Proceedings of the 2007 IEEE Symposium on
Foundations of Computational Intelligence (FOCI 2007)

# Model of driver eye motion based on driving plan and prediction of changes in the environment

Yuki Togashi
Graduate School
of Computer Science and Technology,
Hokkaido University
y_togashi@complex.eng.hokudai.ac.jp

Takashi Omori
Tamagawa University
Research Institute
omori@lab.tamagawa.ac.jp

Koichiro Yamauchi
Graduate School
of Computer Science and Technology,
Hokkaido University
yamauchi@complex.eng.hokudai.ac.jp

***Abstract*— Modeling of human cognitive process is one of important research field for computational intelligence. In this study, we try to construct a computational model of car driver that can explain real world driver behavior. Through the research, we try to establish foundation for a bridge between basic science and engineering requirement. In our previous work, we proposed a model of driver's eye movements that consists of bottom-up and top-down eye motion sub models.**

**This paper proposes an extended version of the previous eye motion model. In the current extended model, top-down eye motion is determined based on the driver's driving plan, whereas the bottom-up eye motion is determined based on predicted locations of moving objects.**

**Keywords:** eye motion model, driving, prediction of intention, driving plan, map, fixation point, peripheral vision

## I. Introduction

Car safety systems are improving as intelligent transport systems (ITS) evolve. One way of making driving safer is supporting drivers by reminding them of objects that may be hazards. However, if a safety system announces that there are hazardous objects around, and the driver has already noticed them, the system will only annoy the driver. To avoid this problem, the safety system should alert the driver only to the objects that the driver has not noticed. To achieve this, the system has to recognize the cognitive status of the driver precisely. However, it is hard to achieve such precise recognition only by observing the driver's behaviors. To overcome this problem, the support system needs to predict the cognitive status of the driver using a model of human behavior. Several researchers have already proposed models of human driving behavior [1] [2] [3]. These models, however, describe only driver behaviors but not the driver's internal cognitive process. In this study, we focused on eye motion during driving and constructed a cognitive model for the motion of human eye. We also tried to re-create the driver's behavior from the model.

We have already constructed an eye motion model that consists of two sub models of eye motion: the eye motion process to active collection of information for steering operation, and passive eye motion in response to normal visual stimuli. This model yields a probability distribution function of eye motion at any given time. This model enabled us not only to estimate what kind of driver information processing is going on, but also to detect driver's status at any given time by calculating likelihood of actual eye motion data at a given time [4][5]. However, identifying actual state of the cognitive process was difficult. To overcome this issue, we introduced an effect of intention (driving plan) into the model of internal process for eye motion, and we calculated a probability distribution of eye motion based on driving plan. That is, as a glance movement model, the probability distribution of glance movement from a more general driving plan, as indicated by steering wheel operation for example, is added to the previous model. In our experiment, we examined a driver's glance, and evaluated validity of the model by comparing the driver behavior to that of the driving simulator [6]. We found that the model describes a driver's driving behavior well but some behavior does not agree with the model. Therefore, we added an additional sub model that predicts the environment by identifying hazardous objects, and included them into the bottom-up eye motion model.

## II. Driver Eye Motion Model

### A. Overview

As mentioned above, in our previous work we have already constructed a model of drivers' eye motion to detect the essential factor of eye motion [4][5]. The previous model predicts the driver's eye motion and steering operation with reference to bottom-up and top-down eye motion related to road and objects near the car. Therefore, we could estimate internal factors of actual eye motion and check whether the driver observes hazardous objects or not by comparing the model's behavior and an actual driver's behavior. For example, when formulating an eye motion model, we can estimate a factor of current eye motion by calculating the likelihood of the model's probability distribution function based on current eye motion data (Fig. 1).

The bottom-up eye motion model was constructed by the saliency map [7], which was based on findings about the human visual system. According the model, eye motion depends on saliency, which can be calculated by local texture and motion of objects. The probability distribution of bottom-up eye motion : $P_{Eye_{BU}}(\boldsymbol{x}, t)$, where $\boldsymbol{x}$ is the position in view area and $t$ is time, is calculated using the saliency.

The top-down eye motion model was constructed by a driving model based on reinforcement learning method proposed by Koike et al. [1]. Koike's driving model was based on a finding that drivers paid attention both to a area near and to a area far from the car [8][9]. Therefore, the

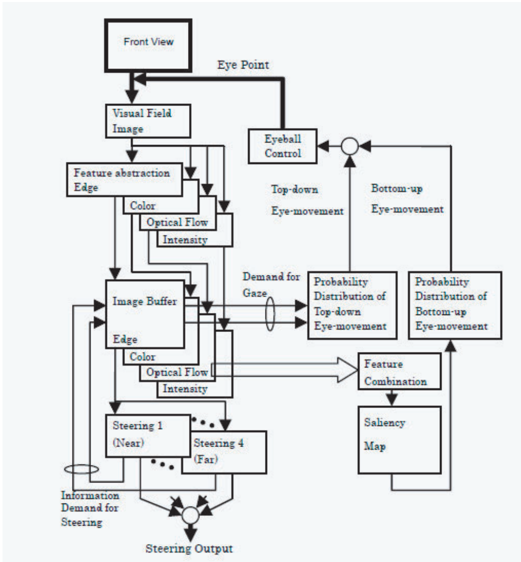



Fig. 1. Model of eye motion and steering operation.

model determined its steering operation based on current and future (two or three seconds later) car position. The model consisted of several modules, each of which output appropriate steering angles in each particular situation. The value function was calculated by the output value of each module, and the reward was given as a distance between the car and a edge of the road. So, the model learned to output appropriate steering angles so as to smooth operation of the car along a curve of the road. A steering angle was calculated by a weighted sum of module outputs. The weights were determined by a soft-max function, in turn calculated by a responsibility value of each module. The top-down eye motion probability distribution: $P_{Eye_{TD}}(\boldsymbol{x}, t)$ was detected based on the weights. The final eye motion probability distribution function was calculated by a sum of the bottom-up and top-down eye motion probability distribution functions.

$$P_{Eye}(\boldsymbol{x}, t) = rP_{Eye_{BU}}(\boldsymbol{x}, t) + (1 - r)P_{Eye_{TD}}(\boldsymbol{x}, t) \quad (1)$$

Here, we used 0.5 for $r$ value in our simulation. But it is reported that eye motion while driving mainly depends on objects outside and the value should be reconsidered to satisfy evidence.

We simulated this model in a simulation environment and evaluated its performance by calculating similarity to eye motion data for an actual driver. The results showed appropriate similarity to each situation[5]. This suggests that the former model may be used to predict a driver's status based on the eye motion probability distribution function.

### B. Eye motion model based on intention

The former model, however, worked well only in situations where there were no pedestrians or other cars. Moreover, the model was tested only on simple shaped roads. Generally speaking, it was hard to know what a driver really recognizes in the real environment. We constructed a model of the recognition process that reflected intention of the driver. The eye motion was also affected by the intention of the driver. A rough sketch of the processes is divided into the three steps below(Fig. 2).

(1) Recognition of general representation
Observing stimuli in front of the car and recognizing discrete objects. The recognition results are stored in the "Object recognition results buffer".

(2) Selection of a driving plan from several plan candidates
Based on the contents of the object recognition results buffer, an appropriate plan is selected from several candidates.

(3) Eye motion depending on the plan
To execute the plan, the driver looks around more carefully.

In the model, the system repeats (1) until a driving plan is selected in (2). When a driving plan is selected, (2) and (3) are executed repeatedly, but if the selected plan in (2), has been executed, the system repeats (1).

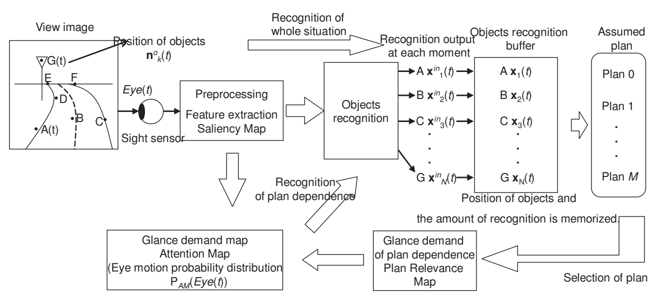



Fig. 2. Object recognition model.

### C. Saliency map

The eye motion process was simulated using the model that incorporates the saliency map [7][10]. We denote $SM(t, \boldsymbol{x})$ as the saliency map at each time $t$, where $\boldsymbol{x}$ denotes the position vector of view area $\boldsymbol{x} = (x, y)$. Then, $SM(t, \boldsymbol{x})$ is normalized to acquire the eye motion probability by being normalized by the summarized value as in the equation below.

$$SM(t, \boldsymbol{x}) = \frac{SM(t, \boldsymbol{x})}{\sum_{\boldsymbol{x} \in all} SM(t, \boldsymbol{x})} \quad (2)$$

Eye is supposed to move maximum point $\boldsymbol{x}^*$ where $\boldsymbol{x}^* = \text{argmax}_{\boldsymbol{x}} SM(t, \boldsymbol{x})$.

### D. Driving plan

During driving, plans compete their execution by calculating their score values based on an object's attributes and distance between the car and objects. This information is stored in the recognition results buffer, and, the most appropriate driving plan with the highest score is selected. Each plan consists of a conditions for selection, action, and observation to monitor the evolving situation. Driving actions are determined by the selected driving plan, which is described by steering angle, accelerator operation, and observation for a moment of plan selection.

### E. Selection of driving plan

The driving plan can be divided into master plans and sub plans. A master plan is a global plan such as "go straight, turn left, turn right" and so on. The sub plan denote condition for executing the corresponding master plan. The master plan is selected if all conditions are satisfied. If one of master plans is selected, the corresponding sub plans are activated. Then, the activated sub plans are evaluated by calculating estimation value $E$, and a sub plan with the maximum value is selected for execution.

### F. Plan relevance map

To realize actions in the selected sub plan, a plan relevance map (PRM) is introduced. The PRM denotes the position $\boldsymbol{c}_i = (c_{ix}, c_{iy})$ and width $\boldsymbol{s}_i = (s_{ix}, s_{iy})$ of target objects using normal distributions (Fig. 3).

(a) (b)

Fig. 3. Example of plan relevance map (b) when parking vehicle (a)

We denote $PRM(t, \boldsymbol{x})$ as PRM at each time $t$, where $\boldsymbol{x}$ denotes the position vector of view area $\boldsymbol{x} = (x, y)$. A plan that requires several objects includes multiple position and width value for target objects. Therefore, there are several partial $PRM(t, \boldsymbol{x})'s$ in a view area and a total $PRM(t, \boldsymbol{x})$ as weight sum is calculated as below.

$$PRM(t, \boldsymbol{x}) = \sum_i w_i R_i \left(\boldsymbol{x}, \boldsymbol{c}_i(t), \boldsymbol{s}_i(t)\right) \tag{3}$$

$$R_i(\boldsymbol{x}, \boldsymbol{c}_i, \boldsymbol{s}_i) = \frac{1}{\sqrt{2\pi s_{ix}^2 s_{iy}^2}} \exp\left\{-\frac{(x - c_{ix})^2}{s_{ix}^2} - \frac{(y - c_{iy})^2}{s_{iy}^2}\right\} \tag{4}$$

where $w_i$ denotes weighting variables to a object $i$ based on a recognition strength to the objects.

$PRM(t, \boldsymbol{x})$ is also converted to a probability distribution by being normalized by the total sum of $PRM(t, \boldsymbol{x})$ of all targets within the view. $PRM(t, \boldsymbol{x})$ is calculated as below.

$$PRM(t, \boldsymbol{x}) = \frac{PRM(t, \boldsymbol{x})}{\sum_{\boldsymbol{x} \in all} PRM(t, \boldsymbol{x})} \tag{5}$$

(6)

### G. Plan dependent eye motion probability distribution

According to $PRM(t)$ and the saliency map, attention map $AM(t)$ is calculated as in the equation below.

$$AM(t, \boldsymbol{x}) = SM(t, \boldsymbol{x}) \otimes PRM(t, \boldsymbol{x}) \tag{7}$$

where $\otimes$ is the operator for multiplying the values at the same position. The eye moves to the position where $AM(t)$ is highest.

When the car goes down a straight road, it is well known that drivers usually see the vanishing point of the road [8]. To describe this phenomenon, our model creates an additional PRM corresponding the vanishing point when no sub plan is selected.

## III. Driving Simulation

### A. Outline of driving simulation

The model was tested in an experiment by fabricating and using a driving simulator. This described 3D space with VRML and used Virtual Reality Toolbox of MATLAB for drawing and control. The vehicle model in the simulator used a fixed velocity running model of two wheels as a four-wheeled vehicle was used.

In the simulation environment, one car is stopped in one lane on a 10-m-wide two-lane road. Two oncoming cars go straight at a speed of 10[m/s] in the opposite lane. A virtual urban area was created complete with a house, a shop, a building, and trees by the side of the street. The task is to safely pass the stopped vehicle.

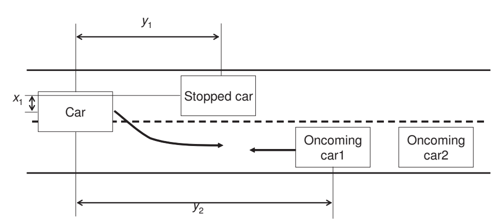



Fig. 4. Simulation environment

In this experiment, we considered the following five sub plans.

- Plan 0: Going straight.
  Driver goes straight in the lane of the appropriate direction.

- Plan 1: Stopping car.
  Driver observes a distance to oncoming car and stops self until the car passes through to a safe distance.
- Plan 2: Turning into the oncoming lane.
  Driver observes the distance with the oncoming car. And driver goes into the opposite lane turning the steering wheel to the right at a very safe distance.
- Plan 3: Going straight in the oncoming lane
  Driver observes the distance to stopped car and recovers the steering wheel, going straight in the oncoming lane at a safe distance from the stopped car.
- Plan 4: Going into the appropriate lane.
  Driver observes presence of person or object behind the stopped car, and goes into the appropriate lane, turning the steering wheel to left after passing the stopped car.

We defined the value calculation method, amount of steering wheel operation, and parameters of PRM $\boldsymbol{c}, \boldsymbol{s}$, and then simulated the execution of the plan.

We defined distance to the stopped car $x$ axially as $x_1$ and $y$ axially as $y_1$, and distance to the oncoming car $x$ axially as $x_2$ and $y$ axially as $y_2$. Value is calculated with the following equation.

- Plan 0 $E_0(t) = 0$
- Plan 1 $E_1(t) = -y_1 + c_1$
- Plan 2 $E_2(t) = y_2 + c_2$
- Plan 3 $E_3(t) = x_1 + c_3$
- Plam 4 $E_4(t) = y_1 + c_4$

$c_1, \ldots, c_4$ are safe distances between the car and the object.

- Safe distance to the stopped car ($y$) $c_1 = 15$
- Safe distance to the oncoming car ($y$ $c_2 = 40$
- Safe distance to the stopped car ($x$) $c_3 = 3$
- Safe distance from the stopped car after passing it ($y$) $c_4 = 2$

In this experiment, no surprise occurrences, such as a pedestrian suddenly appearing, were included in the simulation.

### B. Results

A change in value of each plan is shown in Fig.5. The value of all plans starts at 0 because the stopped car has not been observed yet, and Plan 0 is selected at the start of the simulation. Plans 1 and 2 are evaluated when approaching the stopped car. At this time, because oncoming car 2 is approaching, the value of Plan 2 becomes 0. Plan 1 is selected, and the driver stops the car. The value of Plan 1 falls when the oncoming car 2 passes, Plan 2 is selected, and the driver goes into the opposite lane. Plans 3 and 4 are executed according to the timing of each and the driver returns to the appropriate lane afterwards. At this time, the value of all plans returns to 0 and Plan 0 is selected again.

## IV. Human Eye Motion as Measured by Driving Simulator

### A. Experimental task and participants

A driving simulator scene was projected on a .75m by 1.02m screen. An eye camera and a keyboard were put on

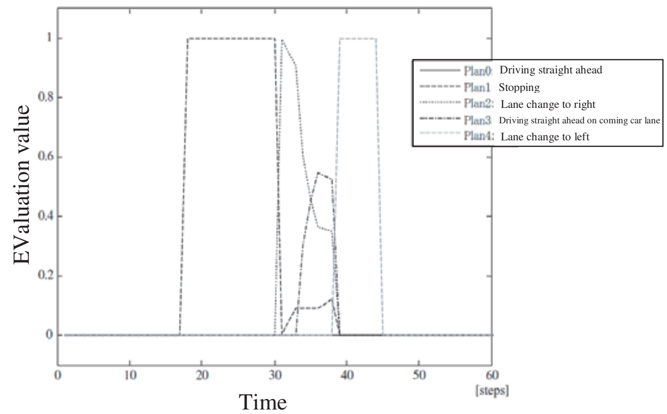



Fig. 5. Calculation of value

a desk, which is .825m from the screen, and participants sit in a chair, which was 1.3m from the screen.

Participants looked at the screen and simulates driving using a keyboard. Glance distribution of the participant was measured with non-contact eye camera EMR-AT VOXER (Nac Image Technology) and recorded every $1/60[s]$ .

A task is to drive a car while avoiding collision with the stopped and oncoming cars. In the simulated environment, the stopped car is in front of the driver, and there are four oncoming cars in the opposite lane. A simulated road is straight and 3-m-wide, and there are other objects like buildings and signs at side of the road.

The participant operates the car with two buttons.

- Button 1 Going forward and braking
- Button 2 Steering wheel operation

The participant turns a steering wheel to the right or left by pushing button 2. The speed of the participant's car is fixed at 50[km/h]. We prepared several oncoming cars of various speeds. Therefore, the two oncoming cars in near sides is a fixed 60[km/h] and one oncoming car in far sides is a fixed 45[km/h]. The participant repeats this task ten times. In each trial, the initial position of each car changes at random within the range of -20[m] $\sim$ 20[m] from the standard position. The task continues until the driver collides with another car, runs off the road, or successfully passes the stopped car and returns to the appropriate lane.

Four participants (all mail, age was 26.25 years in average) joined to the task, and their driving experience were 6.3 years in average. One of them drove several times a week, two of them drove several times a month, and one rarely drove.

### B. Results

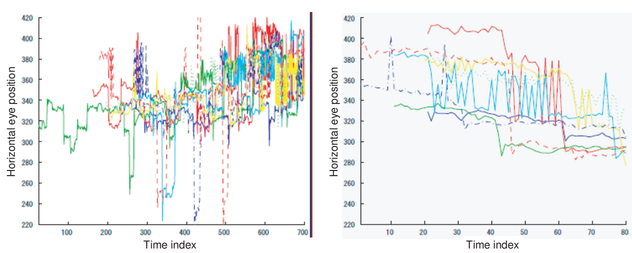
The direction of participant's glances as he pushes the button is shown in Fig. 6. These correspond to Plans 2 and 4, presented in Section III-A. Strictry speaking, Fig. 6 shows the all direction of the participant's glance during executing the sub-plan immediately before pushing button. Participants tended to glance to the right and glanced most frequently at the vicinity to the right of the lane (Fig. 6 (a)). Glances directed to the left, by contrast tended to be distributed in the vicinity of the left side of the center lane (Fig. 6 (b)).

(a) Plan2 :Change to right lane

(b) Plan4 :Lane change to left

Fig. 6. Measured eye motion

Glance measurement results for the actions corresponding to the plans described in Section III-A are shown in Fig. 7. The graph end position corresponds to a moment of the button push and the lines display eye motion immediately before the driving action was undertaken.




(a) Eye movement corresponding to Plan 1.

(b) Eye movement corresponding to Plan2 .

Fig. 7. Eye motion of one participant (Each line denotes the eye motion for one trial)

As Fig.7 (a) shows, when a participant glanced at vicinity of road center, his glance was frequently directed to the right of screen immediately before he took action. Fig.7 (b) shows that the glances move a lot from their initial position to left side immediately before the participant takes action.

In Fig.6, it can be seen that eye motion distribution varied with the driving situation during the task of passing stopped car. Moreover, from Fig.7, the glance that followed the driving plan corresponds to Plans 1 and 2, was observed. We also observed eye motion distribution that corresponds to Plans 3 and 4, but the results are omitted due to space limitations.

These results suggested that glance direction depends on the planned driving action of a participant. However, the individual variation among participants in the amount of eye motion was large, and was different between trials even for the same participant.

## V. Comparison of Model and Human

We compared the eye motion predicted by the model with the eye motions of a driver during the task mentioned above(III,IV). The directions of the model's galances and the directions of the participant's glances are shown in Fig. 8. These correspond to plan 1: Stopping car(III-A).

(a) The directions of model's galances

(b) The directions of participant's glances

Fig. 8. Glance distribution wahile plan 1

As a result, the distributions of eye motion were comparatively near immediately before beginning of passing the stopped car. However, human bottom-up glance movements to observe surroundings were less frequent than those predicted by the model. Moreover, the lack of eye motion suggests that people often observe the environment using peripheral vision.

## VI. Eye Motion Model Based on Prediction of Changes in the Environment

The fact that human bottom-up eye motion is less frequent than that predicted by the model is thought to be due to factors other than saliency. For instance, a driver directs his glance to objects that appear suddenly or get in his way. That is to say, at first, the driver recognizes what's going on in the environment, predicts where each relevant object will be in the next few moments, identifies objects that are potential hazards, and finally glances at them to observe them in more detail.

This means that people note the location of objects and track information on each object in the environment, which amounts to a space map. They then use the map to predict changes in objects' locations. People pay more attention to objects whose locations are not what they initially predicted and glance directly at them. This means the space map takes account of the current location of objects as well as their predicted location in the near future. This process offers drivers better choices of possible actions.

### A. Overview

The structure of the new model is shown in Fig. 9.

The bottom-up eye motion part of the previous model is included in the saliency glance movement calculated in the image processing system (Fig. 9 (a)). The top-down eye motion is included in the plan dependence glance movement calculated in the driving plan system (Fig. 9 (d)). A new model corresponds to the eye motion calculated in the object

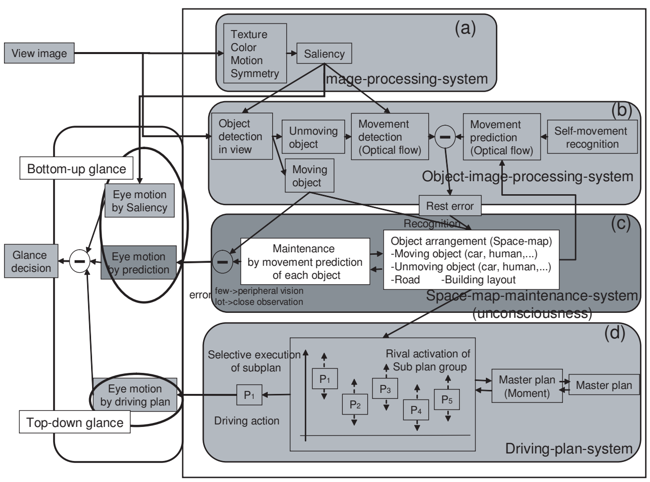



Fig. 9. Eye motion decision process

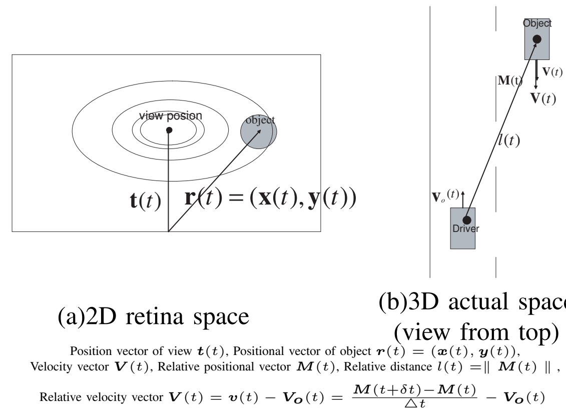



Fig. 10. Model parameter in retina and actual space.

image processing system (Fig. 9 (b)) and the space map maintenance system (Fig. 9 (c)).

That is, in the image processing system, moving objects and stationary objects are classified with optical flow. In the space map maintenance system, first, the stored recognition results in the object image processing system as a space map, next, the next situation of the moving object is predicted with recognition results, and finally, the driver glances at objects that do not conform to his expectations or at objects that suddenly appear in view. Information in the space map is also used in making driving decisions, and recognition processing from observation of the environment and of driving decision-making is executed simultaneously and concurrently. Decisions about glance direction at any given time are made based on competing demands between the bottom-up eye motion, as calculated by saliency or the predicted environment, and the top-down eye motion, as calculated by driving intention.

### B. Eye motion model based on prediction of object movement

To predict movement of objects, it is necessary to calculate the positional and rate vectors in three dimensions from the image of the object on the retina. We denote $\boldsymbol{M}(t)$ and $\boldsymbol{r}(t)$ as the 3D position vector and 2D retinal image vector at time $t$, respectively. We assume that $\boldsymbol{M}(t)$ is approximately derired from $\boldsymbol{r}(t)$ using the computational model of vision [11].

People also perceive objects in the environment with peripheral vision. To express this, the conviction level of recognition to each object is defined. The conviction level of object $M$ rises when the object is nearer the center of view and drops when it is farther from the center of view. Moreover, recognized information attenuates with time. $\alpha$, $\beta$ is constant.

$$\triangle C_M = -\alpha C_M + \beta \exp\left\{ -\frac{\| \boldsymbol{r}(t) - \boldsymbol{t}(t) \|^2}{2\sigma^2} \right\} \quad (8)$$

The position of the object in the next moment is predicted by $\tilde{\boldsymbol{M}}(t + \triangle t) = \boldsymbol{M}(t) + \triangle t \boldsymbol{V}(t)$. The accuracy of the predicted position is shown by the next expression, where $\gamma$ is constant.

$$Prob\,(\boldsymbol{M}(t+\triangle t)) = \frac{1}{\sqrt{2\pi}\sigma(C_M)} \exp\left\{ -\frac{\| \boldsymbol{M}(t+\triangle t) - \tilde{\boldsymbol{M}}(t+\triangle t) \|}{\sigma(C_M)^2} \right\} \quad (9)$$

$$\sigma(C_M) = \frac{\gamma}{C_M} \quad (10)$$

Because this is proportional to the inverse of the conviction level of decentralization, a driver can be confident of his prediction for an object with a high level of conviction and less confident of his prediction for a small object.

A driver's glance will probably be directed to object $M$ at this time, as described by the equation below.

$$Prob\,(\boldsymbol{M}(t+\triangle t)) < \theta \quad (11)$$

$$\theta = \frac{\delta}{l(t)} \quad (12)$$

$\delta$ is a constant. Once, the retinal position of the target object $\boldsymbol{r}(t + \triangle t)$ is equivalent to the center of the view $\boldsymbol{t}(t+\triangle t)$, $C_M$ becomes large, so that a driver will not likely glance at this object (see Eq (8)). On the other hand, if the target object is far from a driver's car, $l(t)$ becomes larger so that the driver is likely to glance at it.

## VII. Summary

In this paper, we discussed the processing necessary to construct a model of human eye movement during driving. In the model, a driver predicts the movement of objects in the environment and directs glances at objects that move differently than expected. The model is expected to predict observational eye movement that is similar to human behavior. It will be necessary to test the model by comparing it with eye movements during real driving situations. It will also be necessary to add the element of risk to the model.

## REFERENCES

[1] Y. Koike and K. Doya, "A driver based reinforcement learning with multiple-step sttae estimation," *IEICE Transactions(D-II)*, vol. J84-D-II, no. 2, pp. 370–379, 2001, (in japanese).

[2] D. D. Salvucci, "Integrated driver modeling in the act-r cognitive architecture," *Driver Modelling*, 2003.

[3] T.AI-Shihabi and R.R.Mourant, "A framework for modeling human-like driving behavior for autonomous vehicles in driving simulators," in *The Fifth International Conference on Autonomous Agents,Montrel,Canada*, 2001, pp. 286–291.

[4] K. Mizutani, G. Saito, T. Omori, and A. Ogawa, "A feasibility study of cognitive computation model for driver's process estimation from driving behavior," IEICE Technical Report NC2003-198, Tech. Rep., 2004, (in japanese).

[5] K. Mizutani and T. Omori, "A feasibility study of driver's cognitive process estimation from driving behavior," *IEEJ Transactions on Electronics, Information and Systems*, vol. 125-C, no. 6, pp. 967–975, 2005, (in japanese).

[6] T. Yuki, K. Mizutani, G. Saito, T. Omori, and K. Yamauchi, "Modeling of eye motion decision process considering car driver's driving plan," in *Info-Hokkaido 2006*, 2006, (in japanese).

[7] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[8] M.F.Land and D.N.Lee, "Where we look when we steer," *Nature,369*, pp. 742–744, 1994.

[9] M. Land and J. Horwood, "Which parts of the road guide steering?" *Nature, 377*, pp. 339–340, 1995.

[10] S.-J. Park, S.-W. Ban, J.-K. Shin, and M. Lee, "Implementation of visual attention system using bottom-up saliency map model," in *Artifical Neural Network and Neural Infomation Processing - ICANN/ICONIP*, 2003, pp. 678–685.

[11] M. Kawato and T. Inui, "Computational theory of the visual cortical areas," *IEICE Transactions(D-II)*, vol. J73-D-II, no. 8, pp. 1111–1121, 1990, (in japanese).