Modeling individual differences in risk feeling of autonomous driving behavior with a prediction error

Takashi HASHIMOTO* and Hideyoshi YANAGISAWA*
* The University of Tokyo, Department of Mechanical Engineering, Design Engineering Lab.
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8654, Japan
E-mail: hide@mech.t.u-tokyo.ac.jp

Received: 3 February 2020; Revised: 10 May 2020; Accepted: 15 June 2020

Abstract
Studies on autonomous driving systems are being conducted to realize a safe and secure mobile society. According to these studies, the safety of the system and the driver’s feeling of safety often do not match. To enable automatic driving control according to the driver’s comfort, previous researchers introduced a risk feeling index to quantify the driver’s feeling of safety. These studies, however, did not consider the effects of the drivers’ individual differences for the risk feeling index. We hypothesized that the risk perception was formed by the difference between the actual autonomous driving behavior and the driver's prediction for it. The driver's risk feeling is expected to vary depending on his/her prior prediction of autonomous driving behavior. We considered that this difference, which we term as the prediction error, explains an individual difference in the risk perception. The purpose of this study is to reveal the effect of the prediction error on the risk perception. We conducted an experiment to investigate the effect of these factors on the risk feeling in the overtaking scene using a driving simulator. We obtained two types of subjective evaluations to examine the effects of the prediction errors on risk perception. One is perceived risk based on the driver’s original prediction, and the other is perceived risk based on the predictions that we manipulated by conducting a learning session during the experiment. The results suggest that the risk is perceived based on both the experimentally manipulated prediction and the individual driving characteristic of anxiety while driving. These findings will enable the development of a secure automatic driving system that suits each driver by controlling the system appropriately based on the driving characteristic of the driver.

Keywords: Design, Autonomous driving, Risk feeling, Individual differences, Prediction error

1. Introduction

Research and development of an automatic driving system has been conducted to realize a safe and secure mobile society (Suganuma, 2017). The technological improvement of safety is expected, but driver safety includes psychological factors that cannot be achieved only by technical aspects. Hence, the safety of the system itself does not always match the driver’s feeling of safety. When the driver does not feel comfortable with the system even though the autonomous driving technology guarantees adequate safety, autonomous vehicles may not become prominent. Thus, it is important to design a system based not only on technical safety but also on the driver’s feeling of safety.

Recently, studies on safety for autonomous driving systems have been conducted. For example, Yoon et al. (2016) proposed an algorithm for improving the lane change detection accuracy as a safe and reliable autonomous driving system. McAllister et al. (2017) proposed a method of building reassurance by giving explanations such that the behavior of autonomous driving can be easily interpreted. These studies, however, did not consider the driver's perception of the driving behavior. It is necessary to deal with the driver's feeling of safety in a safe autonomous driving system. When considering studies that are focused on automotive control based on the driver’s emotion, the risk feeling index has been investigated (Kondoh et al., 2004, 2008, Nakano et al., 2018). However, there are not enough studies on the effects of individual differences in the risk perceptions. It is necessary to consider the effect of individual differences while indexing...
the risk feeling. Based on the expectation effect theory that we previously proposed (Yanagisawa, 2016), we hypothesized that the risk perception was formed by the prediction error: the difference between the actual autonomous driving behavior and the driver's prediction. We considered that risk perceptions are relative to the prior expectation of the driving behavior. For example, even if the autonomous driving behavior is the same, we can assume that those who drive aggressively are less likely to perceive the risk, and those who are safety-oriented are more likely to perceive the risk. We, therefore, considered that the prediction error explains an individual difference in the risk perception. The purpose of this study is to reveal the effect of the prediction error in autonomous driving behavior on the risk perception and the model driver’s risk feeling for autonomous driving behavior.

2. Related works

2.1 Risk feeling indexes in automotive driving behavior

During a study that indexed the risk feeling perceived by the driver, Kondo (2008) proposed that the risk perceived by the driver while approaching a preceding vehicle can be given as follows.

\[ RF = \frac{5}{TTC} + \frac{1}{THW} \]  

where RF is the risk feeling that is perceived by the driver, time to collision (TTC) is the time required for two vehicles to collide if they continue at their present speed and are on the same path, and time head way (THW) is the time difference between any two successive vehicles when they cross a given point. Equation (1) does not take individual differences into account.

2.2 Studies on individual differences in the risk feeling of the driving behavior

It is known that there are individual differences in the timing of the brake operation when approaching a preceding vehicle. Kondo (2004) proposed a personalized risk feeling index that is normalized by individual driving characteristic when approaching the preceding vehicle as follows.

\[ RF_{ind} = \left(\frac{1}{THW}\right)/\left(\frac{1}{THW_{ind}}\right) + \left(\frac{1}{TTC}\right)/\left(\frac{1}{TTC_{ind}}\right) \]  

Here, \(TTC_{ind}\) and \(THW_{ind}\) are the values of \(TTC\) and \(THW\) at the time of the brake operation while approaching the preceding vehicle. It has been confirmed that this personalized risk feeling index has improved the accuracy of identifying when the driver changes from the accelerator to the brake, compared to the conventional index. Since this index is intended for a driving assistance system equivalent to an autonomous driving level of 1 or 2, in which the driver is responsible for monitoring and responding to safe driving, it cannot be applied to autonomous driving systems of level 3 or higher. This is because the system is responsible for monitoring and responding to safe driving. It may, however, be applied to a limited situation such as in the context of our study. In this paper, we analyzed the factors of individual differences that affect the risk perception and indexed the risk perception according to the extracted factors. We modeled the risk perception at the automatic driving level of 3, in which it is necessary to control the autonomous driving system that is based on the individual belief of the driving behavior.

2.3 Studies on individual differences in the risk feeling of driving behavior

In this study, based on the fact that the factors determine the individual differences in the risk perception and include personality, situation cognition, and option cognition (Ueichi, 1998), we analyzed these factors of individual differences. Situation cognition is an individual’s way of thinking and knowledge for each situation, and option cognition is the evaluation of the actions that can be selected. Personality is a relatively stable characteristic of a person, and therefore the effect of option cognition is small in
autonomous driving, where the system is responsible for the safe driving. In the quantification of the individual differences in risk feeling, it is necessary to control the autonomous driving system in accordance with the expectations of the individual. Therefore, we examined the individual differences that result from the situation cognitive factor. We made an assumption that the situation cognitive factor corresponded to the prior prediction in the expectation effect theory (Yanagisawa, 2016). In this study, we examined the expected value and the uncertainty of the probability distribution of the prior prediction in the expectation effect theory as a factor of individual differences in the perception of the risk.

### 3. Modeling individual differences of the risk perception in autonomous driving behavior

#### 3.1 Modeling individual differences in risk feeling based on the expectation effect theory

Based on the analysis of the factors that define the individual differences in the risk perception, we indexed the risk feeling by using the expected value and the uncertainty of the prior distribution (probability distribution of the prior prediction) in the expectation effect theory (Yanagisawa, 2016) as factors of individual differences. In the expectation effect theory, perception is defined as the Bayesian estimate of the physical quantities. The Bayesian posterior, the distribution of the perception, is formed by multiplying the prior. The distribution of perception and the likelihood function are formed from the sensory stimuli. Thus, the prior expectation affects the posterior perception as shown in Fig. 1. This effect is called the expectation effect. The prior \( p(\theta) \) stochastically predicts the physical quantity \( \theta \) by learning the frequency distribution of the physical quantity. The brain estimates the physical quantity based on the neural signal pattern \( R \) as the likelihood function \( p(R|\theta) \). The posterior \( p(\theta|R) \) follows the Bayesian estimate, and the posterior estimate \( \hat{\theta} \) is the value of the maximum probability of the posterior. The expectation effect \( \epsilon \) is defined as follows.

\[
\epsilon = \hat{\theta} - \theta_{\text{lik}}
\]

The prediction error \( \delta \) is the difference between the expected value of the prior distribution \( E[p(\theta)] \) and the maximum likelihood estimate \( \theta_{\text{lik}} \) and is defined as follows.

\[
\delta = \theta_{\text{lik}} - E[p(\theta)]
\]

It is known that the larger the prediction error, the more attention a person dedicates to obtaining information and to learning new stimuli to update the existing knowledge (Christoph, 2018). There are two types of expectation effects: assimilation, which underestimates the prediction error when the prediction error is small, and contrast, which overestimates the prediction error when the prediction error is large. The variance \( V[p(\theta)] \) of the prior distribution is called uncertainty, and the magnitude of the uncertainty affects the strength of the expectation effect. When the uncertainty is high, the magnitude of the expectation effect decreases, and when the uncertainty is low, the magnitude of the expectation effect increases. As the prediction error affects the expectation effect, we examined the effect of the prediction error on the driver’s risk perception. We assumed that the difference between the expected value of the driving characteristic of the individual and the driving characteristic of the automatic driving system corresponds to the prediction error.

Fig. 1  Attractive influence of the prior in the Bayesian estimate to the posterior. A posterior distribution is formed by the product of a prior distribution and a likelihood distribution. The value of the maximum probability of the posterior is the posterior estimate \( \hat{\theta} \).
Based on the expectation effect theory, we hypothesized that the difference between the prediction based on the individual driving characteristic and the actual driving behavior (i.e., prediction error) affects the risk perception. In other words, if the actual driving behavior is more dangerous than predicted, the sense of risk increases, and if it is safer, the sense of risk decreases.

### 3.2 Formulation of individual differences

The most common cause of a traffic accident is a rear-end collision with a vehicle that has stopped in a lane (National Police Agency Traffic Bureau, 2018). In this study, based on the following hypothesis, we examined the risk feeling index while considering the effect of the individual differences by using the distance from the preceding vehicle while overtaking a vehicle that has stopped.

Hypothesis: the difference between a prediction based on the individual driving characteristic and the actual driving behavior (i.e., prediction error) provides a sense of risk.

In the condition where the prediction error from the expected value of the individual driving characteristic (e.g., inter-vehicle distance) is small, it is believed that the risk feeling is evaluated to be smaller than the actual risk due to the effect of the assimilation of the expectation effect. In the condition where the prediction error is large, the risk feeling is evaluated to be larger than the actual risk due to the effect of the contrast effect. In addition, it is believed that there is no change in the risk feeling even if the prediction error further increases. Therefore, we considered that the risk perception follows the S-curved-function such as the sigmoidal function. Based on the hypothesis, we formulated the risk feeling index while taking into account the effect of the prediction error as in Eq. (5). The plot is shown in Fig. 2.

\[
\text{Risk feeling considering the effect of the prediction error} = \frac{a}{1 + \exp (b \times (d_{\text{actual}} - d_{\text{ind}}) + c)}
\]  

(5)

Where, \(d_{\text{actual}}\) is the inter-vehicle distance between the host vehicle and the vehicle in front that has stopped in the lane, and \(d_{\text{ind}}\) is the expected value of the individual driving characteristic. In addition, \(a\) and \(b\) are positive constants, and \(c\) is an arbitrary constant.

![Fig. 2 The plot of the risk feeling index considering the effect of the prediction error (Eq. (5)). This curve shows that the perceived risk changes depending on the difference between a prediction based on the individual driving characteristic and the actual driving behavior (prediction error).](image-url)
4. Experiment on the effect of the prediction error for the risk feeling

4.1 Method

To verify the hypothesis discussed in section 3.2, we investigated the interaction effects of the actual driving behavior and its prior prediction on the risk perception in the overtaking scene. We used a driving simulator (the specification is shown in Table 1) and presented an overtaking scene on a two-lane road on one side to the participants. Figure 3 shows the details of the overtaking scene. In the overtaking scene, a participant’s vehicle approaches the vehicle that stopped in the center of the driving lane from behind and it changes lanes to overtake it. We manipulated the prediction and the actual driving behavior by using the inter-vehicle distance right before changing lanes in the overtaking scene and obtained the participants’ subjective evaluations of the risk perception. We used the inter-vehicle distance right before changing lanes based on the individual prediction, \( d_{\text{ind}} \), as an individual driving characteristic because we considered that the drivers change lanes with the inter-vehicle distance based on their own sense. Ten male students from the University of Tokyo in their twenties who have driver's licenses took part as participants. All of the participants were supposed to drive less than once a week to avoid the effect of past driving experience on the experimental results. We assumed that by having less driving experience, it would be easier to control the individual driving characteristic in the experiment. In order to obtain the participants’ evaluation, we prepared three types of scenes: the manual driving scene, the learning scene, and the evaluation scene. The manual driving scene is a scene of manual driving to obtain the individual characteristic \( d_{\text{ind}} \). The learning scene is a scene to learn the autonomous driving behavior that changes lanes at the inter-vehicle distance \( d_{\text{adapt}} \) in order to manipulate the individual driving characteristic from \( d_{\text{ind}} \) to \( d_{\text{adapt}} \). The evaluation scene is a scene for the participants to evaluate the perceived risk in the autonomous driving behavior that changes lanes at the inter-vehicle distance \( d_{\text{actual}} \). We manipulated the inter-vehicle distance in both the evaluation scene and the learning scene as shown in Table 2. We examined the effects of these three distances (\( d_{\text{ind}} \), \( d_{\text{adapt}} \), and \( d_{\text{actual}} \)) on the participants’ risk perception. We obtained two types of evaluations for the participants: perceived risk before the learning session (Evaluation 1) and the perceived risk after the learning session (Evaluation 2). We investigated the effect of the individual driving characteristic on the risk perception by Evaluation 1, and the effect of the manipulated individual driving characteristic on the risk perception by Evaluation 2.

Table 1  Specifications of the driving simulator used in the experiment.

| Equipment       | Model number               | Remarks                        |
|-----------------|----------------------------|--------------------------------|
| Driving simulator | FORUM 8 UC-win Road Ver.10 DS |                                               |
| Steering and pedal | Logitech G29               |                                               |
| Display         | DELL E248WFPb             | 24 inches (1920×1200)               |
| Display         | DELL 2407WFPб             | 24 inches (1920×1200)               |
| Display         | BenQ ET-0009-B            | 24 inches (1920×1200)               |
| PC              | DELL Precision Tower 7810 |                                               |

Fig. 3  The overtaking scene and the definition of the parameters.
Figure 4 shows the flow of the experiment. In the manual driving scene, we asked the participants to drive manually ten times to obtain the initial individual driving characteristic $d_{ind}$. After the manual scene, we conducted the evaluation scene and obtained Evaluation 1 for each of the 16 evaluation scenes. In order to obtain Evaluation 2, we asked the participants to evaluate the evaluation scene in the same way as Evaluation 1 after presenting the learning scene. We presented the learning scene under the same conditions 15 times to make the participants believe that the autonomous vehicle behaves like the learning scene. We obtained Evaluation 2 for each learning condition. We presented the 16 scenes with different levels of inter-vehicle distance in a random order that was arranged in advance for Evaluations 1 and 2. We divided the participants into two groups, and they were counterbalanced by reversing the presentation order of the evaluation scenes between the two groups. At the end of the experiment, we asked the participants to answer the "driving style check sheet" (DSQ) (Ishibashi, 2007) to obtain their personal attitudes and initiatives when driving (e.g., anxiety tendency).

Table 2  Experimental condition of the manual driving scene, evaluation scene, and learning scene.

|                | Length of $d_{actual}$ [m] | Length of $d_{adapt}$ [m] | Initial speed [km/h] | Traveling speed [km/h] | Initial inter-vehicle distance [m] | Number of presentations |
|----------------|-----------------------------|---------------------------|----------------------|------------------------|-----------------------------------|------------------------|
| Manual driving scene | -                           | -                         | 0                    | 100                    | 250                               | 10 times               |
| Evaluation scene    | 10, 12, 15, 18, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55 | -                         | 100                  | 100                    | 250                               | 1 time                 |
| Learning scene      | -                           | 10, 55                    | 100                  | 100                    | 250                               | 15 times               |

![Fig. 4  The flow of the experiment.](image)

In the manual driving session, the participants operated the steering wheel and the accelerator pedal and changed lanes to avoid a stopping car. We obtained the inter-vehicle distance, $d_{ind}$, in which the participant started to change lanes and marked it as the individual’s driving characteristic. We instructed the participants to operate the steering wheel only when changing lanes so that there is no steering input while approaching the stop vehicle ahead. We defined the timing right before the lane change as the timing of the steering input. In the learning session, we manipulated the individual driving characteristic $d_{ind}$ by presenting the learning scene to the participants to make them believe that the autonomous driving vehicle changes lanes at the inter-vehicle distance $d_{adapt}$. We instructed participants to watch the autonomous driving behavior in the learning scene. While presenting the learning scene to the participants, as a dummy task, we instructed them to press a button on the steering wheel an arbitrary number of times when they sense a risk. We introduced the dummy task to maintain the participants' concentration for the learning scene. In the evaluation session, we manipulated the inter-vehicle distance in which the autonomous vehicle switches lanes and obtained the evaluation of the risk perceived by the participants on nine levels as listed in Table 3.

Table 3  Evaluation words used in the evaluation session and the corresponding evaluation values.

| Evaluation word | very safe | safe | relatively safe | slightly safe | neither | slightly dangerous | relatively dangerous | very dangerous |
|-----------------|-----------|------|-----------------|---------------|---------|---------------------|----------------------|--------------|
| Evaluation score| -4        | -3   | -2              | -1            | 0       | +1                  | +2                   | +3           | +4           |
4.2 Results

Figure 5 shows the results of Evaluations 1 and 2 for two participants as examples of the evaluation data. The $d_{\text{ind}}$ of subject 1 (57.5 m) was shorter than that of subject 2 (88.9 m). Subject 1 showed shorter $d_{\text{actual}}$ where the score of risk feeling changes from positive to negative (21.3 m) than that of subject 2 (42.6 m). Figure 5 also shows that the evaluation of the risk feeling after learning at 55 m generally tends to be greater than that after learning at 10 m for both results. Figure 6 shows the relationship between $d_{\text{actual}}$ and the participants’ evaluation of the risk feeling based on the individual driving characteristic (Evaluation 1) for all participants. Figure 7 shows the relationship between $d_{\text{actual}}-d_{\text{ind}}$ and the participants’ evaluation of the risk feeling based on the individual driving characteristic (Evaluation 1). Figures 6 and 7 show that the negative correlation coefficient between $d_{\text{actual}}-d_{\text{ind}}$ and the risk feeling was greater than that between $d_{\text{actual}}$ and the risk feeling.

Fig. 5  Example of Evaluations 1 and 2 for the two participants. “Before learning” is the result of Evaluation 1, and “learning in 10 m” and “learning in 55 m” are the results of Evaluation 2. Subject 1’s average inter-vehicle distance just before changing lanes that was obtained in the manual driving scene (individual driving characteristic) was 57.5 m, whereas subject 2 was 88.9 m. Subject 1’s before learning linear approximation x-intercept is 21.3 m, whereas subject 2 is 42.6 m.

Fig. 6  Relationship between $d_{\text{actual}}$ and the risk feeling based on the individual driving characteristic. The correlation coefficient was $r = -0.770$. 
Fig. 7  Relationship between $d_{\text{actual}}$-$d_{\text{ind}}$ and the risk feeling based on the individual driving characteristic. The correlation coefficient was $r = -0.818$. The x-axis is the inter-vehicle distance $d_{\text{actual}}$ in the evaluation scene based on the inter-vehicle distance $d_{\text{ind}}$ acquired for each participant.

We defined the x-axis intercept as the distance $d_{\text{zero}}$ in which the risk feeling was estimated to be zero by linear approximation for each participant. Table 4 shows the correspondence between the estimated x-axis intercept $d_{\text{zero}}$ and the individual driving characteristic $d_{\text{ind}}$ for each participant. Figure 8 shows the relationship between DSQ’s anxiety tendency and $d_{\text{ind}}$-$d_{\text{zero}}$. We found significant positive correlations between them.

Table 4  Correspondence between the estimated x-axis intercept and the individual driving characteristic for each participant.

| Estimated x-axis intercept ($d_{\text{zero}}$) [m] | Individual driving characteristic ($d_{\text{ind}}$) [m] |
|-----------------------------------------------|-----------------------------------------------|
| 35.44                                         | 60.44                                         |
| 30.23                                         | 55.94                                         |
| 34.91                                         | 71.53                                         |
| 53.87                                         | 83.70                                         |
| 83.24                                         | 112.3                                         |
| 33.19                                         | 57.01                                         |
| 42.64                                         | 88.91                                         |
| 28.32                                         | 61.10                                         |
| 28.32                                         | 59.17                                         |
| 21.28                                         | 57.53                                         |

Fig. 8  Relation between the anxiety tendency and $d_{\text{ind}}$-$d_{\text{zero}}$. The correlation coefficient was $r = 0.703$.

Figure 9 shows the average and linear approximation of the risk feeling with respect to the inter-vehicle distance for each learning condition. The linear approximation correlation coefficients in the condition of the inter-vehicle distance (10 m, 55 m) were $r_{10m} = -0.984$ and $r_{55m} = -0.983$. Pearson’s uncorrelated test was significant ($p_{10m} = 0.00$, $p_{55m} = 0.00$). The analysis of variance (ANOVA) for the effects of the learning condition and the inter-vehicle distance for the risk feeling showed that the interaction between the factors was not significant ($F = 1.21$, $p = 0.275$) and the main effect of each factor was significant (learning condition: $F = 32.4$, $p = 0.00$, inter-vehicle distance: $F = 22.3$, $p = 0.00$).
In order to examine the sensitivity to the risk, we classified the participants’ evaluation values into a danger zone and a safety zone as shown in Fig. 10. In the danger zone, the participants’ evaluation values are greater than zero. In the safety zone, the participants’ evaluation values are less than zero. We calculated the slope of the linear approximation of the risk feeling in the danger and safety zones, respectively. Figure 11 shows the average value of the slope. We found that only the main effect of the zone condition was significant (F = 19.6, p < 0.01) as a result of the ANOVA for the effects of the learning condition and the zone condition on the slope of the linear approximation of the risk feeling.

Fig. 9  Relation between the risk feeling after learning and $d_{\text{actual}}$. The plot is an average of the evaluation values after learning for each learning condition (10 m, 55 m).

Fig. 10  Danger zone and safety zone classifications. The danger zone is configured with evaluation values after learning with a value of 0 or more, and the safety zone is configured with a value of 0 or less.
5. Discussion

From Fig. 6 and Fig. 7, we found that the correlation between the inter-vehicle distance based on the individual driving characteristic and the risk feeling is greater than that between the inter-vehicle distance alone and the risk feeling. This result supports the hypothesis that the risk is perceived based on the prediction error, the difference between the actual and the predicted driving behavior (i.e., $d_{\text{actual}} - d_{\text{ind}}$). We observed that the distance in which the risk feeling was estimated to be zero by linear approximation ($d_{\text{zero}}$) was not zero, and this was less than $d_{\text{ind}}$ in Table 4. This result indicates that there is a difference between the inter-vehicle distance in which the driver does not feel risk, and the expected value of the individual driving characteristic obtained in the manual driving scene. It is thought that the evaluation of the risk feeling is influenced by another factor, which has not been studied yet. The strong correlation between DSQ’s anxiety tendency and $d_{\text{ind}} - d_{\text{zero}}$ in Fig. 8 suggests that the personality tends to affect the evaluation of the risk feeling. This result supports the fact that the personality is a factor of individual differences in the risk perception.

Figures 6 and 7 show that the expected values for the individual driving characteristic affects the individual’s evaluation of the risk. Because the learning condition was significant as shown in Fig. 9, we can assume that the expected value of the individual driving characteristic was successfully manipulated by learning the lane change at a constant inter-vehicle distance. In addition, from Fig. 9, we observed that the evaluation of the risk under the learning condition of 10 m is smaller than that under the learning condition of 55 m at all inter-vehicle distances. This indicates that the expected values of the driving characteristic approached the learning condition, and the participants perceived the risk based on the updated belief of the autonomous driving characteristic (i.e., inter-vehicle distances).

From Fig. 11, we observed that the slope of the risk feeling line does not depend on the learning conditions, and that the slope in the danger zone is significantly greater than that in the safety zone. This suggests that there is asymmetry between the safety zone and the danger zone, and the perceptual sensitivity of the degree of danger is higher than the degree of safety. Prospect theory (Daniel at el., 1979) in behavioral economics proposes that losses are more sensitive than gains, which indicates that the participants are more sensitive to risk than safety. We believe that a person has a high sensitivity to risk as a defense instinct because the risk needs to be dealt with. We have proposed that the risk perception can be divided into a danger zone and a safety zone. Considering that a negative correlation was found between $d_{\text{actual}} - d_{\text{ind}}$ and the risk feeling based on the individual driving characteristic in Fig. 7, we modeled the risk feeling indexes for the danger zone and for the safety zone as Eq. (6) and Eq. (7).
\[
\text{Risk feeling considering the effect of the prediction error in the danger zone} = \frac{4}{1 + \exp(0.167 \times (d_{\text{actual}} - d_{\text{ind}}) + 8.12)}
\]

\[
\text{Risk feeling considering the effect of the prediction error in the safety zone} = \frac{4}{1 + \exp(0.260 \times (d_{\text{actual}} - d_{\text{ind}}^4) + 4.41)} - 4
\]

The adjusted coefficient of determination for Eq. (6) was \( R^2 = 0.240 \) and it was significant at a level of 1% (\( F = 31.9, p = 0.00 \)). The adjusted coefficient of determination for Eq. (7) was \( R^2 = 0.357 \) and it was also at a level of 1% (\( F = 43.2, p = 0.00 \)). The value of the adjusted coefficient of determination was not high. The regression curves using Eq. (6) and Eq. (7) are shown in Fig. 12, but the curves do not fit well. This is due to the learning conditions \( d_{\text{adapt}} \) in place of the individual driving characteristic \( d_{\text{ind}} \) when performing the regression analysis. The poor fit of the regression curve may be due to the individual driving characteristic which is not the same as the learning conditions, or the narrow range of the inter-vehicle distance in the evaluation scene. Therefore, the risk feeling function should be examined further by using the appropriate value of the individual driving characteristic or expanding the range of the inter-vehicle distance in the evaluation scene.

Fig. 12  Danger zone and the safety zone regression curves.

6. Conclusion

The purpose of this study was to extract the factors that define the individual differences in the risk perception of autonomous driving behavior, and to model the risk perception while taking into account the effects of the individual differences. We hypothesized that the prediction error, which is the difference between the driving behavior based on the individual sensation and the autonomous driving behavior, explained the individual differences. We experimentally manipulated the individual driving characteristic (i.e., prior prediction) and the actual autonomous driving behavior to control the prediction error and obtained subjective reports of feeling of risk from the participants for the varied prediction errors. As a result, we obtained the following findings:

- The risk was perceived based on the prediction error, i.e., the difference between the participants’ belief of the inter-vehicle distance in which the participants switched lanes to avoid the preceding vehicle and the actual inter-vehicle distance in the autonomous driving mode.
- The tendency towards anxiety, as a personality factor, affected the driver’s risk perception for the driving behavior of an autonomous vehicle.
- By learning the autonomous driving behavior regarding certain inter-vehicle distances, the risk feeling adapted to the learning conditions.
- There was an asymmetry in which the sensitivity of the perception of the degree of danger is higher than the safety. This indicates that the risk feeling function is different when facing danger in which a driver feels safer, and that it is necessary to apply an appropriate index formula according to the purpose and situation during autonomous driving control.
These findings will help quantify the individual differences in the risk perception by further examination and improvement. Nevertheless, the regression curve did not fit very well due to the limited experimental data. The model prediction and the experimental results indicate that the prediction error explained the individual differences. This result suggests that our model based on the expectation effect theory has potential to express the individual differences in the risk perception of autonomous driving behavior. Although the experiment was conducted with a limited number of participants, it shows a significant difference, and there should be larger surveys in the future.

Acknowledgement

We thank Prof. Tamotsu Murakami and the members of the Design Engineering Laboratory at the University of Tokyo for supporting this project. A part of this study was supported by KAKEN grant number 18H03318 from the Japan Society for the Promotion of Science.

References

Daniel, K. and Amos, T., Prospect Theory: An Analysis of Decision under Risk, Econometrica, Vol. 47, No. 2 (1979), pp. 263-292.
Ishibashi, M., Okuwa, M., Doi, S. I. and Akamatsu, M., Indices for characterizing driving style and their relevance to car following behavior, in SICE Annual Conference 2007, IEEE, pp. 1132–1137.
Kondoh, T., Hirose, T. and Furuyama, N., A proposal of tailor-made risk feeling formula when closing on the lead vehicle, Japan Ergonomics Society, Vol. 50, No. 6 (2004), pp. 350–358 (in Japanese).
Kondoh, T., Yamamura, T., Kitazaki, S., Kuge, N. and Boer, E. R., Identification of visual cues and quantification of drivers’ perception of proximity risk to the lead vehicle in car-following situations, Journal of Mechanical Systems for Transportation and Logistics, Vol. 1, No. 2 (2008), pp. 170–180.
McAllister, R., Gal, Y., Kendall, A., Van Der Wilk, M., Shah, A., Cipolla, R. and Weller, A., Concrete problems for autonomous vehicle safety: Advantages of Bayesian deep learning, International Joint Conferences on Artificial Intelligence, Inc., (2017), pp. 4745-4753.
Nakano, T. and Yanagisawa, H, Formulation of risk feeling for autonomous vehicle behaviour (Considering environment factors), Proceeding of the JSME Design & Systems Conference, (2018). (in Japanese), https://doi.org/10.1299/jsmedsd.2018.28.1203
National Police Agency Traffic Bureau: The situation of traffic accidents during 2017 (online), 2018, available from <https://www.npa.go.jp/publications/statistics/koutsuu/H29zennjiko.pdf>, (accessed on 28 January, 2020).
Rheinberger, C. M. and Hammitt, J. K., Dinner with Bayes: On the revision of risk beliefs, Journal of Risk and Uncertainty, Vol. 57 (2018), pp. 253–280.
Suganuma, N., Technical overview of autonomous driving system, Journal of the Japan Society of Applied Electromagnetics and Mechanics, Vol. 25, No. 4 (2017), pp. 373–378 (in Japanese).
Ueichi, H. and Kusumi, T., Effects of personality, cognitive, and situational variables on risk taking behavior, Shinrigaku kenkyu: The Japanese journal of psychology, Vol. 69, No. 2 (1998), pp. 81–88.
Yanagisawa, H., A computational model of perceptual expectation effect based on neural coding principles, Journal of Sensory Studies, Vol. 31, No. 5 (2016), pp. 430–439.
Yoon, S. and Kum, D., The multilayer perceptron approach to lateral motion prediction of surrounding vehicles for autonomous vehicles, IEEE Intelligent Vehicles Symposium (IV), IEEE (2016), pp. 1307-1312.