Development of a Situational Awareness Estimation Model Considering Traffic Environment for Unscheduled Takeover Situations

Hiroaki Hayashi[1] · Naoki Oka[1] · Mitsuhiro Kamezaki[2] · Shigeki Sugano[1]

Abstract

In semi-autonomous vehicles (SAE level 3) that requires drivers to takeover (TO) the control in critical situations, a system needs to judge if the driver have enough situational awareness (SA) for manual driving. We previously developed a SA estimation system that only used driver's glance data. For deeper understanding of driver's SA, the system needs to evaluate the relevancy between driver's glance and surrounding vehicle and obstacles. In this study, we thus developed a new SA estimation model considering driving-relevant objects and investigated the relationship between parameters. We performed TO experiments in a driving simulator to observe driver's behavior in different position of surrounding vehicles and TO performance such as the smoothness of steering control. We adopted support vector machine to classify obtained dataset into safe and dangerous TO, and the result showed 83% accuracy in leave-one-out cross validation. We found that unscheduled TO led to maneuver error and glance behavior differed from individuals.

Keywords Autonomous driving · Situational awareness · Cognitive behavior · Unscheduled takeover

1 Introduction

Automated driving (AD) systems have the potential to achieve safer car-society and to enhance the quality of life of people. However, there are still many issues for achieving a fully AD system. SAE International defines six levels of automation, from level 0 (no automation) to level 5 (fully vehicle autonomy) [1]. In level 3 systems, drivers are not required to monitor the road environment and are allowed to engage in non-driving related tasks (NDRTs), such as reading books. A

takeover request (TOR) will be issued when the AD system reached functional limits, and then the drivers are required to intervene the vehicle control (takeover: TO) and start manual driving (MD) [2].

TO would be dangerous especially in 'unscheduled situations', in which the AD vehicles encounters unplanned roadworks or sudden accidents that cannot be dealt with the AD system. Figure 1 shows an example of unscheduled TO, where an obstacle (a crashed car) suddenly appears. The time from when TOR is issued until the ego-vehicle collides with the obstacle is called the time to collision (TTC). After TOR is issued, the driver holds the steering wheel and places his/her foot on the pedals. This is called physical engagement. Then or simultaneously, the driver checks the surrounding traffic situation and determines actions to be taken. This is called cognitive engagement. AD vehicles have a human interface (typically, a button) to enable the driver to approve the TOR. After completing physical and cognitive engagement, the driver approves the TOR by pressing the button to change the driving mode from AD to MD [3].

However, drivers engaging in NDRTs are often distracted and may have low or zero situational awareness (SA). The related works have reported that drivers of a highly automated vehicles are likely to perform NDRTs and this leads to degrading drivers' SA [4] and NDRTs do not influence

✉ Hiroaki Hayashi
   h_hayashi@sugano.mech.waseda.ac.jp

   Naoki Oka
   n_oka@suganos.mech.waseda.ac.jp

   Mitsuhiro Kamezaki
   kame-mitsu@aoni.waseda.jp

   Shigeki Sugano
   sugano@waseda.jp

1   Department of Modern Mechanical Engineering, Waseda University, 17 Kikui-cho, Shinjuku-ku, Tokyo, Japan

2   Research Institute for Science and Engineering(RISE), Waseda University, 17 Kikui-cho, Shinjuku-ku, Tokyo, Japan

Fig. 1 Flow of unscheduled takeover. TOR stands for takeover request. Drivers are required to complete physical and cognitive engagement after TOR is required
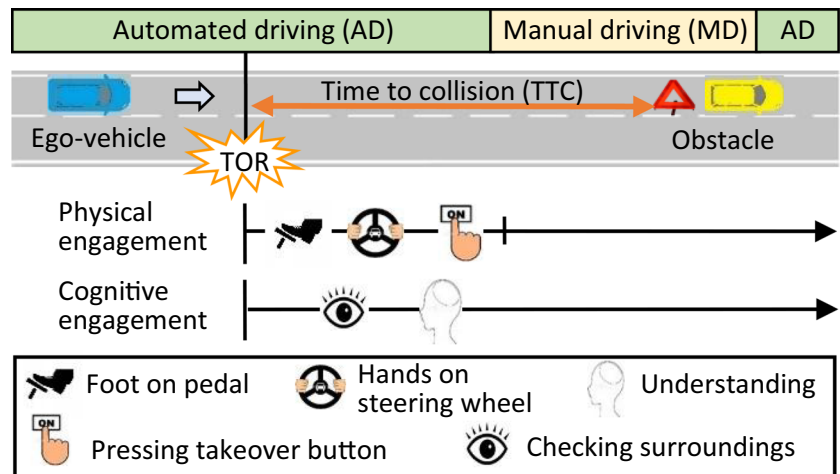
physical engagement (can be performed on reflex) but cognitive engagement [5]. These studies indicate that there is a risk that the drivers approve the TOR without sufficient cognitive engagement and this would lead to car accidents. It is thus necessary to develop a system to evaluate driver's cognitive engagement and judge if the driver is ready to start MD.

Psychophysiological approach using eye gaze was proposed to estimate human's cognitive engagement [6] and this had been applied to drivers during AD (SAE level 2: drivers must monitor the road) [7]. However, there are no effective methods for estimating driver's SA in unscheduled TO situation (SAE level 3). Thus, we previously proposed a SA estimation model based on driver's glance data [3]. We derived the standard glance model including the glance area (mirrors or windows where the driver looked) and glance time (how long driver looked at each area) to estimate driver's SA. We then developed a SA assistant system to highlight areas where was insufficiently looked. The experimental results revealed that the assistant system improved the driving performance and reduced the number of accidents during TO. The study indicates that the driver's glance behavior greatly affected driver's SA. However, it was a preliminary study of SA estimation, that is, the SA estimation system must be improved in terms of 'considering the surrounding situation' and 'treating glance features as combination'.

One is if a driver looks at glance areas to understand the environment (Fig. 2). For example, the driver surely looked a glance area, but if other vehicles to be checked were not appeared in the glance area at the time, the driver would not detect the vehicles. SA consists of the perception of elements in current situation, comprehension of current situation, and projection of future states [8]. Also, detecting driver inattention requires accurate real-time driver-gaze monitoring combined with a vehicle and road scene state [9]. Thus, driving-relevant objects in traffic environments are important to estimate driver's SA.

The other is how glance features relate to one another. For example, if the driver checked behind via the side mirrors, the driver could spend less time (or does not need) to check the rearview mirror. Moreover, if the driver checked the front in an enough time and then looked other areas in a few second, the driver might not have sufficient information of the front road. Thus, driver's glance features should be treated as a total glance behavior.

From the above, our hypothesis is that the SA estimation model can be improved by 'considering the surrounding situation' and 'treating glance features as combination'. In this study, we developed a new SA estimation model considering driving-relevant objects in traffic environments. To evaluate the relationship among glance features, we created an explanatory variable set which was made by combinations of glance

Fig. 2 Combination of glance and traffic environment. Drivers can be aware of traffic environment through glance areas
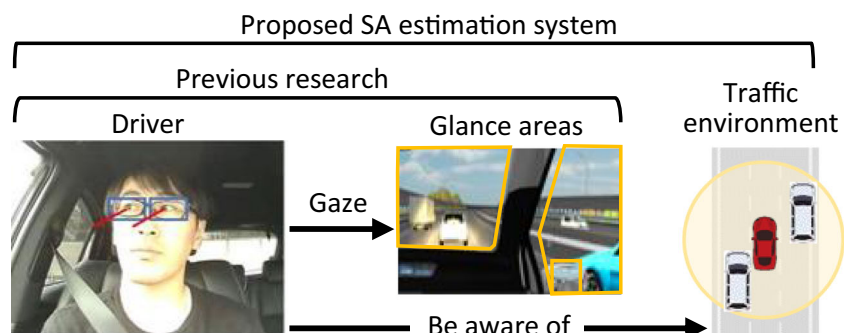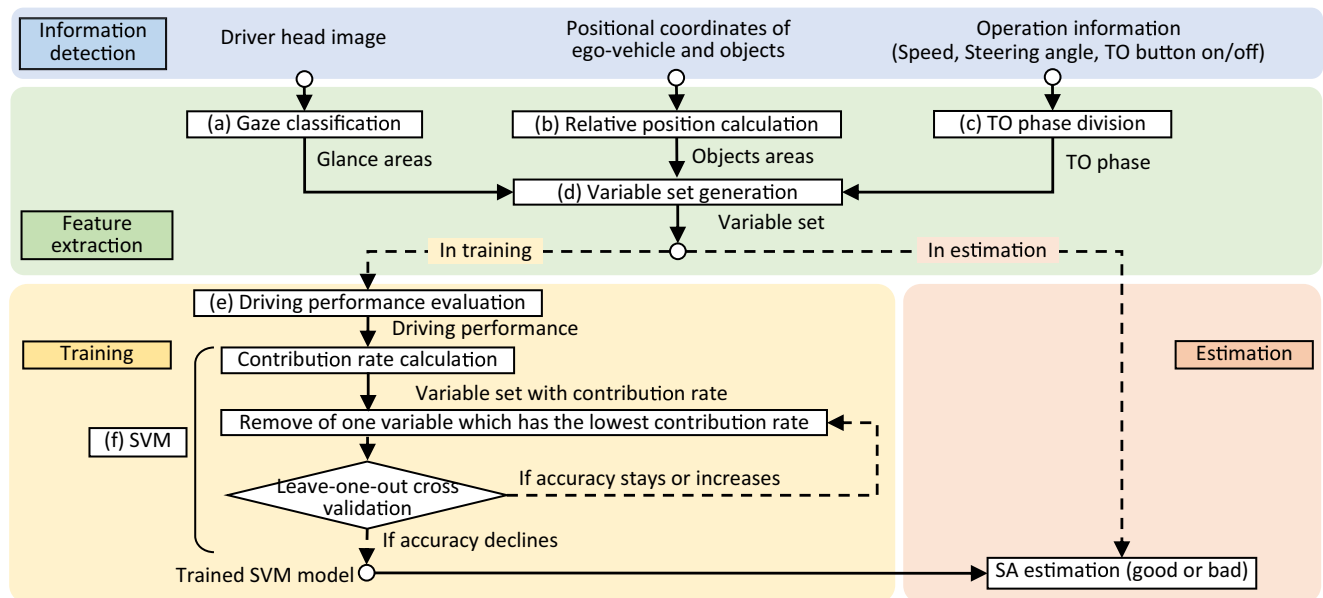
Fig. 3 Architecture of SA estimation model. **a** Driver's gaze direction is classified into eight glance areas. **b** Driving-relevant objects' position is classified into ten objects area. **c** Series of TO is classified into five phases. **d** Extracted features are combined to generate explanatory variable set. In training phase, **e** driving performance is evaluated for features and introduced support vector machine (SVM) for holistically analyzing these variables. This study contributes to reveal that driver's gaze on driving-relevant objects has positive effect on SA estimation and to derive determining variables which significantly influence the SA.

predictor variable using TO phases 1–5. **f** Perform variable selection and train the model based on support vector machine (SVM) using TO phases 1–2. After offline training, the model can estimate driver situational awareness (good or bad) in phases 1–2

2 Methodology

In this section, we explain the proposed SA estimation model including input and output information and classification method (Fig. 3).

Fig. 4 Definition of glance and object areas. Glance area has eight and object area has ten areas (glance area + LBS + RBS). **a** Definition of glance areas with eight areas. **b** Definition of object areas with ten areas
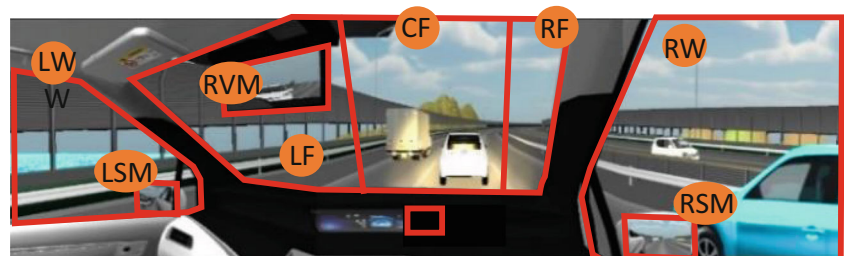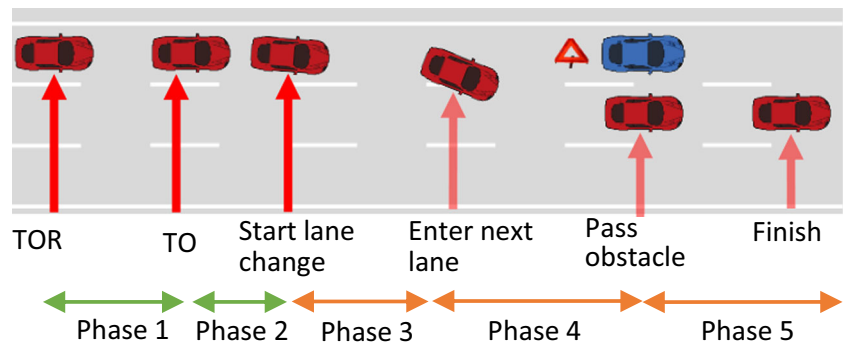

(a) Definition of glance areas with eight areas
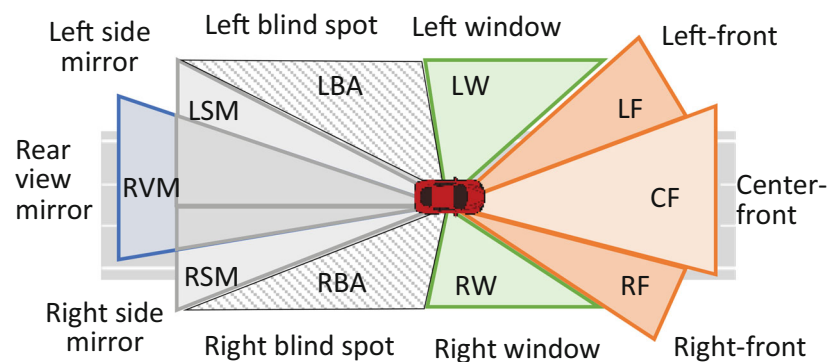
(b) Definition of object areas with ten areas

Fig. 5 Series of TO which has five phases. Phases 1–5 were used for evaluating driving performance, and phases 1–2 were used for SA estimation

2.1 Glance Area and Traffic Environment

As stated in Section 1, precise estimation of driver's SA needs to analyze driving-relevant objects in traffic environment. It is important to know if the driver looked the objects and its timing and duration time of the glance. To analyze them, we estimate the driver's gaze direction (Fig. 3a) and the position of the objects (Fig. 3b).

As shown in Fig. 4a, driver's gaze direction was classified into eight areas, including left-front (LF), center-front (CF), right-front (RF), right side mirror (RSM), rear view mirror (RVM), left side mirror (LSM), right window (RW), and left window (LF). These areas are known as the common areas where drivers should check [10, 11]. We divided the front area into three according to the traffic lane, including CF (for objects in the same lane as ego-vehicle), RF (for objects in the right lane), and LF (for object in the left lane). A gaze classification system is described in Section 3.1. Driving-relevant objects in traffic environment include obstacles, other vehicles, pedestrians, and traffic signals. In this paper, we target a TO scenario where a driver is required to change a lane (as described in Section 4), so we set other vehicles and obstacles as the driving-relevant objects. We calculated the relative position of these objects to ego-vehicle and classified them to ten objects areas based on the glance areas (Fig. 4b). As blind spots where the drivers cannot see, we added right blind spot (RBS) and left blind spot (LBS).

If driver's glance area is equal to an object area, we assume that the driver looked the object and understood the attribute of the vehicle such as the position and speed.

2.2 TO Phases for Lane Change

Throughout a TO for lane change, drivers behave differently in different phases. In this paper, we thus divided a series of TO into five phases (Fig. 3c). Definition of each phase is shown in Fig. 5.

- Phase 1: From TOR is issued until the TO button is pushed.

- Phase 2: From TO button is pushed until lane change is started (steering angle >5°).
- Phase 3: From lane change is started until ego-vehicle entered the next lane.

Table 1 Explanatory variables and contribution rate

#	Explanatory variables (time s)	Weight *3	CR % *3
1	[Phase 1] Gaze on F *1	−207.5	1.35
2	[Phase 1] Gaze on RVM *1	−199.1	1.30
3	[Phase 1] Gaze on RSM *1	624.8	4.07
4	[Phase 1] Gaze on RW *1	−190.0	1.24
5	[Phase 1] Gaze on RVM (FV exists) *2	−398.8	2.60
6	[Phase 1] Gaze on RSM (FV exists)	−9.3	0.06
7	[Phase 1] Gaze on RW (FV exists)	0	0.00
8	[Phase 1] Gaze on FV	−408.1	2.66
9	[Phase 1] Gaze on CF (obstacle exists) *1	14.9	0.10
10	[Phase 2] Gaze on F	899.7	5.86
11	[Phase 2] Gaze on RVM	61.0	0.40
12	[Phase 2] Gaze on RSM	468.2	3.05
13	[Phase 2] Gaze on RW	367.9	2.40
14	[Phase 2] Gaze on RVM (FV exists)	173.7	1.13
15	[Phase 2] Gaze on RSM (FV exists)	447.9	2.92
16	[Phase 2] Gaze on RW (FV exists)	−71.7	0.47
17	[Phase 2] Gaze on FV	549.9	3.58
18	[Phase 2] Gaze on CF (obstacle exists)	549.5	3.58
19	[Phase 1&2] Gaze on F	692.2	4.51
20	[Phase 1&2] Gaze on RVM	−138.1	0.90
21	[Phase 1&2] Gaze on RSM	1093.0	7.12
22	[Phase 1&2] Gaze on RW	177.9	1.16
23	[Phase 1&2] Gaze on RVM (FV exists)	−225.0	1.47
24	[Phase 1&2] Gaze on RSM (FV exists)	438.6	2.86
25	[Phase 1&2] Gaze on RW (FV exists)	−71.68	0.47
26	[Phase 1&2] Gaze on FV	141.85	0.92
27	[Phase 1&2] Gaze on CF (obstacle exists)	564.4	3.68
28	From TOR–TO	−1165.5	7.60
29	From TO–LC	720.5	4.70
30	From TOR to first look at F	−1841.2	12.0
31	From last look at F–LC	341.6	2.23
32	From last look at RSM–LC	−803.3	5.24
33	From last look at RW–LC	−284.0	1.85
34	From last look at RVM–LC	−1002.6	6.53

*1 F includes CF, LF and RF, each acronym is shown in Fig. 4

*2 FV: following vehicle, FV exists: driver looked at the glance area while following vehicle was there

*3 These variables are described in Sections 2.5 and 5.2

- Phase 4: From the ego-vehicle entered the next lane until the ego-vehicle passed the crashed car.
- Phase 5: From the ego-vehicle passed the crashed car until the ego-vehicle runs for a while (5–10 s).

To judge if the drivers could safely change the lane, SA should be estimated before the ego-vehicle enters into the next lane. Thus, we only used information within phases 1–2 for SA estimation. To understand if the driver safely changed the lane, information within phases 1–5 was used.

2.3 Explanatory Variables

There are generally two approaches to select explanatory variables for classification.

- A small number of explanatory variables is set based on a hypothesis for predictor variables. The purpose is to evaluate how much the hypothesis could explain the predictor variable [12, 13]. However, other potential key variables could be overlooked due to the limited explanatory variables.
- Many available explanatory variables are introduced, and necessary explanatory variables are extracted by variable selection methods [14, 15]. The extracted explanatory variables are called determining variables. The purpose is to increase estimation accuracy. However, it cannot always explain the relationship between the determining and predictor variables.

In this study, we adopted both of two approaches. We extracted several important features that would be related to SA and created an explanatory variable set that combines them (Fig. 3d). We then extracted determining variables by using a variable selection method. The explanatory variable set was the combinations of the driver's glance, position of following vehicle and obstacle, and phases of TO. As listed in Table 1, 34 explanatory variables in total were introduced.

- To know how drivers had glance behavior, the total time that drivers gaze on each glance area in each phase was introduced. Driving-relevant objects are not here considered, so for the simplicity, we combined LF, CF, and RF as the front (F).
- To know if drivers were aware of following vehicles and obstacles, the total time that the drivers gazed on these objects was introduced. To know if the driver checked the obstacle which can only be seen in CF, we do not combine LF, CF, and RF here.
- To know the time spent on each phase, and drivers' reaction time to TOR, the time from TOR until TO, from TO until lane change starts, from TOR until drivers' first look at front were introduced.
- Traffic environments vary with the time, so drivers should check the surroundings just before changing the lane. The time from the last look at each glance area until lane change starts was thus introduced.

Fig. 6 **a** SVM chooses support vectors and makes a separating hyperplane with the max margin. **b** C-SVM tolerates a few points to get misclassified. **c** RBF kernel performs non-linear classification
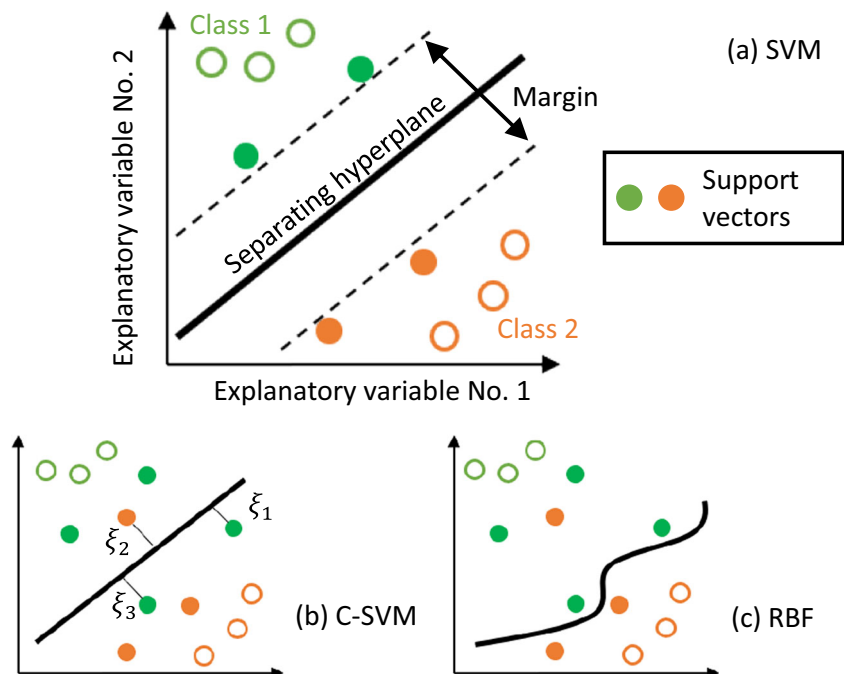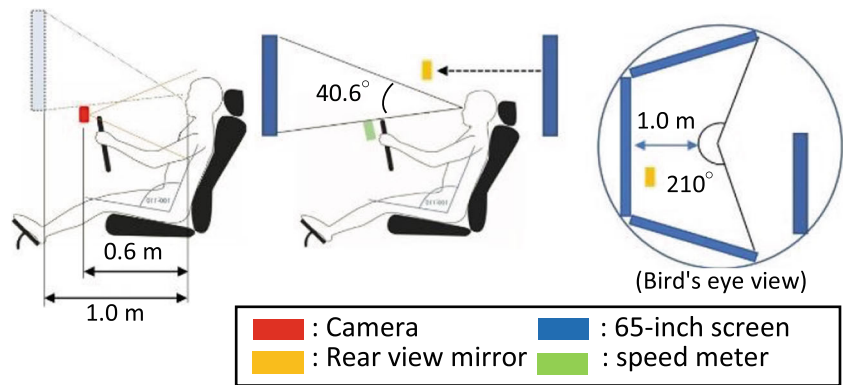
Fig. 7 Simulator arrangement. Four screens cover a field of view as wide as an actual vehicle

2.4 Predictor Variables

It is difficult to determine objectively if SA is good or bad. Driving performance can be the most important criteria. Drivers should drive smoothly and keep safe distances from other vehicles. Thus, we used driver performance instead of SA for the predictor variable (Fig. 3e). In this study, we defined good driving performance when following two conditions are satisfied.

- The driver kept more than 20 m away from other vehicles and obstacles. This is because the vehicles' speed in highway is limited as 22 m/s and 1 s is regarded as required driver's response time.
- The driver kept the derivative value of the steering angle $\dot{\theta}$ less than three times of baseline. We found from exploratory experiments that $\dot{\theta}$ related with SA and it was almost more than three times of baseline when they rushed into changing a lane. The baseline is the maximum value that the driver performed at a simple lane change TO without any other vehicles.

Driving performance was evaluated between phases 1–5. If a driver had good performance, his/her SA can be assumed as good, otherwise as bad.

2.5 Classification Using Support Vector Machine

Our SA estimation model requires a binary classification to judge if a driver has good or bad SA, and we need to holistically analyze the effect of each related variable, so we introduced support vector machine (SVM) (Fig. 3f) [16]. SVM is attractive for classification with high flexibility and estimation accuracy. Figure 6 shows how SVM works. SVM finds a separating hyperplane that maximizes the margin (distance between those closest points to the line) (Fig. 6a). The closest points are called support vectors, and is given by

$$SV_i = \left(a_i, x'_{i,1}, x'_{i,2}, x'_{i,3}, \ldots, x'_{i,n} \right) \quad , \tag{1}$$

where SV_i is the i-th support vector, a_i is the coefficient of SV_i, and $x_{i,j}$ is the coordinate of SV_i, i.e. the value of explanatory variables.

We adopted LIBSVM (ver. 3.14), which is a commonly used SVM library [17]. Among several types of SVM, we chose C-SVC since it was robust for uncertainty and ambiguity of data by misclassification errors. C-SVC works by solving an optimization problem denoted in (2).

$$\min_{w, \xi} \left[\frac{1}{2} \|w\|^2 + c\sum_{i=1}^{n} \xi_i \right] \quad , \tag{2}$$

Fig. 8 Drivers engage on NDRTs while AD working. AD system issues TOR by indicator and beep, and drivers push TO button to approve TOR
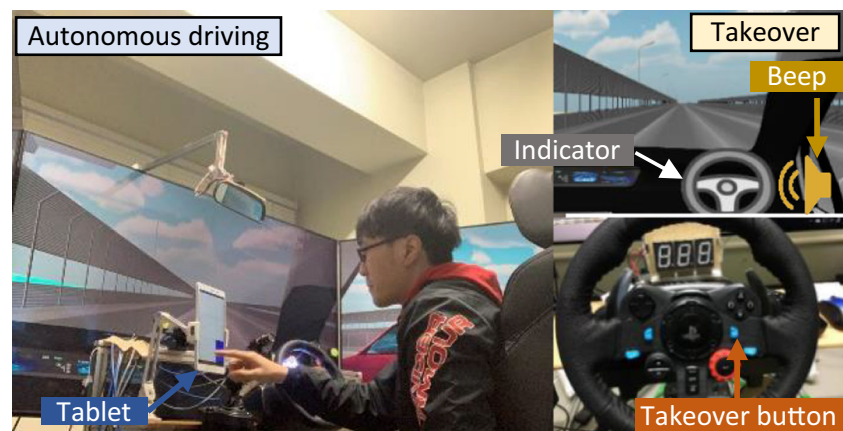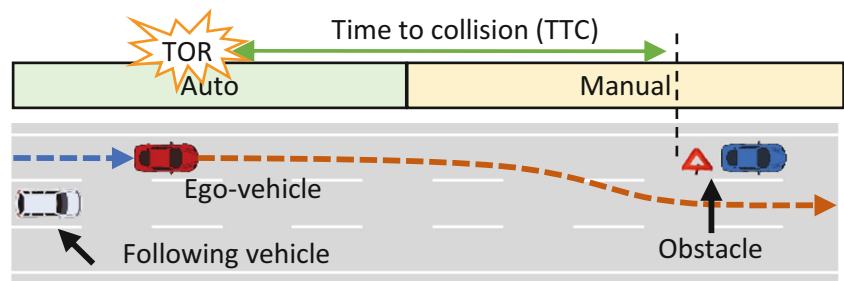
Fig. 9 Takeover scenarios. An obstacle prevents the autonomous system, so drivers must manually change lane

$$w_j = \sum_{i=1}^{n} a_i x'_{i,j}, \tag{3}$$

where w is the weight on each explanatory variable of the separating hyperplane and $\|w\|$ is the size of w. ξ_i is the distance of misclassified samples to their correct region. c is the coefficient and larger c gives solutions with less misclassification but may lead to overfitting (Fig. 6b).

To examine the relative impact of explanatory variables, we calculated the contribution rate (CR) of each explanatory variable using w, and it is given by

$$CR_j = \frac{|w_j|}{\sum_{j=1}^{34} w_j} \ . \tag{4}$$

LIBSVM contains several kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid. We adopted RBF kernel since it could deal with non-linear problems and systems with the complexity. In RBF, the separating hyperplane is defined by

$$y = \sum_{i=1}^{N} a_i \cdot \exp\left(-\gamma \|x'_i - x\|^2\right) \ , \tag{5}$$

where vector x is the test data, y is the estimation result of the test data (good or bad), $\|x'_i - x\|$ is the distance between the test data and each support vector. γ is a coefficient to present the complexity of hyperplane (Fig. 6c).

The two coefficients γ and c directly influence the performance of SVM. To achieve accurate classification, we used a grid search to find the optimal hyperparameters for each training dataset. We explored the determining variables and tested the final model by following steps.

1. Set all data as training data and find optimal parameter by grid search. Use the optimal parameter to create a primary SVM model.
2. Calculate the estimation accuracy with the primary SVM model by leave-one-out cross validation (LOOCV) [18], as primary accuracy.
3. Calculate w and CR of each variable by support vectors, which can be obtained from the SVM model.
4. Remove one explanatory variable with the lowest CR, and calculate the estimation accuracy by LOOCV. Then, remove one more variable with the lowest CR from the rest and redo LOOCV. This step is repeated until the accuracy starts to decline.
5. The combination of explanatory variables with the greatest accuracy is defined as the determining variables and the accuracy is the final accuracy.

3 Implementation

We set an unscheduled TO situation for developing the SA estimation model. This section describes a gaze classification system and a driving simulator (DS).

3.1 Gaze Classification

To obtain driver's glance data, we used Smart Eye Pro with three cameras to measure driver's gaze direction [19]. To ensure the accuracy of glance data, we set another RGB camera in front of the driver to record the driver's face, and we manually compared the recorded videos with the outputs of Smart
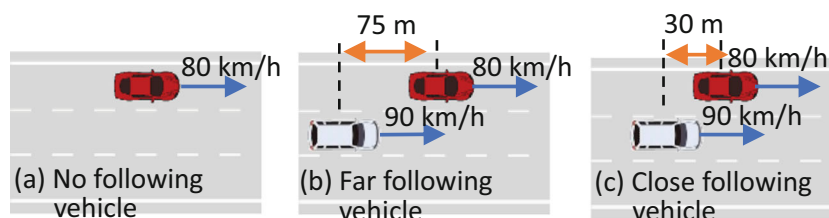

Fig. 10 Three patterns of following vehicle according to its existence and position. The distances and speeds are the ones at the time when TOR is issued. **a** No following vehicle. **b** Far following vehicle. **c** Close following vehicle

Table 2 Pattern of TO scenario

	No vehicle	Far	Close
TTC: 10 s	P1	P3	P5
TTC: 7 s	P2	P4	P6

Eye Pro. If the outputs were clearly wrong, we fixed them. As for implementing the system in real vehicles, we will introduce a gaze classification system using convolution neural networks [3].

3.2 Driving Simulator

The driving conditions must be the same among subjects, so we performed experiments in a DS based on Unity [20]. The dynamic behavior of vehicles can be autonomously controlled. This study focuses on driver's glance behavior, so the driver's eyes movement in the DS must be the same as possible as that in real vehicles. The simulator consists of four 65-in. LED screens, and three of them arranged at the front with a horizontal and vertical field of view of 240° and 40.6°, respectively, and the fourth one is placed the back to show the traffic environment behind the vehicle (Fig. 7). A rearview mirror is also set to enable the driver to check behind the vehicle like being in a real vehicle. Cameras for gaze classification are located behind the steering wheel.

A tablet computer is set in the left side of the steering wheel for NDRTs (Fig. 8). When the system issues TOR, the front screen displays 'manual driving' and makes a beep, and then driver pushes the TO button placed on the steering wheel. It might be better that the system shows more information about current traffic environment, but TTC is too short to convey the amount of information. We also need to address efficient ways to convey information to drivers when TOR. These issues will be dealt in future. In this paper, to simply examine driver's behavior, we only used beep which contains the minimum information.

4 Design of Experiments

This section describes the simulated TO scenarios, data acquisition, subjects, and experimental procedure.

4.1 Unscheduled TO Scenario

We set a sudden car accident situation in a highway as an unscheduled TO scenario in the DS (Fig. 9). A crashed vehicle stopped on the first lane as an obstacle, so the ego-vehicle that runs on the lane is required to move to the next lane. We set six different TO patterns by changing TTC and other vehicles. Considering that the TTC should be 7 s at least for safe TO, we set the TTC to 7 s (shorter) and 10 s (longer) [21]. Three patterns of other vehicles were set, and drivers were expected to perform different driving behavior under the different patterns (Fig. 10).

- No following vehicle. There is no other vehicle on the next lane. After a driver checks the traffic environment, the driver can perform lane change at any timing and with any running speed.
- Far following vehicle. A vehicle far behind in the next lane runs faster than the ego-vehicle. A driver is expected to percept the vehicle, comprehend that it is faster than the ego-vehicle, and speed up and move to the front of it. If the driver enters the next lane with a low speed, a collision will occur.
- Close following vehicle. A vehicle near behind in the next lane runs faster than the ego-vehicle. A driver is expected to percept the vehicle, comprehend that it will pass the ego-vehicle, and speed down and move to behind it. If the drive enters the next lane without speed down, a collision will occur.

During AD system working, drivers engaged in the NDRT until they received TOR. The time until TOR was issued was 3–5 min randomly. NDRT was a n-back cognitive task using the tablet.

## 4.2 Data Acquisition, Subjects, and Procedure

We recorded vehicle's telemetry including position, speed, steering angle, and pedal position from the DS at 100 Hz. Smart Eye Pro measured subjects glance area. Driver's head image was recorded in Full HD (1080p) at 16 Hz. Smart Eye Pro was not connected to the DS. To synchronize their time

Table 3 Performance results

Subject No.	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Rate %
P3	1	0	0	1	0	0	0	1	1	1	1	1	58
P4	0	1	1	1	0	1	1	1	1	1	0	1	75
P5	0	1	0	1	0	1	0	1	1	0	1	0	50
P6	1	0	1	1	0	1	0	1	0	1	0	0	50

1: Good Performance, 0: Bad Performance

Fig. 11 Examples of glance maps. The vertical axis represents glance area. The blue bars show the area where the driver looked. The pink and yellow bars show where the following vehicle and the obstacle was, respectively. When the blue bar overlaps the pink or yellow bar, we assume that the driver understood the following vehicle or the obstacle. The horizontal axis represents time and the sequence of glance is divided into phases by vertical brown lines. **a** Good performance (S1 in P5). **b** Bad performance (S2 in P6)

(a) Good performance (S1 in P5)

(b) Bad performance (S2 in P6)
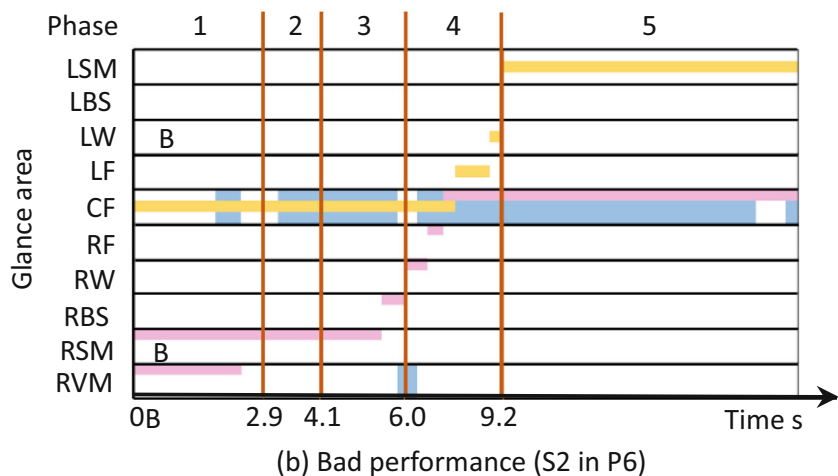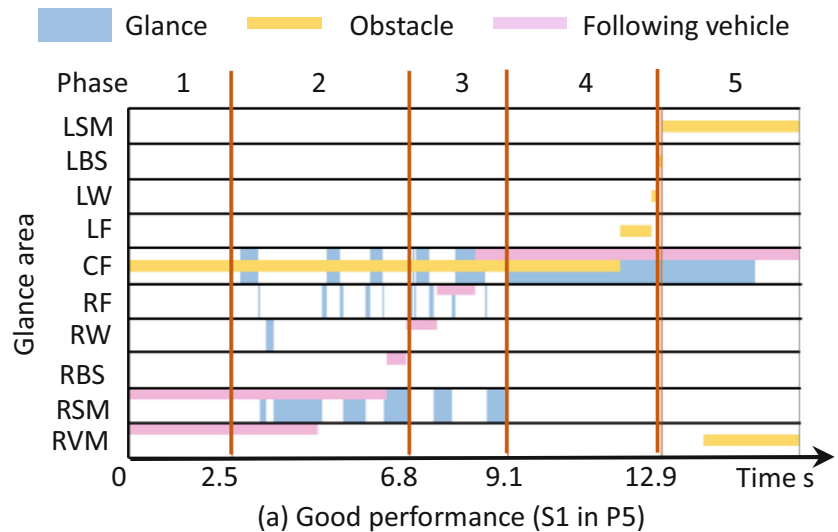
stamps as much as possible, we set a time flag on gaze data on the point when TOR was issued and synchronized them afterward. A total of twelve subjects (11 males and 1 female, age M: 23 years, SD: 1.1) participated in the experiments. All of them had normal or corrected-to-normal vision. Before the experiment, we explained them about the TO process and

how to make the TO. The subjects practiced driving in the DS as much as they wanted until they have confidence. As listed in Table 2, the subjects performed all six different TO scenarios, respectively (one time). The order was randomized. Subjects were provided with monetary compensation for their contribution.

Fig. 12 Accuracy increased from 70 to 83% and the number of rest variables decreased from 34 to 14 after removing variables with the lowest CR
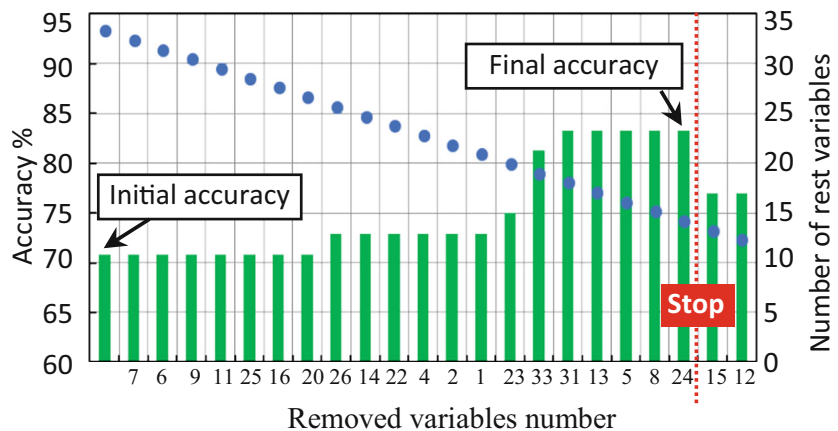
Table 4   Determining variables

#	Explanatory variables (time s)	Weight	CR %
3	[Phase 1] Gaze on RSM	0.1	0.01
10	[Phase 2] Gaze on F	128.0	6.51
12	[Phase 2] Gaze on RSM	144.6	7.36
15	[Phase 2] Gaze on RSM (FV exists)	125.2	6.37
17	[Phase 2] Gaze on FV	132.0	6.71
18	[Phase 2] Gaze on CF (obstacle exists)	70.6	3.59
19	[Phase 1&2] Gaze on F	69.1	3.51
21	[Phase 1&2] Gaze on RSM	144.7	7.36
27	[Phase 1&2] Gaze on CF (obstacle exists)	83.0	4.22
28	From TOR–TO	−225.7	11.48
29	From TO–LC	195.9	9.96
30	From TOR–first look at F	−282.6	14.38
32	From last look at RSM–LC	−205.7	10.46
34	From last look at RVM–LC	−158.7	8.07

5 Results and Analysis

In this section, we describe the results of experiments and SVM, and discuss the implications.

5.1 Takeover Performance and Glance Map

The driving performance is judged as good or bad by the method described in Section 2.4. Table 3 lists the evaluation results of performance in each subject in each TO scenario. Here, P1 and P2 were used for the baseline. Following analysis focuses on P3–P6. We found that only 58% of participants safely performed the TO. As stated in Section 1, drivers of a highly automated vehicles are likely to have low or zero SA, and which can be a reason for the above result. Using obtained drivers' glance areas and driving-relevant objects' position, we made glance maps to visually show drivers' glance behavior. Figures 11a, b are for good and bad performance, respectively. In Fig. 11a, S1 looked at the following vehicle several times through RSM, and then safely made lane change after the following vehicle passing in P5. In contrast, in Fig. 11b,

S2 almost did not check surroundings, so S2 was not aware of the following vehicle and made a collision in P6. This suggests that the glance behavior has a significant impact on SA and the driving performance.

5.2 Determining Variables and Accuracy

As stated in Section 2.5, we explored the determining variables by following steps. At first, we calculated the CR of all explanatory variables, and the result is listed in the right-side two columns of Table 1. Then, we removed variables with a low CR one by one. Figure 12 shows how the estimation accuracy changed over the removal. The result shows that the accuracy increases as the number of explanatory variables decreases, and the accuracy reached the maximum 83% when the number of variables was 14. The removed variables, such as the time driver gazed on F (#1) in phase 1 and RVM (#2, #11), can be regarded as no or negative impact on SA estimation. Instead of RVM, as an example shown in Fig. 11a, many drivers gazed on F (#10) and RSM (#12) to check surroundings. The remaining variables are defined as determining variables (Table 4). From the result, we could find followings.

- The number of variables in phase 2 was greater than that in phase 1, which indicates that glance in phase 2 is more important for good SA. The traffic environment changes over time, so drivers should keep attention to it. Moreover, in unscheduled TO, drivers are suddenly imposed on large burden. Even if drivers checked the surrounding in phase 1, they need to check the surrounding again in phase 2.
- w of determining variables #28, 30, 32, and 34 was a negative value, which indicates that the shorter the time spent, the better SA is. For example, drivers will have a high probability of good SA if they start looking at F and approve the TO fast.
- The determining variables #15, 18, and 27 show that the awareness to following the vehicle and the obstacle is significant to SA estimation. This meets our hypothesis. However, their CRs were smaller than the variables #10, 12, 17, and 21, which did not consider driving-relevant objects. This indicates that looking at glance areas is not only for checking if
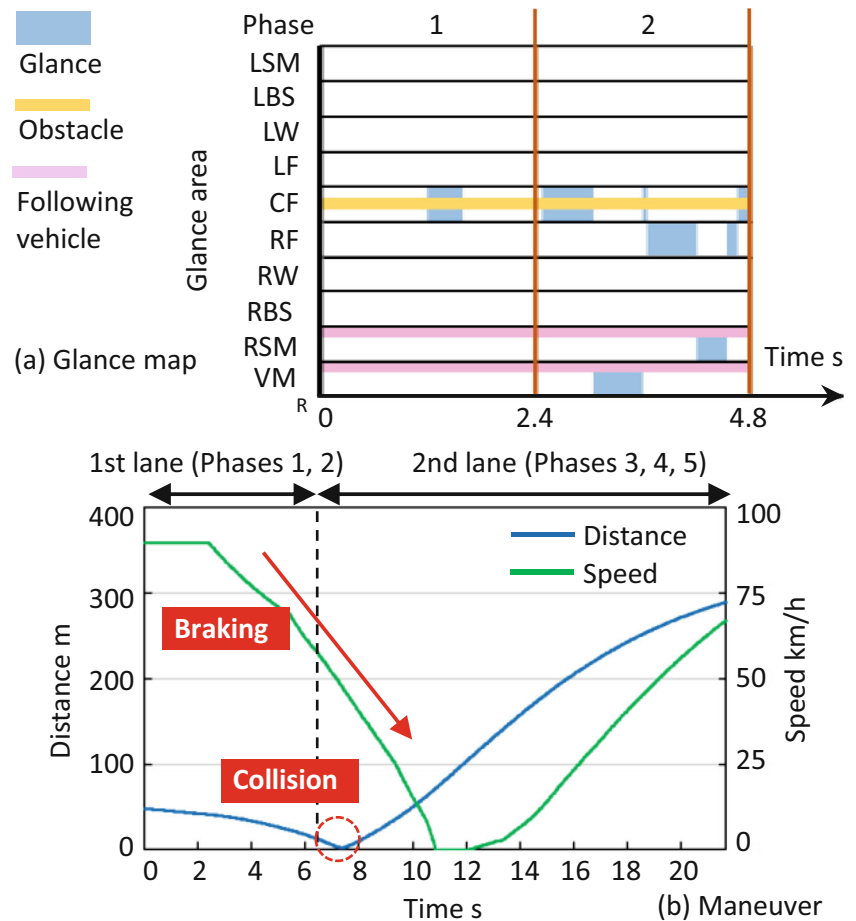
Table 5   Result of leave-one-out cross validation (LOOCV)

Subject No.	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
P3	1	0	1	1	0	0	1	1	1	1	0	1
P4	0	1	1	1	0	1	1	1	1	1	1	1
P5	1	0	1	1	0	1	0	1	1	1	1	0
P6	1	1	1	1	0	1	0	1	1	1	0	0

■: Incorrectly estimate good performance as bad

■: Incorrectly estimate bad performance as good

Fig. 13 Glance map and maneuver information of S1 in P5. **a** Glance map. **b** Maneuver

(a) Glance map

(b) Maneuver

there is the object but also for understanding the state of ego-vehicle, such as speed, direction, and position.

6 Discussion

6.1 Reason of Estimation Error

The estimation accuracy eventually came to 83% and left a 17% error. The result of SA estimation is listed in Table 5. To prevent accidents, the accuracy is expected to be near to 100%. Here, we discuss about reasons of error.

6.1.1 False Estimation (Bad as Good)

From Table 5, S1 and S2 had bad performance in P5 and P6, respectively. However, the model estimated them as good performance.

- Figure 13a shows the glance map of S1 in P5. The glance behavior looks good because S1 checked both the obstacle and following vehicle several times. However, as Fig. 13b shows, S1 entered the next lane at 0 km/h, which can be regarded as mis-stepping on pedals, and then caused the collision.
- Figure 14a shows the glance map of S2 in P6. S2 also checked both the obstacle and following vehicle and this seems good SA. However, as Fig. 14b shows, S2 moderately controlled the steering wheel but found that the amount was not enough, and then steeply controlled it, which led to bad performance.

From the above, we found that unscheduled TOs likely caused maneuver errors. This indicates that AD could make drivers lose their sense of pedal and steering operation and the sudden TOR stressed on drivers.

6.1.2 False Estimation (Good as Bad)

From Table 5, S2 had good performance in P5, but the model estimated it as bad performance. Figure 15a shows the glance map of S2 in P5. S2 checked the environment but S2 missed the following vehicle. Thus, S2's SA could be bad. However, as

Fig. 14 Glance map and maneuver information of S2 in P6. **a** Glance map. **b** Maneuver
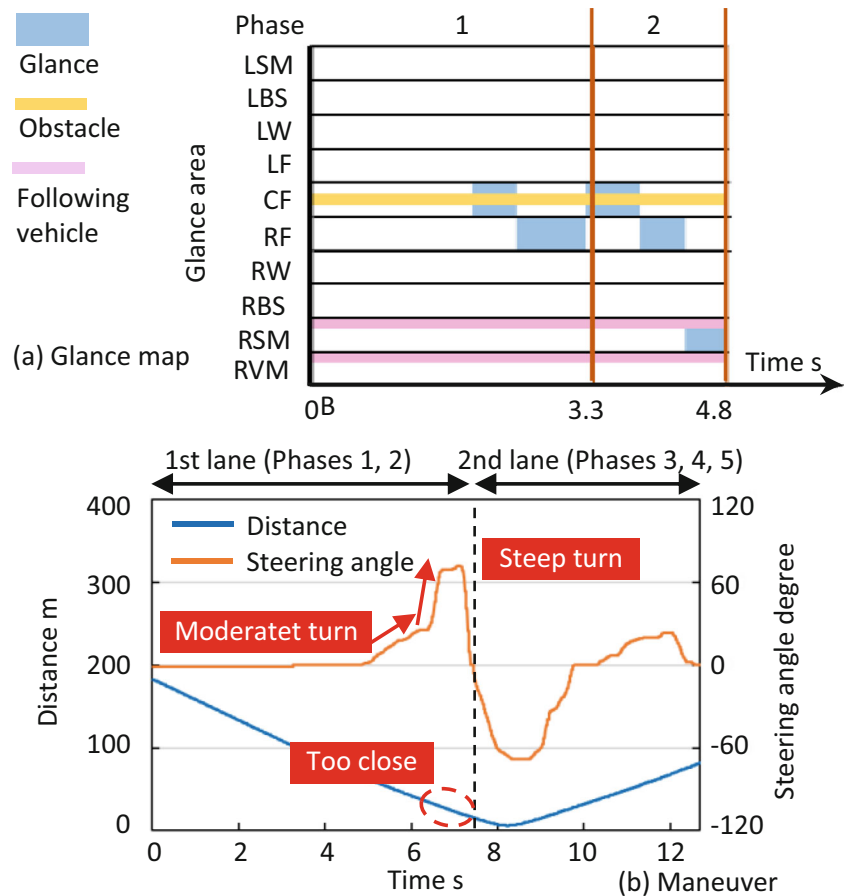

(a) Glance map

(b) Maneuver

Fig. 15b, c show, S2 speeded up and started lane change fast, so that the ego-vehicle successfully entered the next lane before the following vehicle catches it. From the above, we found that the driver could safely complete TO by chance, even if they did not check surroundings well. However, in such case, the system should estimate the driver's SA as bad, and assist the SA or reject the TO.

From the above analyses, we could find that the performance estimation error was caused by maneuver error under good SA or occasional success under bad SA, not SA estimation error. A steering and pedal control support would be necessary to make drivers maneuver smoothly the vehicle and reduce their maneuver effort.

6.2 Difference of Individual Glance Behavior

Here, we performed leave-one-subject-out cross validation (LOSO-CV), which uses one subject's data for testing and other subjects' data for training [22], to analyze individual differences of glance behavior (i.e., generality of the model). If the accuracy
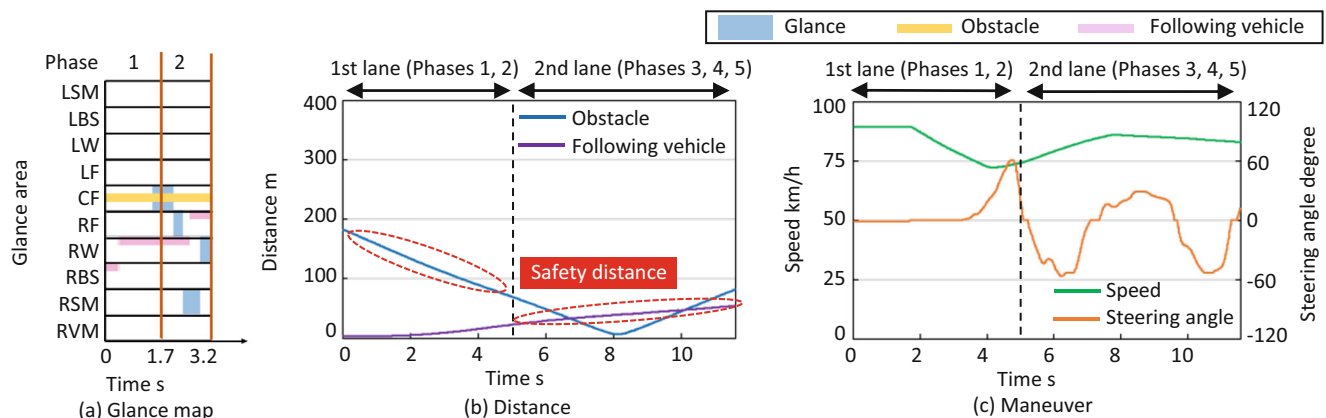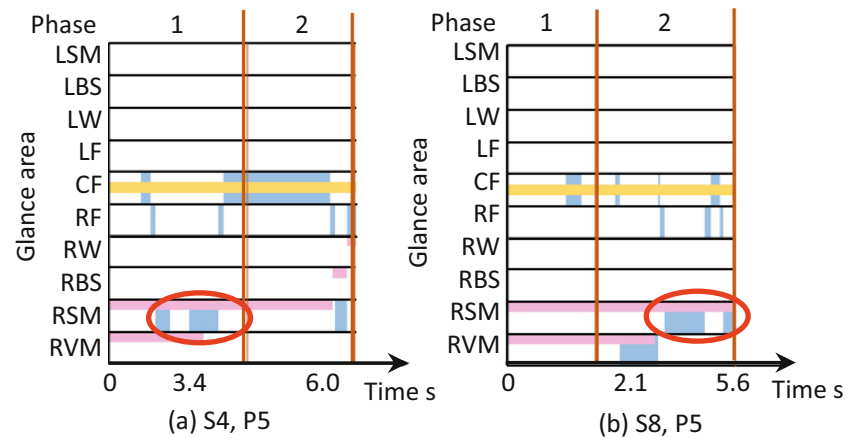

(a) Glance map

(b) Distance

(c) Maneuver

Fig. 15 Glance map and maneuver information of S2 in P5. **a** Glance map. **b** Distance. **c** Maneuver

Fig. 16 Individual difference for good performance. S4 checked surroundings most in phase 1, while S8 in phase 2. **a** S4, P5. **b** S8, P5

(a) S4, P5

(b) S8, P5

of LOSO-CV is much lower than that of LOOCV, the glance behavior differs individually, otherwise, the difference does not exist. As a result, we obtained that the accuracy of LOSO-CV was 52%, which was much lower than that of LOOCV (83%). The result indicates that glance behavior surely differed individually. Next, we discuss how individual glance behaviors are different.

6.2.1 Glance Behavior of Good Performance

From Table 5, S4 and S8 had good performance at all patterns, while they had different glance behavior. Figure 16 shows the glance maps of S1 and S8 in P5. Both two subjects checked surrounding well and had good SA, but S4 checked surrounding most in phase 1, while S8 almost checked surrounding in phase 2. Subjects should push the TO button when they understand the surrounding traffic environment. However, some subjects forgot this because of the long AD working time and the pressure of TOR.

6.2.2 Glance Behavior of Bad Performance

From Table 3, S5 and S7 had bad performance at almost all patterns, while they had different glance behavior. Figure 17

shows the glance maps of S5 and S7 in P4. From Table 4, we found that the shorter the time between the last look at RSM and lane change (#32), the better SA was. However, the time S5's last look at RSM to lane change was long, so it led to bad performance, On the other hand, S7 looked RSM just before lane change, but the following vehicle was not already there, so S7 missed the following vehicle and had bad performance.

From the above analysis, we found that the determining variables interacted with each other and affected on the performance of SA. This meets our hypothesis. Moreover, a SA assistant method should be designed according to ways led to bad performance. These findings suggest that our proposal to use a combination of various explanatory variables are useful for SA estimation.

6.3 Comparison with Previous Model

We compared the accuracy of proposed model with the previous model [3]. We trained and evaluated the previous SA estimation model and evaluated it using the data collected in this study. Since the previous system only considers the glance time and areas but driving-relevant objects, we extracted the time on the three glance areas (F, RSM, and RVM) as input to derive

Fig. 17 Individual difference for bad performance. S5's last look at RSM to LC is long, while S7 missed following vehicle. **a** S5, P4. **b** S7, P4

(a) S5, P4

(b) S7, P4
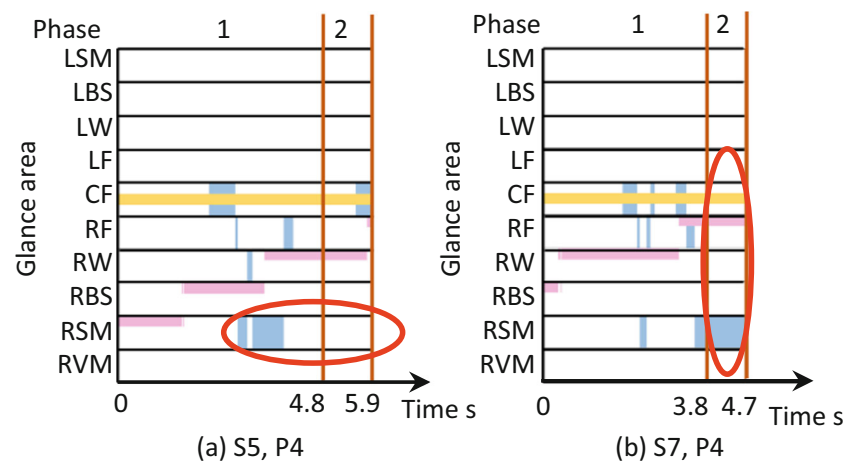
standard glance model. If a driver looked all standard glance area within the standard glance time range respectively, the driver's SA is good. We performed LOOCV. The result was 56%, which was lower than the proposed model (83%). Moreover, to find factors that improves the accuracy, we evaluated the accuracy of SVM by using the same data as the previous model. The LOOCV resulted in 69%, which was higher than the previous model (56%) but lower than the proposed model (83%). We confirm from the result that the proposed model could improve the accuracy by 'considering the surrounding situation' and 'treating variables as combination'.

7 Conclusion and Future Works

To estimate driver's situational awareness (SA), it is important to analyze how a driver percepts objects in the traffic environment and consider the interactions between glance features. We thus proposed a traffic environment considered SA estimation model. We conducted a takeover experiment on a driving simulator, and created an explanatory variable set that include not only driver's glance information, but also other vehicles and obstacles information. We introduced support vector machine (SVM) to develop the SA estimation model. After a data processing, determining variables were obtained, and they indicate that the drivers gaze on objects is important for estimating drivers SA. As a result of SA estimation, SVM obtained 83% accuracy with the determining available and the awareness to driving-relevant objects was significant to SA estimation. The comparison with the previous model revealed that the accuracy of the proposed model was improved by 'considering the surrounding situation' and 'treating variables as combination'.

In this paper, we only analyzed glance behavior after TOR, so we will address estimating drivers' SA before TOR (i.e., while AD). We will increase the data sample including different genders and ages. Moreover, we will integrate the proposed model to an assistant system, which can improve driver's SA or stop the vehicle depending on driver's SA level, and a driver workload estimation system [23], to provide assistant suitable the driver's workload. Furthermore, since unscheduled TO will lead to maneuver error, we will analyze the impact of driver's maneuver, and design an assistant system for enhancing driver's maneuver ability.

References

1. SAE On-Road Automated Vehicle Standards Committee. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (2016)
2. Kamezaki, K., Hayashi, H., Manawadu, U.E., Sugano, S.: Human-centered intervention based on tactical-level input in unscheduled takeover scenarios for highly-automated vehicles. Int. J. Intell. Transp. Syst. Res. **18**, 1–10 (2019)
3. Hayashi, H., Kamezaki, M., Manawadu, U. E., Kawano, T., Ema, T., Tomita, T., Catherine, L., Sugano, S.: A driver situational awareness estimation system based on standard glance model for unscheduled takeover situations. In: Proc. IEEE Intell. Vehicles Symposium, pp. 718–723 (2019)
4. Winter, J.C.F., Happee, R., Martens, M.H., Stantond, N.A.: Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transp. Res. Part F Traffic. Psychol. Behav. **27**(Part B), 196–217 (2014)
5. Zeeb, K., Buchner, A., Schrauf, M.: Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid. Anal. Prev. **92**, 230–239 (2016)
6. Punitkumar, B., Babji, S., Rajagopalan, S.: Quantifying situation awareness of control room operator using eye-gaze behavior. Comput. Chem. Eng. **106**(2), 191–201 (2017)
7. Tyron, L., Natasha, M.: Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. Transp. Res. Part C Emerg. Technol. **76**, 35–50 (2017)
8. Endsley, M.R.: Toward a theory of situation awareness in dynamic system. Hum. Factors. **37**(1), 33–64 (1995)
9. Fletcher, L., Zelinsky, A.: Driver inattention detection based on eye gaze-road event correlation. Int. J. Rob. Res. **2**(6), 774–801 (2009)
10. Tawari, A., Chen, K. H., Trivedi, M. M.: Where is the driver looking: analysis of head, eye and iris for robust gaze zone estimation. in: Proc. Int. Conf. Intelligent Transportation Systems, pp. 988–994 (2014)
11. Fridman, L., Toyoda, H., Seaman, S., Seppelt, B., Angell, L., Lee, J., Mehler, B., Reimer, B.: What can be predicted from six seconds of driver glances? In: Proc. Conf. Human Factors in Computing Systems, pp. 2805–2813 (2017)
12. Trzcinski, T., Rokita, P.: Predicting popularity of online videos using support vector regression. IEEE Trans. Multimedia. **19**(11), 2561–2570 (2017)
13. Sanaeifar, A., Bakhshipour, A., Guardia, M.: Prediction of banana quality indices from color features using support vector regression. Talanta. **148**(1), 54–61 (2016)
14. Kazutoshi, T., Bono, L., Dragan, A., Takio, K., Mikio, K., Natsuo, O., Takahiro, S.: Prediction of carcinogenicity for diverse chemicals based on substructure grouping and DVM modeling. Mol. Divers. **14**(4), 789–802 (2010)
15. Shimofuji, S., Matsui, M., Muramoto, Y., Moriyama, H., Kato, R., Hoki, Y., Uehigashi, H.: Machine learning in analyses of the relationship between Japanese sake physicochemical features and comprehensive evaluations. Japan J. Food Eng. **21**(1), 37–50 (2020)
16. Suykens, J.A.K., Vandewalle, J.: Least squares support vector machine classifiers. Neural. Process. Lett. **9**, 193–300 (1999)
17. Chang, C. C., Lin, C. J.: LIBSVM—a Library for Support Vector Machines, https://www.csie.ntu.tw/~cjlin/libsvm/, Last Accessed: 2020-06-01

18. Tzu-Tsung, W.: Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recogn. **48**(9), 2839–2846 (2015)
19. Smart eye, https://smarteye.se/research-instruments/se-pro/, Last Accessed: 2020-06-01
20. Manawadu, U. E., Ishikawa, M., Kamezaki, M., Sugano, S.: Analysis of individual driving experience in autonomous and human-driven vehicles using a driving Simulator, in: Proc. IEEE/ASME Int. Conf. Advanced Intelligent Mechatornics, pp. 299–304 (2015)
21. Eriksson, A., Stanton, N.A.: Takeover time in highly automated vehicles: noncritical transitions to and from manual control. Hum. Factors. **59**(4), 689–705 (2017)
22. Scheurer, S., Tedesco, S., Brown, K. N., O'Flynn, B.: Sensor and feature selection for an emergency first responders activity recognition system, in: Proc. IEEE SENSORS, pp. 1–3 (2017)
23. Manawadu, U. E., Kawano T., Murata, S., Kamezaki, K., Muramatsu, J., Sugano, S.: Multiclass classification of driver perceived workload using long short-term memory based recurrent neural network, in Proc. IEEE Intell. Vehicles Symp., pp. 2009–2014 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Mitsuhiro Kamezaki received the B.S., M.S., and Dr. of Engineering degrees in mechanical engineering from Waseda University, Japan, in 2005, 2007, and 2010, respectively. Since April 2010 to March 2013, he was a Research Associate in the Department of Mordern Mechanical Engineering, Waseda University. Since April 2013 to March 2018, he was an Assistant Professor in the Research institute for Science and Engineering, Waseda Univ. Since April 2018, he has been an Associate Professor in the Research Institute for Science and Engineering, Waseda Univ. His interests include intelligent machine system, man-machine interface, and operator support system. Dr. Kamezaki received the Young Investigation Excellence Award from RSJ in 2016, the Best Paper Award from the IEEE/ASME AIM 2016, and the Best Paper Awards from RSJ and SICE. He is a Member of the JSAE, SICE, RSJ, JSME, and IEEE RAS.

Hiroaki Hayashi received the B.S. and M.S., in Mechanical Engineering from Waseda University in 2017, and 2019, respectively. Currently he is a Graduate Student in the Department of Modern Mechanical Engineering in Waseda University. His research interests include human-machine interface systems, driving simulation, and driving automation systems. He received young fellow award from the Japan Society of Mechanical Engineers in 2020. He is a member of the JSME.

Shigeki Sugano received the B.S., M.S., and Dr. of Engineering degrees in mechanical engineering from Waseda University, Japan, in 1981, 1983, and 1989, respectively. From 1986 to 1990, he was a Research Associate in Waseda University. Since 1990, he has been a Faculty Member in the Department of Mechanical Engineering, Waseda University, where he is currently a Professor. From 1993 to 1994, he was a Visiting Scholar in the Mechanical Engineering Department, Stanford University. He is a Member of the Humanoid Robotics Institute, Waseda University. Since 2014, he serves the Dean of the School of Creative Science and Engineering, Waseda University. His research interests include human-symbiotic anthropomorphic robot design, dexterous and safety manipulator, and human-robot communication. He served as the General Chair of IROS2013 in Tokyo and SICE Annual Conference 2011. Dr. Sugano is a fellow member of JSME, SICE, RSJ, and IEEE RAS.

Oka Naoki received the B.S. in Mechanical Engineering from Waseda University in 2018, and currently is a Graduate Student in the Department of Modern Mechanical Engineering in Waseda University. His research interests include human-machine interface systems, driving simulation, and driving automation systems. He is a member of the JSME