Development of Control Assistant System for Robotic Wheelchair - Estimation of User's Behavior based on Measurements of Gaze and Environment -

Yoshihisa Adachi, Kenji Goto*, Yoshio Matsumoto, Tsukasa Ogasawara
Graduate School of Information Science
Nara Institute of Science and Technology
http://robotics.aist-nara.ac.jp/
Sony Corporation*

Abstract

We are developing a control assistant system for a robotic wheelchair as a guide robot. This system detects the head pose and gaze direction of the user and recognizes a position and surrounding environment using sensors and a map. Since the system can detect where the user is looking from the measurements, it can estimate attentions of the user on the wheelchair. Experimental results prove that the user's intentions and attentions are related with the duration of gaze and can be distinguished by the information which our robotic wheelchair system can measure.

1 Introduction

With the increase in the number of senior citizens, there has been a growing demand for human-friendly wheelchairs as mobility aids. However, driving conventional 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 can be classified into major two categories. One is on autonomous capabilities. 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. The other is on interfaces for easy operation. To build simple and intuitive interfaces, multi-modal information can be utilized. In some researches, voice information is utilized to give commands to wheelchairs[3, 4]. Other researches utilized user's face direction to directly specify the direction to move. Since the goal of our research is to build practical robotic wheelchairs, both aspects of the research are inevitable. However in this paper, we are focusing on describing the interface aspects of our research. To build a smart interface for a robotic wheelchair, it is important for the systems to know the user's intention and attention. Since the visual information from human face, especially the motions of the head pose and gaze direction, is deeply related with his/her intention and attention, detection of such information can be utilized for a natural and intuitive interfaces. In the conventional researches on the interfaces of robotic wheelchair using visual information on human face[5, 6, 7, 8], the detected direction of the face is directly used as the direction to which the wheelchair moves. Therefore, human needs to turn his/her face to the direction to go intentionally and there is no significant result regarding the intention recognition of the user.

In our previous research, it was shown that both the direction of the face and that of the gaze almost coincided with each other when the user was concentrated in the operation[9]. In the actual operation using the face, the wheelchair slows down when the user is looking around. In this paper, we describe a robotic wheelchair system which can detect the head pose and gaze direction and can recognize a position and surrounding environment using sensors and a map. Since the system can detect where the user is looking from the measurements, it can estimate attentions of the user on the wheelchair. Experimental results prove that the user's intentions and attentions are related with the duration of gaze, and can be distinguished by the information which our robotic wheelchair system can measure.

2 System Configuration

Toward the realization of the intention recognition from the measurements of natural behaviors of the user, we are developing the experimental robotic wheelchair system. **Figure** 1 and **Figure** 2 show the overview of our experiment system overview and the hardware architecture. We adapted a commercial electric wheelchair for computer control by installing a notebook PC. It has a stereo CCD camera pair to capture the facial image of the user, and a laser range finder to recognize the sur-

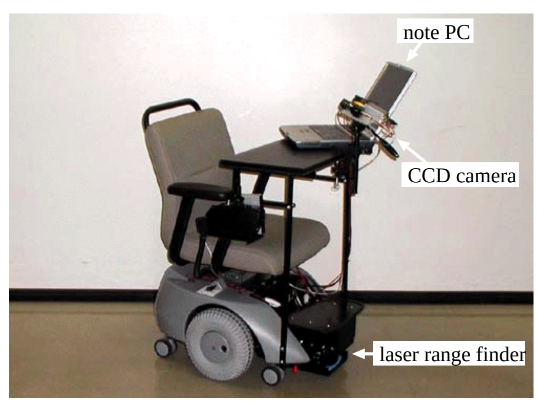


Figure 1: System overview

Table 1: The specification of our system

Electric wheelchair	M-Smart (MISAWA)
PC	Pentium III 800MHz
Laser range finder	LMS 200 (SICK)

Table 2: The specification of laser range finder

The measurement range of distance	Max 80[m]
Resolution of the angle	0.5[deg]
The range of angle	0∼180[deg]
The number of data	361
Response time	26[ms]
Resolution of the measurement	1.0[cm]
System error	±4.0[cm]

rounding environment. **Table** 1 shows the specification of our system and **Table** 2 shows the specification of the laser range finder in the system.

2.1 Face and Gaze Measurement System

Our system can detect the head pose and gaze direction using a stereo vision system [10] which has following advantages: non-contact, passive, real-time, robust, accurate, and compact. It can detect natural behaviors without applying a burden to a user. **Figure** 3 shows an example of the result of the measurement process. The squares indicate the positions of the features and two lines indicate the gaze direction.

2.2 Localization of Robot

The wheelchair can also recognize the position and surrounding environment using a laser range finder and a map. The map consists of sets of line segments. **Figure** 4 shows an example of the map. The initial position within environment is given. The robotic wheelchair observes front scanning data of 180 degrees using a laser range finder. Then, the wheelchair extracts scanning data at each angle on line segments of the map in the previous position and posture. **Figure** 5 shows the extraction of scanning data from the map. Based on the matching of two data, the wheelchair can detect the position. The matching process utilizes a histogram-based method which is an extension of the matching method proposed by Weiss et al [11]. The outline of the process is described as follows:

1. To calculate angle-histograms of two scans and find the angular displacement $d\theta$ between them which gives the highest correlation.
2. To rotate the current scan by $d\theta$.
3. To calculate translation histograms of two scans for x and y directions.
4. To find the highest correlations dx, dy for x and y directions using histograms.

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." If a scan contains a straight line segment in a environment such as a wall, all the angles for the

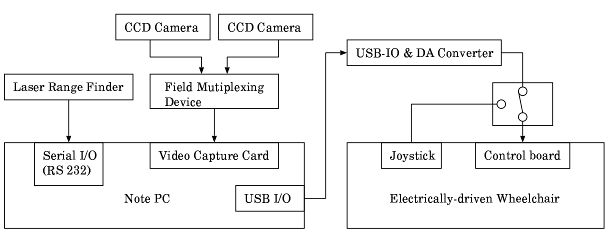


Figure 2: Hardware architecture of the control system

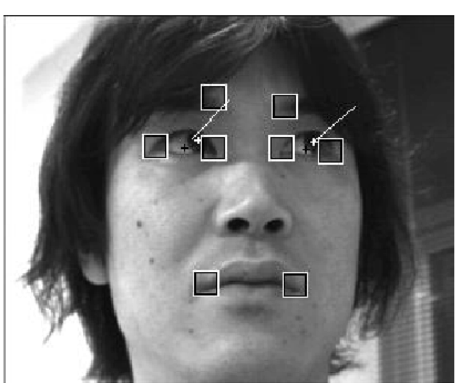


Figure 3: Result of Face and Gaze measurement

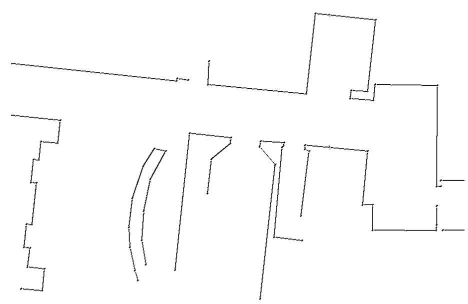


Figure 4: A Map which consists of line segments

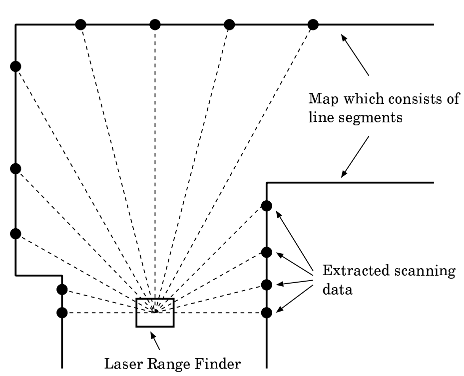


Figure 5: Extracting a scanning data from the map

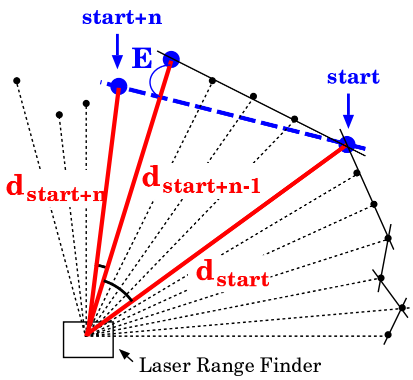


Figure 6: Line segments detection

segment should be the same. However, noises in the scan and the quantization error prevents them from being the same. Since the line segments are quite short, small errors in range data produces relatively large errors with regard to the direction. Thus the angle-histogram tends to be noisy, and grouping of the points which belongs to a line segment is performed before making a histogram.

For grouping points, the value E which is defined as the following equation is utilized:

$$E = d_{start+n-1} - \frac{d_{start} d_{start+n} \sin n\theta}{d_{start} sin(n-1)\theta + d_{start+n} \sin\theta} \quad (1)$$

As shown in **Figure** 6, E represents the distance from a point to a line segment. If the value E is under a threshold, points from $start$ to $(start + n)$ are regarded as linear, and the number of points in the line is accumulated in the angle-histogram $h(x)$. The average filter is also applied to the angle-histograms in order to obtain smooth correlation results. The rotational displacement between the current scan and the extracted scanning data can be obtained by calculation of the cross-correlation of two angle-histograms h_1(current scan) and h_2(extracted scanning data). **Figure** 7(a) shows an example of two angle-histograms, i.e. a current scan and an extracted scanning data. The correlation of them is defined as follows:

$$c(j) = \sum_{i=1}^{N+1} h_1(i) \cdot h_2(i+j) \quad (2)$$

By finding the highest correlation, the rotational displacement between two scans can be obtained. **Figure** 7(b) illustrates the correlation between two angle-histograms shown in **Figure** 7(a). Then one of the angle-histogram is rotated by the obtained rotational displacement $d\theta$. After the orientation of the scans are aligned, the translational displacements in x and y direction, dx, dy are calculated by using x-histogram and y-histogram in the similar manner. **Figure** 8(a) and (b) show an example of the current scan and the extracted data from the map, respectively. **Figure** 8(c) shows the result of matching between the current scan and the extracted data from the map.

2.3 Estimation of User's Attention

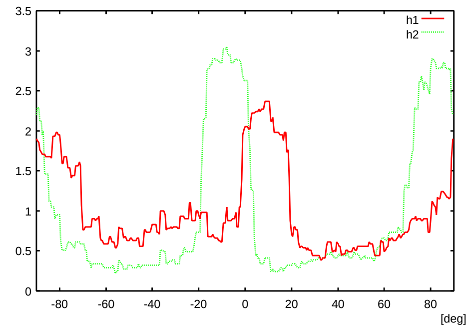
Based on both of the detection of the gaze direction and the one of position within the map, the robotic wheelchair system can detect the gaze point of the user within the environment. Then, it estimates where the user is paying attention during the run based on the length of the time when the user is gazing the same points. Our system utilizes the histogram on the map whose frequency is dependent upon the time of gaze.

The following knowledge is used to define the histogram.

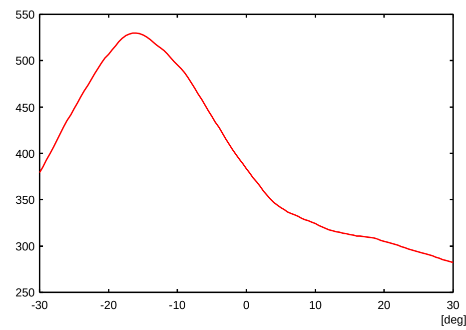
- In the field of view of human, the range of the center of the visual field which has very high spatial resolution is about 2 degrees, and the one of useful field of view is from 4 to 20 degrees. Thus, the frequency of the histogram is enlarged as the gaze direction is approached.
- The attention on each position decreases as the distance from the user becomes large. Thus, the frequency is decreased according as the distance.
- The frequency of the histogram in each position of the map is attenuated.

The definition formula of the fixation is summarized as follows:

$$w_i(t) = \alpha \cdot w_i(t - \Delta\, t) + \begin{cases} 1/L & (\theta \le 5[deg]) \\ 0.5/L & (5 < \theta \le 10[deg]) \\ 0 & (\theta > 10[deg]) \end{cases}$$

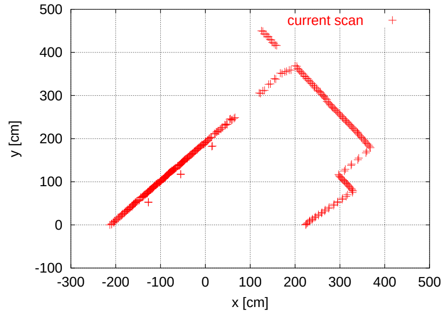


(a) two angle-histograms

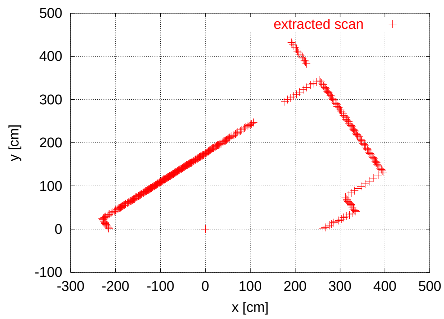


(b) correlations between two angle-histograms

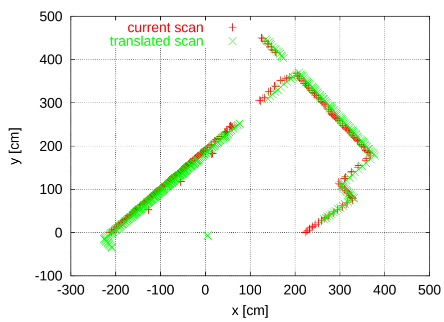
Figure 7: The correlation between two scans



(a) A current scan



(b) An extracted data from the map



(c) Matching two scans

Figure 8: The result of matching between a current scan and an extracted data from map

$w_i(t)$: Frequency of the position i at time t
L : Distance between the position i and the wheelchair
α : Rate of the attenuation of frequency
θ : Angle from the gaze direction of the user

3 Experimental Results

For the intention recognition, we made a preliminary experiment in which the attention of the user and the information on the environment are recorded. In the experiment, the user rides on the wheelchair in the autonomous run, with gazing at a poster on the wall and without paying attentions to the poster.

Figure 9 shows the route of the autonomous run in the experimental environment. **Figure** 10 shows a snapshot of the user and the robotic wheelchair in the experiment. **Figure** 11 shows measured face and gaze direction of the user and the detected position of the wheelchair. In **Figure** 11, the rectangle indicates the position of the wheelchair and the dots indicate the scans from a laser range finder. The long arrow indicates the detected direction of the face and the short arrow indicates the gaze direction respectively. The histogram on the map indicates the histogram of fixation at each position.

Figure 12 shows the sequence of measured face and gaze direction during the autonomous run. From these measurements of the facial and environmental information in **Figure** 12(a) , it can be noticed that the user is paying attention to a poster on the wall. In **Figure** 12(a-2) and (a-3), the frequency of the histogram gets larger and is concentrated on the poster. In contrast to this, the user was not gazing at specified object. From the measured results shown in **Figure** 12(b), it can be easily 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.

Figure 13 and **Figure** 14 show the changes of maximum frequency of the histogram and the position of

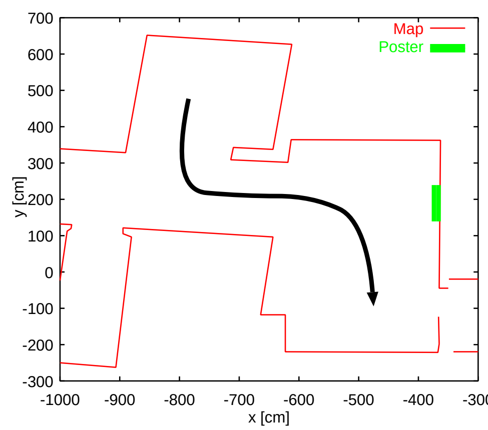


Figure 9: The route in the autonomous run

Figure 10: User in the experiment

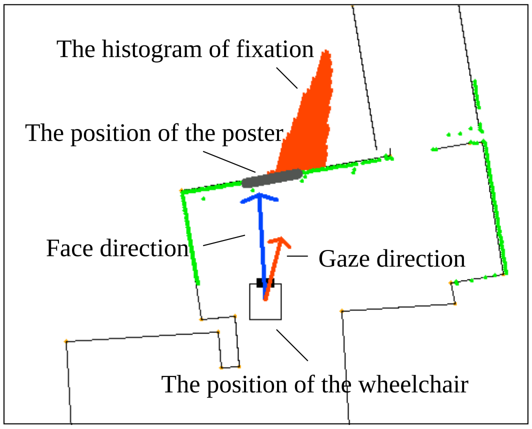


Figure 11: The state in the experiment

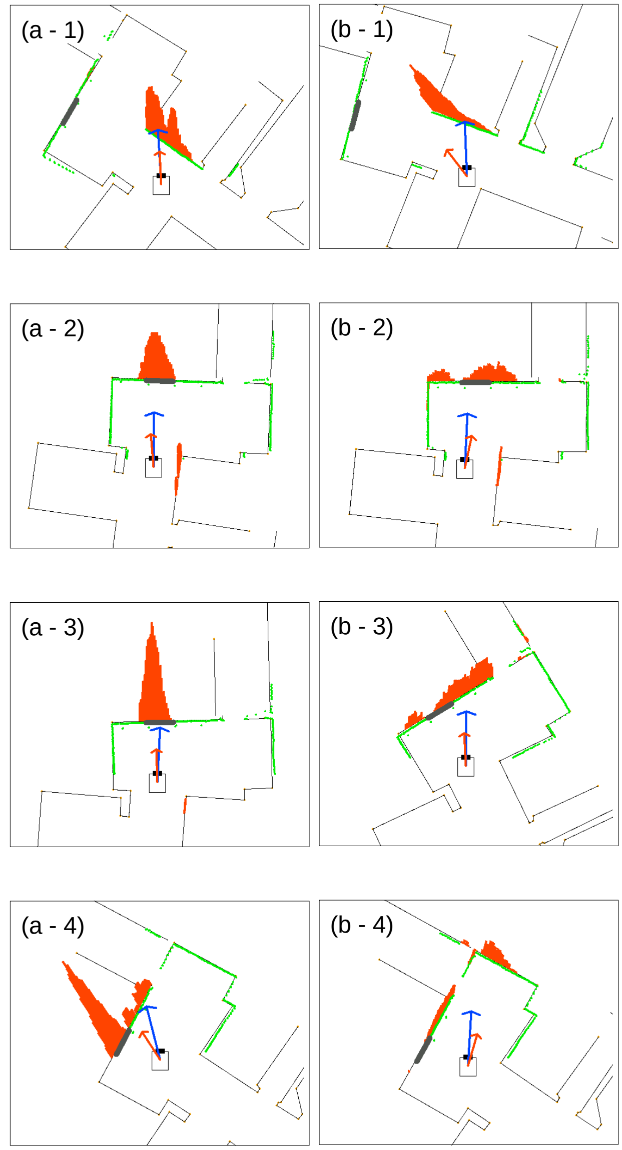
maximum frequency in the map. **Figure** 13 corresponds to the measured results in the experiment of **Figure** 12(a). From the measurements in **Figure** 13(a), it can be noticed that the user is paying attention to the poster since the position of maximum frequency is stable on the poster. **Figure** 13(b) shows that the frequency is large when the user is paying attention to the poster. **Figure** 14 corresponds to the measured results in the experiment of **Figure** 12(b). From the measured results shown in **Figure** 14(a), it can be noticed that the position of maximum frequency is unsteady. **Figure** 14(b) shows that the frequency is small throughout the experiment.

These experimental results indicate that the user's intentions and attentions are related with the direction of both the gaze and the face, and can be distinguished by the information which our robotic wheelchair system can measure. As a next step, we are going to develop an interface which recognize the user's intention and attention. For instance, we are planning to develop a robotic wheelchair system which adjusts its speed according to the estimated attention of the user in the autonomous run.

4 Summary

In this paper, we described a robotic wheelchair system which can detect the head pose and gaze direction using a real-time stereo vision system and can recognize a position and surrounding environment using sensors and a map. Since the system can detect where the user is looking from the measurements, we can estimate intentions of the user on the wheelchair. Experimental results prove that the user's intentions and attentions are related with the duration of the gaze and can be distinguished by the information which our robotic wheelchair system can measure. As an example, while the wheelchair is carrying out the autonomous run, the assist which adjusts the speed of the wheelchair according to the duration of gaze can be considered.

As the future works, we are planning to extend the applicable area of the wheelchair to various situations other than "a poster on the wall." For such extension, we will begin with measuring the gaze point and the user's operation at various situations, and analyze the relationship between them in order to estimate the user's attention.



(a) Paying attention to the poster (b) Not paying attention to the poster

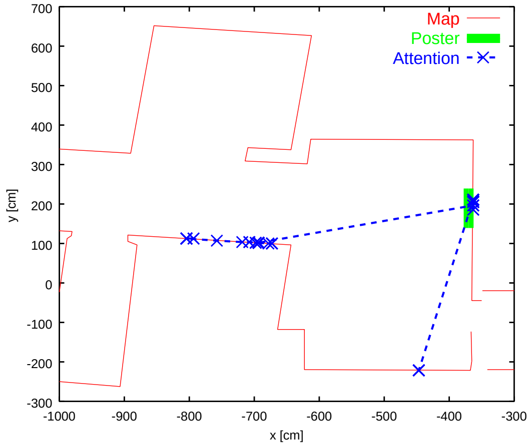
Figure 12: The sequence of measured face and gaze direction and localization during the autonomous run

References

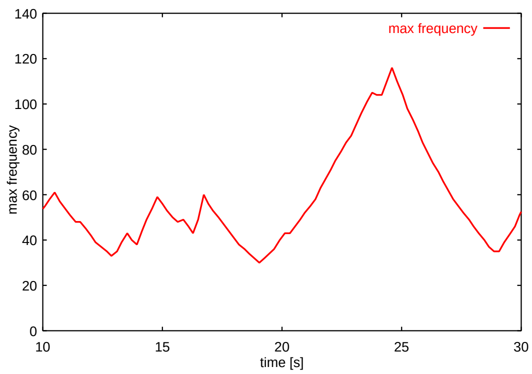
[1] D. P. Miller and M. G. Slack. Design and testing of a low-cost robotic wheelchair prototype. In *Autonomous Robotics*, volume 2, pages 77–88, 1995.

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

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

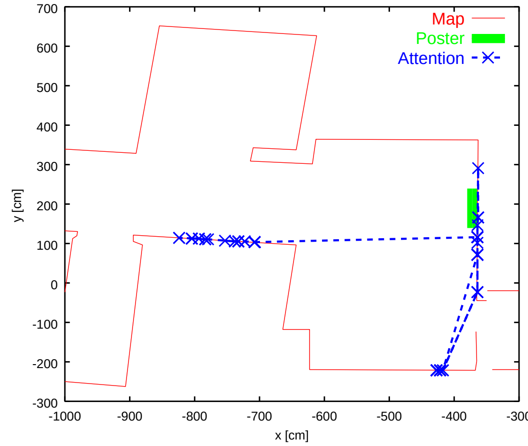


(a) The position of estimated attention

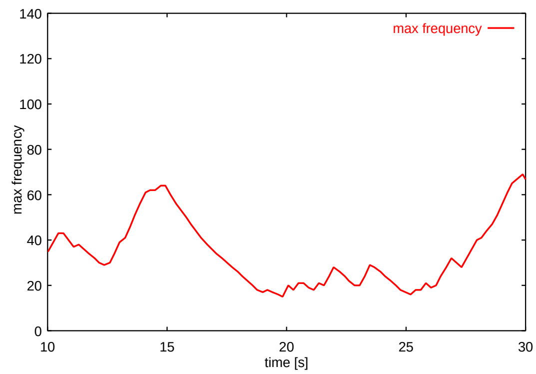


(b) The sequence of maximum frequency

Figure 13: The maximum frequency and the position of attention in the case that the user is paying attention to the poster



(a) The position of estimated attention



(b) The sequence of maximum frequency

Figure 14: The maximum frequency and the position of attention in the case that the user is not paying attention to the poster

[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. In *IEEE Robotics & Automation Magazine*, volume 4, pages 60–70, 1997.

[5] Y. Adachi, N.Shimada, Y. Kuno, and Y. Shirai. Intelligent wheelchair using visual information on human faces. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 354–359, 1998.

[6] S. Nakanishi, Y. Kuno, N. Shimada, and Y. Shirai. Robotic wheelchair based on observations of both user and environment. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 912–917, 1999.

[7] L. M. Bergasa, M. Mazo, A. Gradel, M. A. Sotelo, and J. C. Garcia. Guidance of a wheelchair for handicapped people by head movements. In *International Conference on Field and Service Robotics*, pages 150–155, 1999.

[8] I. Moon, S. Joung, and Y. Kum. Safe and reliable intelligent wheelchair robot with human robot interaction. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 3595–3600, 2002.

[9] Y. Matsumoto, T. Ino, and T. Ogasawara. Development of intelligent wheelchair system with face and gaze based interface. In *Proceedings of 10th IEEE/RSJ International Workshop on Robot and Human Comunication (ROMAN2001)*, pages 262–267, 2001.

[10] Y. Matsumoto, T. Ogasawara, and A. Zelinsky. Behavior recognition conference on intelligent robots and systems. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 2127–2132, 2000.

[11] G. Weiss, C. Wetzler, and E. V. Puttkamer. Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 595–601, 1994.