34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Measuring Cognition Delay Caused by Gaze Movement toward the Direction of Depth

Md. Shoaib Bhuiyan, *Member, IEEE*, Kenji Takagi, Haruki Kawanaka, and Koji Oguri, *Member, IEEE*

*Abstract***— It is difficult to judge from face images and/or from biological signal (such as brain waves) what the driver's state of cognition is. This paper reports experiments measuring the state of cognition accompanied with gaze movement in the direction of depth. While measuring the cognition time, the experimental setup was able to measure the effect of gaze movement in the direction of depth. When the locations of displays were changed, there was no significant difference in the reaction time for the reflective / discriminative experiment. Experiments conducted to know gaze depth movement show a trend where the longer the gaze moving distance becomes, the more the discernment time increases. Results also show that it takes much more time in moving gaze *toward* far in the direction of depth as opposed to moving toward the near.**

I. Introduction

Active safety systems demand to know the state of drivers' cognition. There are some situations when a driver does not pay much attention to objects even though they come into his or her view. So, the oversight detection system needs to understand this thing which is coming into driver's field of view but is not recognized while s/he is driving. How can we know? It is very difficult to judge from face images and /or from biological signal (such as brain waves) what the driver's state of cognition is.

Here, we formulate three hypotheses about the relation between a field of view, gaze, and cognition. Hypothesis 1: A driver responds to the stimulus which occurs in the peripheral vision. Hypothesis 2: The driver moves a gaze in order to recognize the source of the stimulus. Hypothesis 3: The driver recognizes it after moving eye and gazing. In various conventional studies, these are assumed tacitly [1][2]. Although these hypotheses are a little aggressive, it is led from these that a driver recognizes various peripheral hazards in moving the view direction and gazing.

The human field of view can be divided into the central view and the peripheral view. The central view is a high resolution area to recognize a target object. The peripheral view is an area around the central view, and the resolution is lower than the central view. But it can sensitively respond to a movement of the target object and a lighting stimulus. When a driver recognizes the object gazed after moving eye, if we detect just the objects gazed by the driver then some objects which were not gazed by the driver will be left, and we can discriminate between the recognized and non-recognized objects. Miyaji has measured the cognition time for hazardous objects using a driving simulator based on this hypothesis [2].

But, a problem still remains. The useful field of view (or UFOV) is only the two-dimensional eye rotation (pitch angle, meaning up and down; and yaw angle, or left and right of eye ball) as shown in Fig. 1. But the space in which all objects are located has clearly a three-dimensional aspect. Mackworth defined the useful field of view as the visual area over which people can perceive information within a brief glance without eye or head movements [3]. Is it adequate that this location of the gaze point is considered in the two-dimensional field? We have some doubt about this. From the driver's point of view, the bike on the left side and the left turning vehicle on the right side are the same in a yaw angle of a driver's viewing direction as shown in Fig. 2. However, each three-dimensional attention area is significantly different in the direction of *depth*.

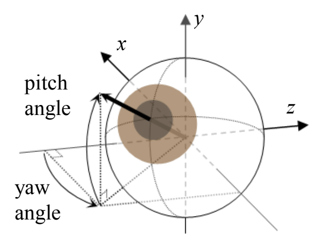

Figure 1. Viewing direction and eye ball rotation angle

In traffic environment, there are many objects such as vehicles, signs and signals at various depths. And the spatial relationship with them is three-dimensionally changing with runs of vehicles continuously. The driver has acquired much information required for driving by changing cautions frequently in three-dimensional space. Miura has described the importance of the depth attention characteristic [4], Today's most studies [5][6] notice only the two-dimensional spread at the measurement of the driver's useful field of view including the simulations done by Miyaji [2].

Earlier, we have proposed the camera measuring system [7] for estimating the driver's gaze point in three-dimensional space. This system has aimed at obtaining the tendency of the driver's gaze (recognition) or the effective cognitive range from a gaze point location three-dimensionally. We apply the word 'useful space of view' (USOV) to this effective cognitive range in the three-dimensions. To estimate the scale of USOV, it is necessary to take into consideration a spread of the cognition toward the *direction of depth*.

Here, in this work, we propose a method of measuring the time to recognize the object with different depth, and perform experiments to find any relationship between the difference of

Md. Shoaib Bhuiyan is a tenured associate professor at the Department of Medical Information Science, Suzuka University of Medical Science, JAPAN 510-0293 (corresponding author; e-mail: bhuiyan@suzuka-u.ac.jp).
Kenji Takagi, Haruki Kawanaka, and Koji Oguri are with Graduate School of Information Science and Technology, Aichi Prefectural University, Aichi 480-1198, JAPAN. (e-mail: oguri@ist.aichi-pu.ac.jp).

cognitive time for different gaze moving distances and/or different directions in depth. We think that it will become a great clue to measure USOV by combining this experimental result with the estimation method of driver's gaze point [7].

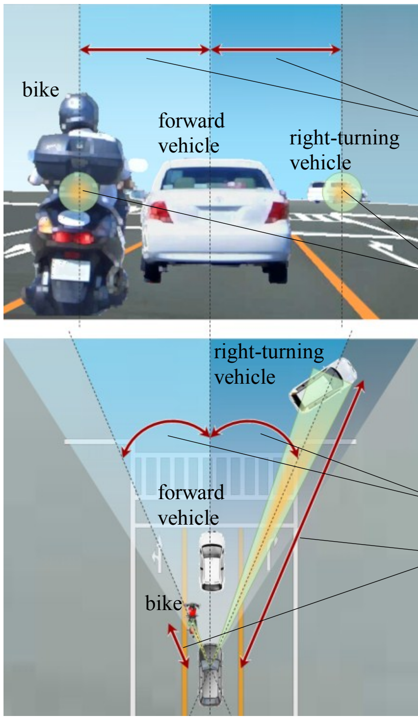

Figure 2. Gaze angle and depth of attention

II. Depth-Specific Experiment to Measure Cognition Time

A. Experimental Setup

A subject sits down at rest and looks at a fixation point. He notices a target displayed on the liquid crystal display (LCD) monitor projected to his peripheral vision, moves the gaze, recognizes it, judges the kind of the target and pushes an appropriate button (left or right) of a keyboard. The time elapsing from a target's appearance to a button push is measured. This 'time' consists of various partial time, such as time to respond to the stimulus within a peripheral vision, time to move the eye direction toward a target, time to focus on the target, time to identify the kind of target, time to choose an appropriate button, and time to push the button, and so on.

We want to know whether cognition time differs at different depth or not. So, we conducted three kinds of measurement experiments. They are "Reflective reaction experiment" where a subject reflects in a stimulus of a display and pushes a button, "Discriminative reaction experiment" where a subject identifies two kinds of targets and operates a button, and "Gaze depth movement experiment" where a subject moves a location of gaze point in the direction of depth, identifies a target and pushes an appropriate button. The result of these three experiments can clarify the relation between cognitive time and the distance and direction of gaze movement in direction of depth.

1) *Reflective Reaction Experiment*

In Reflective reaction experiment, a target (white circle) as shown in Fig. 3 is displayed to a LCD monitor, which is being cleared by black, after a certain period of time. And, the time elapsing from a target's appearance to a button push is measured. Subjects are directed to react and push a button of a keyboard immediately after a target appears on the screen. The number of appearances of a target was set at 20 times and let the average of the measured time be T_1. The size of the target displayed on a LCD monitor is adjusted according to the distance from a subject to be visible as similar in size.

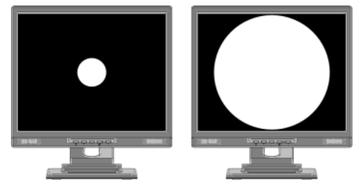

Figure 3. Target of Reflective Reaction Experiment

2) *Discriminative Reaction Experiment*

In Discriminative reaction experiment, a target ('3' or 'ε') as shown in Fig. 4 is displayed to a LCD monitor, which is also being cleared by black, after a certain period of time. Again the time elapsing from a target's appearance to a button push is measured. There are two buttons and a test subject pushes the appropriate button according to the kind of target. The button operation is performed, then the target disappears regardless of the correction and the LCD screen returns to black again. The number of appearances of a target is again 20 times and let the average of the measured time be T_2. The size of the target displayed on a LCD monitor is also adjusted similar to the Reflective reaction experiment described above.

Figure 4. Target of Discriminative Reaction Experiment

3) *Gaze Depth Movement Experiment*

In this experiment, as shown in Fig. 5, two LCD monitors for a target display are installed in front of a subject, and a fixation point as shown in Fig. 6 is located in between the monitors. The position of each display and the fixation point is fixed so that they aren't overlapped when a subject watches them. The rotation angle of eye moving from the fixation point to each display is the same and the two monitors are placed symmetrically. Let distance to each LCD monitor from the fixation point be x [m] equally. The distance to a subject and a near LCD monitor shall be 2 [m].

The targets displayed on LCD monitors are the two same kinds (Fig. 4) as Discriminative reaction experiment. While a test subject gazes on the fixation point, either of the targets appears on one of the monitors at random. The subject moves his viewing direction toward the stimulus, identifies the target and pushes the appropriate button. If a button is pushed, then the target disappears regardless of correctness of a button. So, the subject's viewpoint returns to the fixation point again.

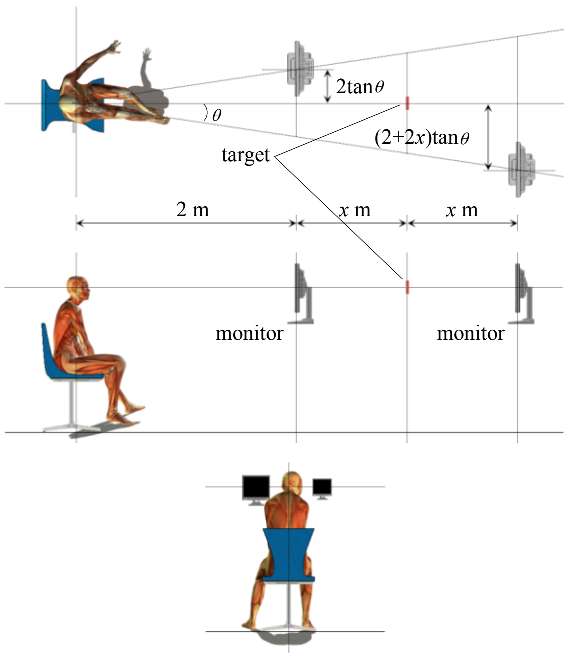

Figure 5. Environment of Gaze depth movement experiment

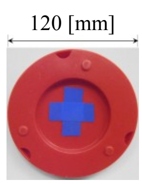

Figure 6. Fixation point

The number of appearances of a target was set at 20 times. We call the time to the far monitor from fixation point as 'Toward-far' and the time to the near monitor as 'Toward-near', and let each average of the measured time be T_3.

B. *Measurement Method of Cognitive Time*

The measured time T_1 in Reflective reaction experiment combines the time for reacting to a stimulus within peripheral view and pushing a button. The measured time T_2 in Discriminative reaction experiment unites the time for reacting to a stimulus within peripheral view, identifying the kind of target, choosing a correct button and pushing that button. The measured time T_3 in Gaze depth movement experiment combines the time for reacting to the stimulus within a view, moving the eye direction toward the stimulus, focusing on the target, identifying the kind of target, choosing a correct button and pushing that button. Therefore, the cognition time T_c for recognizing the kind of target and choosing the correct button is computed by the following formula (1).

$$T_c = T_2 - T_1 \tag{1}$$

And, the cognition time T_d for moving eye direction toward the stimulus and focusing on the target is computed by the following formula (2).

$$T_d = T_3 - T_2 \tag{2}$$

The time for recognizing the object that appeared at various depth can be analyzed by gathering cognition time T_c in case of different distance x. The cognition time influenced by the gaze moving distance and direction in direction of depth can be verified by gathering T_d in case of different distance x.

III. Experimental Results

All of the Reflective reaction experiment, Discriminative reaction experiment and Gaze depth movement experiments were performed with two set condition. The distance x between a fixation point and each display was set to 1 [m] in case A, and to 3 [m] in case B Therefore, distance between a subject and the two displays were 2 [m] and 4 [m] respectively in case A, but were 2 [m] and 8 [m] respectively in case B.

The average age of our seven male subjects was 22.7 years. Each of their reaction time T_1, T_2 and T_3 were actually measured for each case, and are shown in Table I and II. The cognition time T_c and T_d, which were theoretically calculated, are shown in Table III and IV.

TABLE I. Measured Time in Case A

	Near (2 [m])			Far (4 [m])		
	T_1	T_2	T_3	T_1	T_2	T_3
SubA	267.1	386.5	553.2	265.8	433.6	568.1
SubB	343.6	548.7	678.8	366.5	579.7	680.8
SubC	349.2	450.9	715.0	357.1	503.3	694.1
SubD	300.9	471.4	603.1	289.9	457.5	657.8
SubE	310.9	416.3	671.2	307.0	439.5	749.1
SubF	323.5	501.6	749.3	330.6	512.5	732.0
SubG	298.6	426.5	691.4	303.0	414.1	758.0
Ave	313.4	457.4	666.0	317.1	477.2	691.4

[msec]

TABLE II. Measured Time in Case B

	Near (2 [m])			Far (8 [m])		
	T_1	T_2	T_3	T_1	T_2	T_3
SubA	250.9	369.1	599.2	263.4	373.2	660.0
SubB	364.0	521.2	630.4	350.7	528.9	664.8
SubC	348.5	447.8	682.1	321.1	521.8	745.3
SubD	281.1	460.3	705.5	298.4	488.8	848.5
SubE	322.6	415.7	682.9	316.4	412.4	704.8
SubF	299.2	455.0	642.3	311.7	453.3	666.6
SubG	284.4	397.6	836.1	296.8	418.2	854.6
Ave	307.2	438.1	682.6	308.4	456.7	734.9

[msec]

TABLE III. CALCULATED TIME IN CASE A

	Toward-near (2 [m])		Toward-far (4 [m])	
	T_c	T_d	T_c	T_d
SubA	119.4	**166.7**	167.8	134.5
SubB	205.1	**130.1**	213.2	101.1
SubC	101.7	**264.1**	146.2	190.8
SubD	170.5	131.7	167.6	**200.3**
SubE	105.4	254.9	132.5	**309.6**
SubF	178.1	**247.7**	181.9	219.5
SubG	127.9	264.9	111.1	**343.9**
Ave	144.0	208.6	160.0	**214.2**

[msec]

TABLE IV. CALCULATED TIME IN CASE B

	Toward-near (2 [m])		Toward-far (8 [m])	
	T_c	T_d	T_c	T_d
SubA	118.2	230.1	109.8	**286.8**
SubB	157.2	109.2	178.2	**135.9**
SubC	99.3	**234.3**	200.7	223.5
SubD	179.2	245.2	190.4	**359.7**
SubE	93.1	267.2	96.0	**292.4**
SubF	155.8	187.3	141.6	**213.3**
SubG	113.2	**438.5**	121.4	436.4
Ave	130.9	244.5	148.3	**278.3**

[msec]

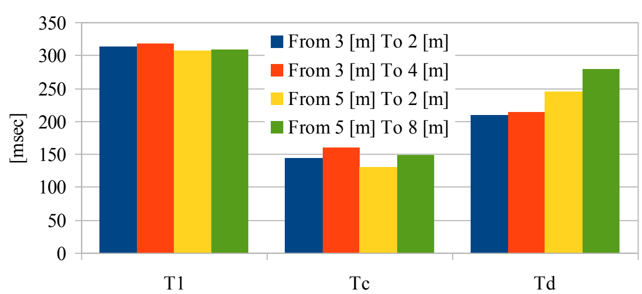

Figure 7. Averaged reaction time and cognition time

As shown in Table I and Table II, there were some subject that has reaction time T_1 about 250 [msec], and some subject that has about 350 [msec] in Reflective reaction experiment. Thus, there is no significant difference between a near monitor (2 [m]) and a far monitor (4 [m], 8 [m]) although individual difference exists in reaction time T_1. From these results, it can be observed that there is no difference in the reaction time T_1, T_2 regardless of the distance to a gazed monitor from a subject.

As cognition time T_c of Table III and Table IV showed, there were not significant changes of reaction time by the position of monitor although individual difference is seen also in Discriminative reaction experiment as well as Reflective reaction experiment. If a subject comes to the point where s/he recognizes that a target has appeared, the reaction time does not change with the appearance of depth of a target.

On the other hand, as cognition time T_d of Table III and Table IV showed, the cognition time of Toward-far becomes longer than *toward-near* among almost all the test subjects in Gaze depth movement experiment. Here, the average of time T_1, T_c and T_d is plotted to Fig. 7 using all the measurement results. From Fig. 7, it is found that the cognition time delay increases with distance of a target to a monitor. And, this rate of increase actually accelerates in *toward-far* of case B.

IV. DISCUSSION

There are various expressive forms of a two-dimensional useful field of view such as a range (distance), cognition time and accuracy rate. We introduced a three-dimensional useful space of view (USOV) on the basis of cognition time distribution. The USOV shall be determined in accordance with the range which can be moved and recognized within a certain definite period of time from the present point of gaze.

According to the result of the experiment, there was a tendency for cognition time of Toward-far to delay, but it will be necessary to measure under more varied conditions and depths when building the model of USOV. For example, how does it change in long distances such as x= 10 [m] or x= 20 [m]? How does human visual capability (visual acuity, eye movement velocity, accommodation ability, visual resolution, and so on) influences cognition time? Additional experiments are needed to build the USOV model since many unexplored items still remain.

Future work is to investigate widely about a cognitive time change for the distance of a target from a fixation point in order to define the USOV model, and to obtain the cognitive time distribution in consideration of not the depth but the three-dimensional translation vector.

REFERENCES

[1] M. Danno, *et al.*, “Measurement of Driver’s Visual Attention Capabilities Using Real-Time UFOV Method,” *International Journal intelligent transportation systems research*, 2011, vol. 9, no. 3, pp. 115–127.

[2] M. Miyaji, *et al.* “Detection of Hasty State by Means of Using Psychosomatic Information,” *International Conference Pattern Recognition Applications & Methods*, 2012, pp.33–40.

[3] N. H. Mackworth, “Stimulus density limits the useful field of view,” in R. A. Monty & J. W. Senders(Ed.), *Eye movements and psychological processes*, Lawrence Erlbaum, 1976, pp. 307–321.

[4] T. Kimura, T. Miura, K. Shinohara, “Effect of fixation point distances on allocation of attention in real three-dimensional space,” *Perceptual and motor skills*, 2010, vol. 109, pp. 327–337.

[5] K. K. Ball, *et al.* “Advances in technology used to assess and retrain older drivers,” *Gerontechnology*, 2002. vol. 1, no. 4, pp. 251-261.

[6] K. Ball and C. Owsley, “Identifying correlates of accident involvement for the older driver”, *Human Factors*, 1991, vol.33, no. 5, pp. 583-95.

[7] K. Takagi, *et al.*, “Estimation of a Three-Dimensional Gaze Point and the Gaze Target from the Road Images,” *International IEEE Annual Conference Intelligent Transportation Systems*, 2011, pp.526–531.