i(t) = \max_j (w_{ij}(t)) \tag{4}$$

Fig. 10 shows the snapshots of the user and the wheelchair system and the states of the detected information in the case that the user is not paying attention. From the measurements of the facial and environmental information in Fig. 10(b) , it can be recognized that the user is not paying attention to specified object in this sequence. The frequency of the histogram is relatively small at each position and is distributed. After arriving at node 1, the robot selects node 3 as the next node because the robot estimates that the user is not paying attention to the mailboxes. Fig. 11 shows the route of the navigation in this case.

Fig. 12 shows the snapshots of the user and the wheelchair system and the states of the detected information in the case that the user is paying attention to the mailbox. From the measurements of the facial and environmental information in Fig. 12(f) and (g), it can be noticed that the user is paying attention to the mailbox. In Fig. 12(g), the frequency of the histogram gets large and is concentrated on the mailbox 1. After arriving at node 1, the robot changes the route and selects the node in front of the mailbox 1 as the next one because the robot estimates that the user is paying attention to the mailbox 1. When the robot estimates that the user stops paying attention to the mailbox in Fig. 12(h), it selects node 2 as the next node calculating the shortest path to the goal. Fig. 13 shows the route of the navigation in the case of paying attention to the mailbox 1.

In Fig. 14 and Fig. 15, the changes of the frequency of the histogram on the positions of mailboxes are shown. Fig. 14 shows the result in the case that the user is not paying attention to the mailboxes. The frequency of the histogram is relatively small at each position and has not reached up to 90%. Fig. 15 shows the result in the case that the user is paying attention to the mailbox 1. This shows that the system can distinguish the attention to the mailbox 1 from the one to the mailbox 2. In this case, the robot arrives at node 1 at 17 seconds and the frequency of the histogram is then concentrated on the mailbox 1 and has reached up to 90% till the user stops paying attention to the mailbox 1 at 31 seconds.

These results indicate the validity of the validity of the control assistance to change the route of the autonomous run using the estimation of user's attention.

IV. Summary

In this paper, we described a robotic wheelchair system which could detect the head pose and gaze direction using a real-time stereo vision system, and could recognize its position within the surrounding environment using a range sensor and a map. Based on the detection of both the gaze direction and the position within the map, the robotic wheelchair can estimate the user's fixation, from which the position where the user is paying attention is estimated. We adopted an attention representation based on histogram. Experimental results indicate the validity of the control assistance to change the route of the autonomous run using the estimation of user's attention.

As future works, we are planning to extend the application area of the wheelchair to various situations. For such an extension, we will start by measuring the fixation and the user's operation at various situations, and then analyze the relationship between them in order to estimate the user's attention.

Acknowledgment

This work is supported in part by 21st Century COE Program "Ubiquitous Networked Media Computing".

References

[1] D. P. Miller and M. G. Slack, "Design and testing of a low-cost robotic wheelchair prototype", Autonomous Robotics, pp. 77-88, Vol. 2, 1995.

[2] E. Prassler, J. Scholz and P. Fiorini "Maid: A robotic wheelchair roaming in a railway station",International Conference on Field and Service Robotics, pp. 31-36, 1999.

[3] R. C. Simpson and S. P. Levine "Adaptive shared control of a smart wheelchair operated by voice control", Proceedings of IEEE International Conference on Intelligent Robots and Systems, pp. 622-626, 1997.

[4] N. Katevas, N. Sgouros, S. Tzafestas, G. Papakonstantinou, P. Beattie, J. Bishop, P. Tsanakas and D. Koutsouris "The autonomous mobile robot SENARIO: a sensor-aided intelligent navigation system for powered wheelchairs", IEEE Robotics & Automation Magazine, pp. 60-70, Vol. 4, No. 4, 1997.

[5] K. Tanaka, K. Matsunaga, N. Kanamori, S. Hori and H. O. Wang: Electroencephalogram-based control of a mobile robot, Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 688-693 (2003).

[6] Y. Adachi, N. Shimada, Y. Kuno and Y. Shirai "Intelligent wheelchair Using Visual Information on Human Faces", Proceedings of IEEE International Conference on Intelligent Robots and Systems, pp. 354-359, 1998.

[7] S. Nakanishi, Y. Kuno, N. Shimada and Y. Shirai "Robotic wheelchair based on observations of both user and environment", Proceedings of IEEE International Conference on Intelligent Robots and Systems, pp. 912-917, 1999.

[8] L. M. Bergasa, M. Mazo, A. Gradel, M. A. Sotelo and J. C. Garcia "Guidance of a Wheelchair for Handicapped People by Head Movements", International Conference on Field and Service Robotics, pp. 150-155, 1999.

[9] I. Moon, S. Joung and Y. Kum "Safe and Reliable Intelligent Wheelchair Robot with Human Robot Interaction", Proceedings of IEEE International Conference on Robotics and Automation, pp. 3595-3600, 2002.

[10] Y. Matsumoto, T. Ino and T. Ogasawara "Development of Intelligent Wheelchair System with Face and Gaze Based Interface", Proceedings of 10th IEEE/RSJ International Workshop on Robot and Human Comunication (ROMAN2001), pp. 262-267, 2001.

[11] Y. Matsumoto, T. Ogasawara and A. Zelinsky "Behavior Recognition Based on Head Pose and Gaze Direction Measurement", Proceedings of IEEE International Conference on Intelligent Robots and Systems, pp. 2127-2132, 2000.

[12] Y. Adachi, K. Goto, Y. Matsumoto and T. Ogasawara: Development of Control Assistant System for Robotic Wheelchair - Estimation of User's Behavior based on Measurements of Gaze and Environment -, Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2003.

[13] G. Weiss, C. Wetzler and E. V. Puttkamer "Keeping Track of Position and Orientation of Moving Indoor Systems by Correlation of Range-Finder Scans", Proceedings of IEEE International Conference on Intelligent Robots and Systems, pp. 595-601, 1994.

[14] Y. Adachi, H. Saito, Y. Matsumoto and T. Ogasawara: "Memory-Based Navigation using Data Sequence of Laser Range Finder", Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2003.

[15] G. Sperling "The information availavle in brief presentations", Psychological Monographs, 74 (Whole No.498), 1960.

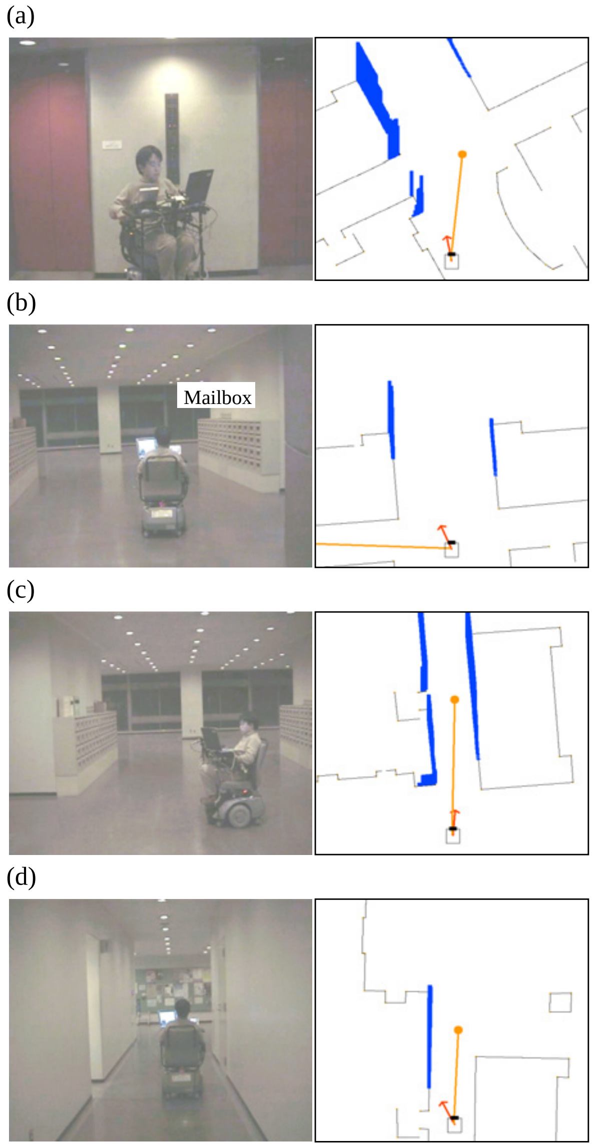

Fig. 10. Navigation in the case that the user is not paying attention.

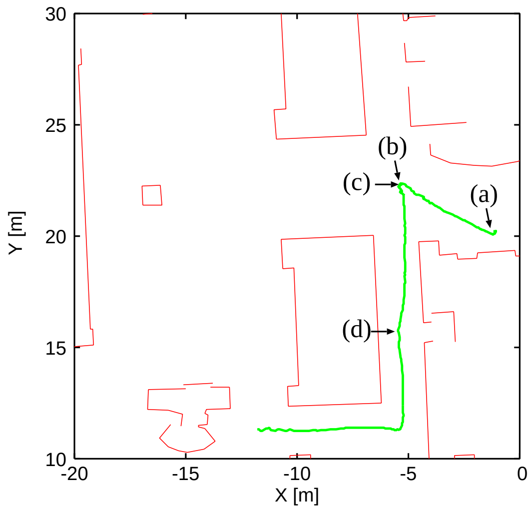

Fig. 11. The route in the case that the user is not paying attention.

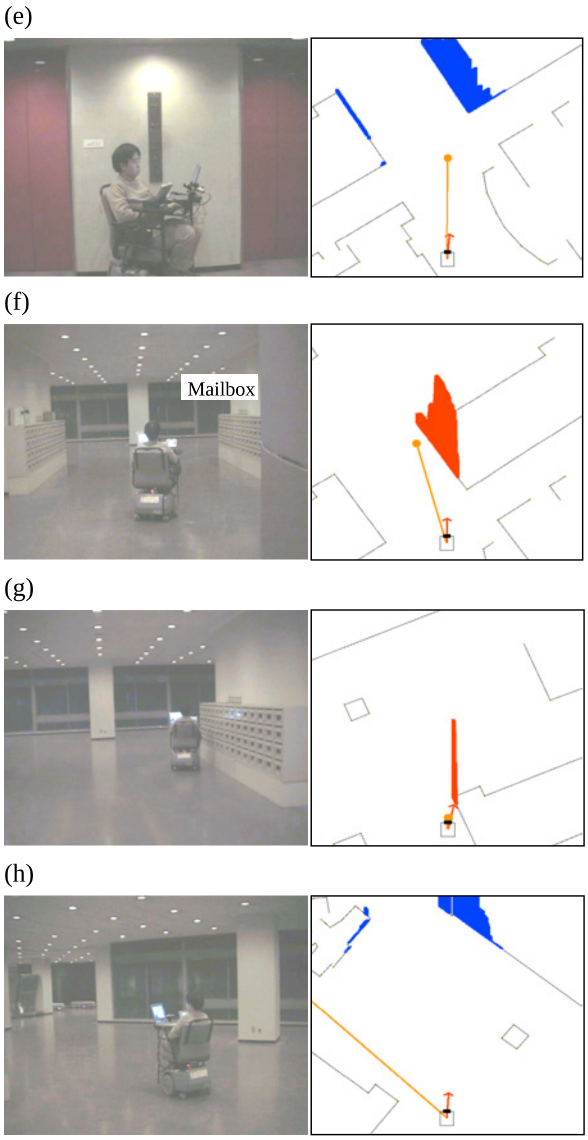

Fig. 12. Navigation in the case that the user is paying attention.

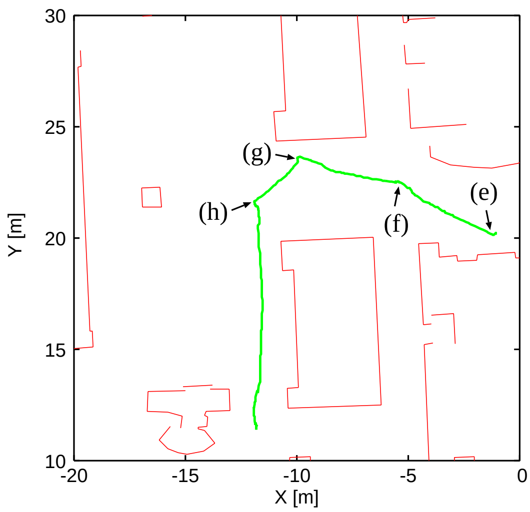

Fig. 13. The route in the case that the user is paying attention.

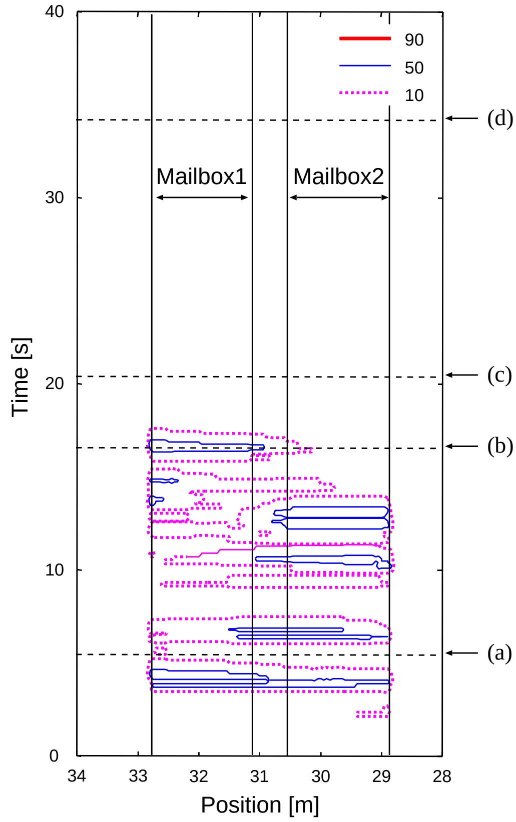

Fig. 14. The frequency in the case that the user is not paying attention.

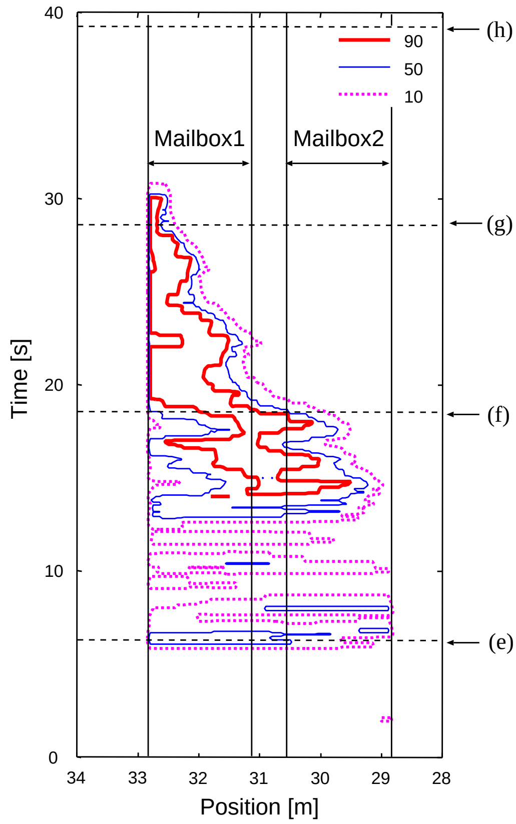

Fig. 15. The frequency in the case that the user is paying attention.