



A forward collision avoidance algorithm based on driver braking behavior

Xiaoxia Xiong^a, Meng Wang^b, Yingfeng Cai^{a,*}, Long Chen^a, Haneen Farah^b, Marjan Hagenzieker^b

^aSchool of Automotive and Traffic Engineering, Jiangsu University, 301 Xuefu Road, 212013, Zhenjiang, Jiangsu, China

^bDepartment of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands



ARTICLE INFO

Keywords:

Collision avoidance
Deceleration curve
Cluster analysis
Fuzzy logic
Driver braking behavior profile
Dynamic time warping

ABSTRACT

Measuring risk is critical for collision avoidance. The paper aims to develop an online risk level classification algorithm for forward collision avoidance systems. Assuming risk levels are reflected by braking profiles, deceleration curves from critical evasive braking events from the Virginia “100-car” database were first extracted. The curves are then clustered into different risk levels based on spectrum clustering, using curve distance and curve changing rate as dissimilarity metrics among deceleration curves. Fuzzy logic rules of safety indicators at critical braking onset for risk classification were then extracted according to the clustered risk levels. The safety indicators include time to collision, time headway, and final relative distance under emergency braking, which characterizes three kinds of uncertain critical conditions respectively. Finally, the obtained fuzzy risk level classification algorithm was tested and compared with other Automatic Emergency Braking (AEB) algorithms under Euro-NCAP testing scenarios in simulation. Results show the proposed algorithm is promising in balancing the objectives of avoiding collision and reducing interference with driver's normal driving compared with other algorithms.

1. Introduction

Traffic accidents cause approximately 1.25 million deaths per year worldwide and between 20–50 million non-fatal injuries with many resulting in disabilities according to the World Health Organization (WHO) ([World Health Organization, 2016](#)). To reduce traffic accidents, collision avoidance system (CAS) has received much attention and efforts from industry and research community over the past years ([Cabrera et al., 2012](#); [Wang et al., 2013](#); [Bränström et al., 2014](#)). CAS is one of the active safety technologies that could identify hazardous driving situations and actively assist drivers in avoiding or mitigating accidents (like Automatic Emergency Braking, AEB, as well as Automatic Emergency Steering, AES) ([Bränström et al., 2014](#)). To determine when the automatic overriding/intervention should be evoked (only automatic braking is considered in the current study), a robust real time risk level classification algorithm is needed, which serves as the core part of a CAS. However, to date risk level classification is still an open question in the literature due to the uncertainty and complexity of the interacting driver-vehicle-road-environment system ([Halim et al., 2016](#); [Mullakkal-Babu et al., 2017](#))

For online CAS implementation, risk associated with a driving situation (situational urgency/criticality) is generally measured using safety indicators based on the initial or predicted state of vehicle

interaction ([Mullakkal-Babu et al., 2017](#)). These safety indicators can be typically categorized into three main types: distance-based, time-based and deceleration-based indicators. Distance-based indicators usually calculate critical braking distance d_{br} based on vehicle kinematics and dynamics, and have been widely used in AEB algorithms ([Fujita et al., 1995](#); [Doi et al., 1994](#); [Seiler et al., 1998](#)). Time-based safety indicators have also been well researched thanks to their intuitive accordance with human natural judgment of situational urgency/criticality. Typical time-based safety indicators include Time To Collision (TTC), Time HeadWay (THW) or time gap, Time to Brake (TTB), and Time to Last-Second-Brake (TLSB) ([Hayward, 1972](#); [Goodrich et al., 1999](#); [Keller et al., 2011](#); [Zhang et al., 2006](#)). Deceleration-based indicator typically measures the required deceleration level a_{req} for the subject vehicle to successfully avoid a potential forward collision with its leading vehicle ([Hillenbrand et al., 2006](#); [Brannstrom et al., 2008](#)). However, these safety indicators are usually calculated under different assumptions (such as constant vehicle speed and/or acceleration/deceleration), which may lead to poor risk estimation performance under situations inconsistent with the assumption. For example, TTC does not reflect well the collision risk when the leading vehicle suddenly brakes (especially in small relative distance scenario), while the performance of THW suffers when the relative speed between the two vehicles is high (with subject vehicle speed being higher). Recently, a field-based

* Corresponding author.

E-mail address: caicaixiao0304@126.com (Y. Cai).

safety indicator was proposed by regarding driving risk as a spatial field (Wang et al., 2015a). However, the auxiliary parameters under this complex safety field assumption (such as driver physiological indices) can vary due to different road conditions and driving styles, making such field-based indicators difficult to calibrate and validate (Liu and Wu, 2009). For all these reasons, no clear/definite quantitative relationship between these safety indicators and risk levels has been established so far.

Another stream of literature attempts to measure situational urgency/criticality from driver's braking profiles. MacCall, et al. develops a Bayesian framework to evaluate situation criticality by jointly assessing the need for braking and the probability of performing braking by driver based on real-world driver braking behavior data (MacCall and Trivedi, 2007). Wang, et al. took average deceleration, maximum deceleration, and maximum kinetic energy difference during deceleration as feature variables and clustered observed deceleration profiles into different risk levels for risk causation analysis using *K-means* clustering technique (Wang et al., 2015b). Kluger et al. extracted feature variables from Discrete Fourier Transformed (DFT) longitudinal acceleration time series and used *K-means* clustering to identify critical events from naturalistic driving data (NDS) (Kluger et al., 2016). As these deceleration measures are actually consistent with the initial state measures in characterizing the risk of a driving situation (in other words, a more critical braking deceleration profile (like with higher deceleration level) generally indicates the initial state of the situation is at higher risk, assuming drivers have made evasive braking reaction to the situation, see Fig. 1), indices developed from naturalistic deceleration profiles have also been used for online driving risk level classification study.

Wada et al. obtained a series of proximity index K_{dbr} and relative distance data at brake initiation (of expert drivers using experimental car) and established a brake judgement line based on these two indicators using least squares method (Wada et al., 2010). Lee et al. used cumulative frequency diagram of acceleration/deceleration at a road section for measuring risk levels of deceleration profiles, and trained a neural network model using NGSIM trajectory data to predict potential severe deceleration for collision warning (Lee and Yeo, 2016). Moon et al. classified offline driving sequence samples into safe (deceleration $> -2 \text{ m/s}^2$), medium risk ($-2 \text{ m/s}^2 \geq \text{deceleration} > -4 \text{ m/s}^2$), and dangerous (deceleration $< -4 \text{ m/s}^2$) levels according to predefined thresholds and determined TTC and Warning Index thresholds of different hazard levels using confusion matrix method (Moon et al., 2009). However, the classified risk levels of the deceleration profiles in these studies were not validated by human inspectors (e.g., validating if a near-crash indeed happens or not by checking video footages in naturalistic driving study (Kluger et al., 2016)), meaning no true risk labels were available for the selected deceleration samples. This can question the validity of the classification results.

As discussed above, the quantification of safety indicator and risk

level relationship still has limitations in the literature, which can be further explored from real driving sequence data. The main objective of this paper is to develop an online driving risk level classification algorithm by optimal selection of safety indicators and classification rules according to naturalistic/offline deceleration profiles. The proposed online risk classification methodology is driven by data available in real time, as opposed to offline techniques which could employ the entire dataset (Shai et al., 1997). The research is conducted in two main steps as shown in Fig. 2. Firstly, offline deceleration profiles will be clustered into different risk levels, using braking deceleration profile to measure the criticality of vehicle interaction development process, as shown in Fig. 1. Based on the clustering results, effective representation of the identified risk levels will be explored using a set of online safety indicators (to avoid drawbacks that may exist for one single indicator) describing the initial state of the situation. In other words, vehicle deceleration is treated as an indirect measurement for real time driving risk in our proposed framework, where offline deceleration clusters provides basis for online safety indicator-based risk level classification rules. The performance of the proposed algorithm will be compared with state-of-the-art algorithms in simulation.

The remainder of the paper is structured as follows. The deceleration profiles for risk level identification are described in Section 2, followed by the deceleration profiles clustering analysis which is described in Section 3. The safety indicator-based risk level classification rule extraction based on risk clustering result is described in Section 4. Simulation test results of the proposed risk level classification rules are presented and discussed in Section 5. The final section, Section 6, summarizes the findings and concludes the paper.

2. Data assembly

Deceleration profile data from the Virginia Tech "100-car" Natural Driving Study (NDS) database (Virginia Tech Transportation Institute, 2014) were used for risk level classification. The database hosts driving sequence time series of 68 crashes and 760 near-crashes (where drivers took an emergency braking and/or evasive steering behavior). Vehicle motion parameters (such as speed, acceleration/deceleration, range from surrounding vehicle and range rate, etc.) values were recorded at 0.1 s interval from 30 s prior to the crash/near-crash till its end. Pre-crash (or pre-near-crash) "critical events" were manually coded for each driving sequence record. In the database, "critical events" are defined as "the state of environment or action that began the crash/near-crash sequence and made the crash/near-crash possible (which occurs outside the vehicle and does not include driver distraction, fatigue, or disciplining child while driving)" (Virginia Tech Transportation Institute, 2014). Such manually labeled critical deceleration profiles (and critical events) can establish a basis for more accurate risk level classification and thus were selected for the study.

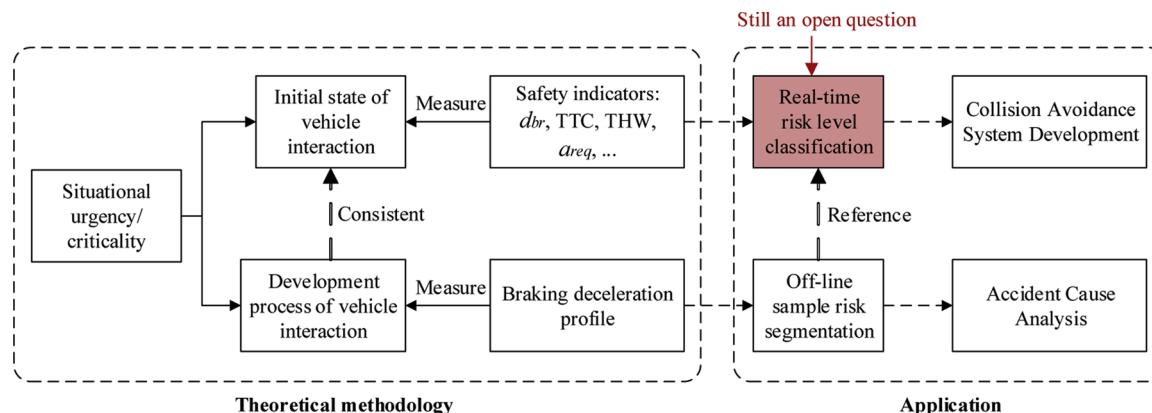


Fig. 1. Measurements of Situational Urgency/Criticality.

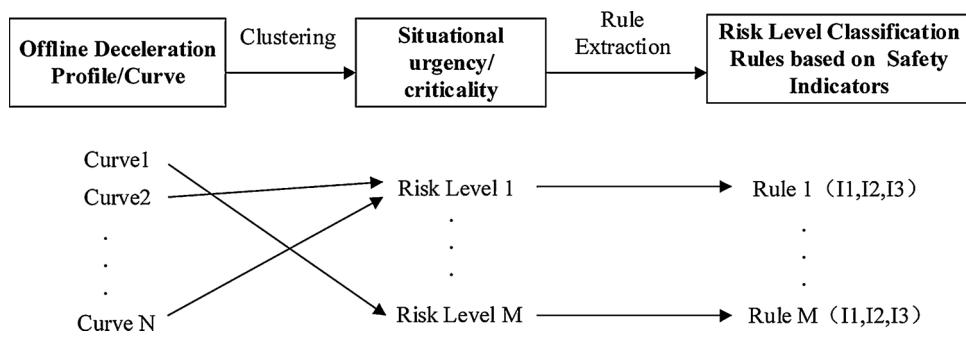


Fig. 2. Risk level classification methodology proposed in the paper.

2.1. Preliminary data selection

Based on the “100-car” database, specific deceleration profile samples were chosen following two main criteria as following:

1) Type of conflict/driver reaction

399 cases of rear-end conflict type were first chosen to extract sample deceleration profiles, as they fall within the scope of forward collision avoidance study. Among them, 385 near-crash cases were selected as they have complete deceleration curves (braking process) that may be interrupted by crash occurrence otherwise. Since the current study only focuses on collision avoidance timing by automatic braking, 121 cases of near-crashes avoided by steering or braking-and-steering were not considered. In addition, 86 cases of rear-end near crashes involving the subject vehicle (SV) changing lane or other vehicle cutting in were dropped from analysis, leading to 178 cases for analysis. The excluded events may feature different driver risk perception processes and result in different deceleration patterns, which may affect deceleration profile-based risk level classification results.

2) Data Validity

Cases without valid and complete vehicle speed, acceleration/deceleration, or range value (i.e., distance from the subject vehicle to leading vehicles), like negative or zero values recorded for missing value cases, were dropped as the analysis method used in this study requires such valid and complete data. Speed and range data were also manually checked via a visualized tool “Natware” (developed by Chalmers University of Technology for “100-car” database analysis) (Dozza, 2013) to validate driving data of the leading vehicle (i.e., principle other vehicle, POV). For example, the range rate between SV and POV should be negative when the subject driver starts to brake in reaction to the recorded “critical event”. A total of 127 rear-end near crashes (from 45 different drivers) were obtained following the two data selection criteria above.

Note that some braking may have already been applied prior to the start of the critical event (probably due to other non-critical event reasons like normal speed adjustment when a SV approaches an intersection/toll booth) and such braking onsets are irrelevant to the research topic (see Fig. 3, where the left red dashed line represents the start of a “critical event”). To maintain consistent analysis of the deceleration profiles, a notation of “critical event-reaction braking” was defined in the next section to identify the start of braking in reaction to the recorded “critical event”. Situational urgency/criticality (measured by online safety indicators) at the start of “critical event-reaction braking” is regarded as the initial state of vehicle interaction.

2.2. Critical braking profile extraction

Specifically, a “critical event-reaction braking” was defined as a sudden increase in brake pedal pressure resulting in a sharp drop in deceleration curve (i.e., the “elbow” point of the deceleration curve after the start of critical event in Fig. 3). To illustrate this concept, consider a typical car-following scenario (near-crash) as an example:

Phase I – subject driver starts normal braking (within $-1\text{--}2 \text{ m/s}^2$ deceleration level) as POV decelerated;

Phase II - subject driver brakes much harder ($> -3\text{--}4 \text{ m/s}^2$ deceleration level) to avoid rear-end striking as POV decelerated at a faster rate (even to a full stop) than the subject driver’s initial expectation.

In this case example, the moment when the driver started braking could be regarded as “normal event-reaction braking”, with the normal event being POV decelerating at a driver-expected moderate rate; in contrast, the moment when the driver braked much harder refers to the start of “critical event-reaction braking”, with the critical event being POV decelerating at a fast rate. It could be noted that the time interval between the start of normal event-reaction braking and critical event-reaction braking could vary (may even overlap) from case to case depending on the driver’s perceived risk level of events (commonly known as hazard perception) (Crick and McKenna, 1992; Haworth et al., 2005; Evans and Macdonald, 2002). As such, it could be believed that the start of “critical event-reaction braking” actually represents the moment when the driver deems the current deceleration level to be unsafe and takes evasive reaction.

Accordingly, in developing a collision avoidance algorithm, to avoid potential driver aversion, the timing of a normal event-reaction braking is expected to be determined by a driver, that is no warning/overriding is needed at this stage; whereas a critical event-reaction braking indicates an unsafe situation and an automatic emergency braking is expected to be implemented by CAS when a critical event-reaction braking of highest risk level is predicted (under the assumption that critical event-reaction braking of highest risk level in a near-crash represents the final “successful” last-second braking, which would be discussed in more details in Section 4 and validated via simulation in Section 5). Thus it is important to distinguish normal vs. critical event-reaction braking for collision avoidance algorithm development.

Based on the “critical event-reaction braking” notation above, the start of critical event-reaction braking were manually identified and checked via Natware. Deceleration profiles with no clear sign of evasive reaction to the critical event (i.e., with no clear “elbow” point in the deceleration curve after critical event started, as shown in Fig. 3) were excluded from analysis to avoid biasness in determining the starting point of “critical event-reaction braking” (a total of 30 such cases were excluded from the 127 rear-end near crashes sample cases obtained above). Finally, deceleration time series for each chosen sample record were extracted from the start of “critical event-reaction braking” to the moment when deceleration value reaching its maximum, yielding a total of 97 such deceleration time series (from 43 different drivers) as the final deceleration profile sample space (for situational urgency/criticality assessment). The statistical measures of the deceleration profile sample are presented in Fig. 4. It could be seen that the initial deceleration of some critical event-reaction braking sequences are below zero, indicating these sample drivers had already started normal event-reaction braking prior to critical event-reaction braking; also, the maximum deceleration in the sample could reach -10 m/s^2 level (close

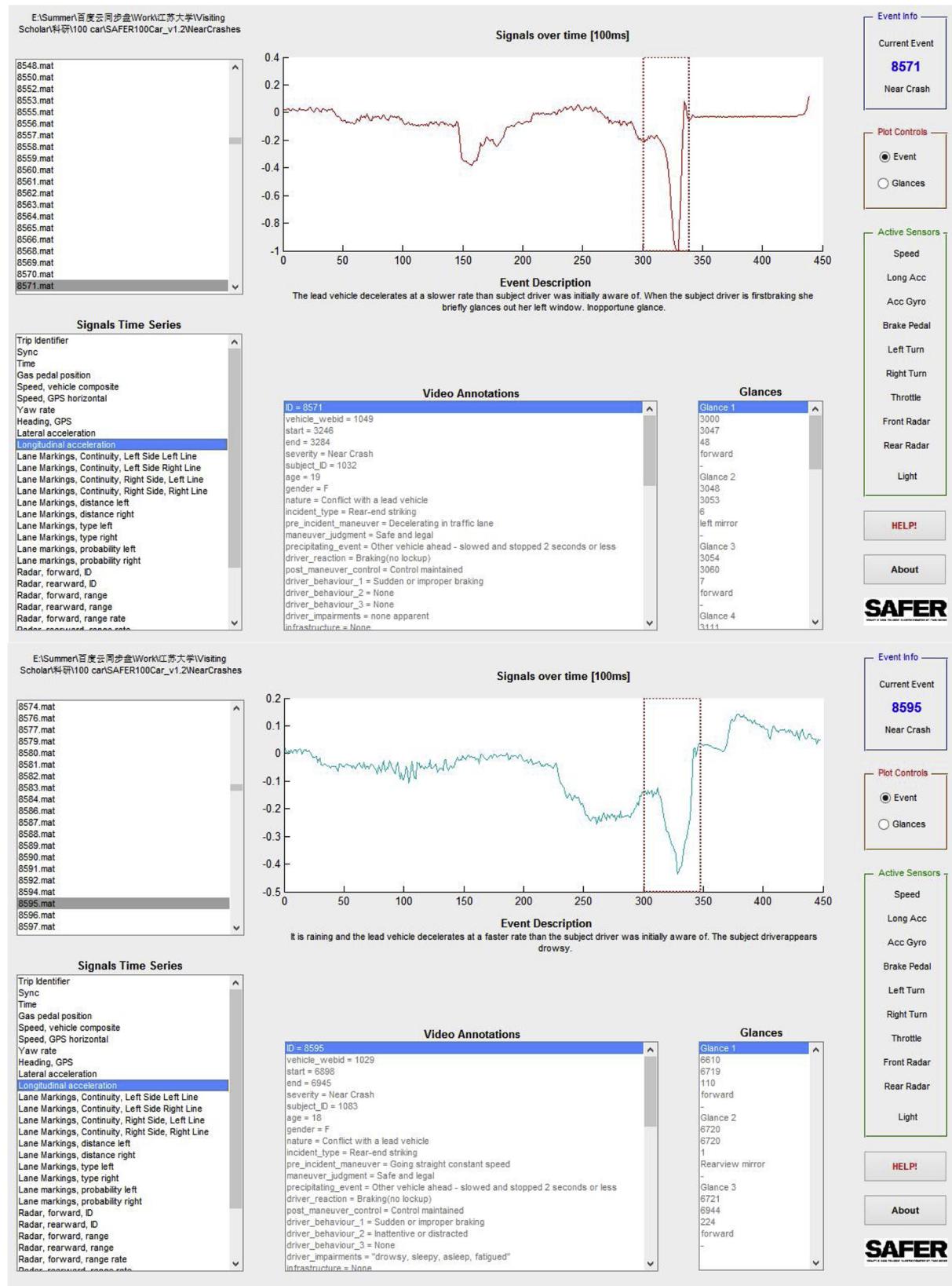


Fig. 3. Deceleration curves of sample #8571 and #8595 (deceleration started before critical event started) extracted from Natware.

to the limit of vehicle braking capability), indicating some of these extracted deceleration profiles actually reflect the driving situations of highest risk level.

It should be noted that at this study stage, the start of critical event-

reaction braking for each sample record was identified manually by checking visualized deceleration curves based on Natware. As results show that the extracted deceleration time series features a continuous decrease in deceleration value with the magnitude over -4 m/s^2 (see

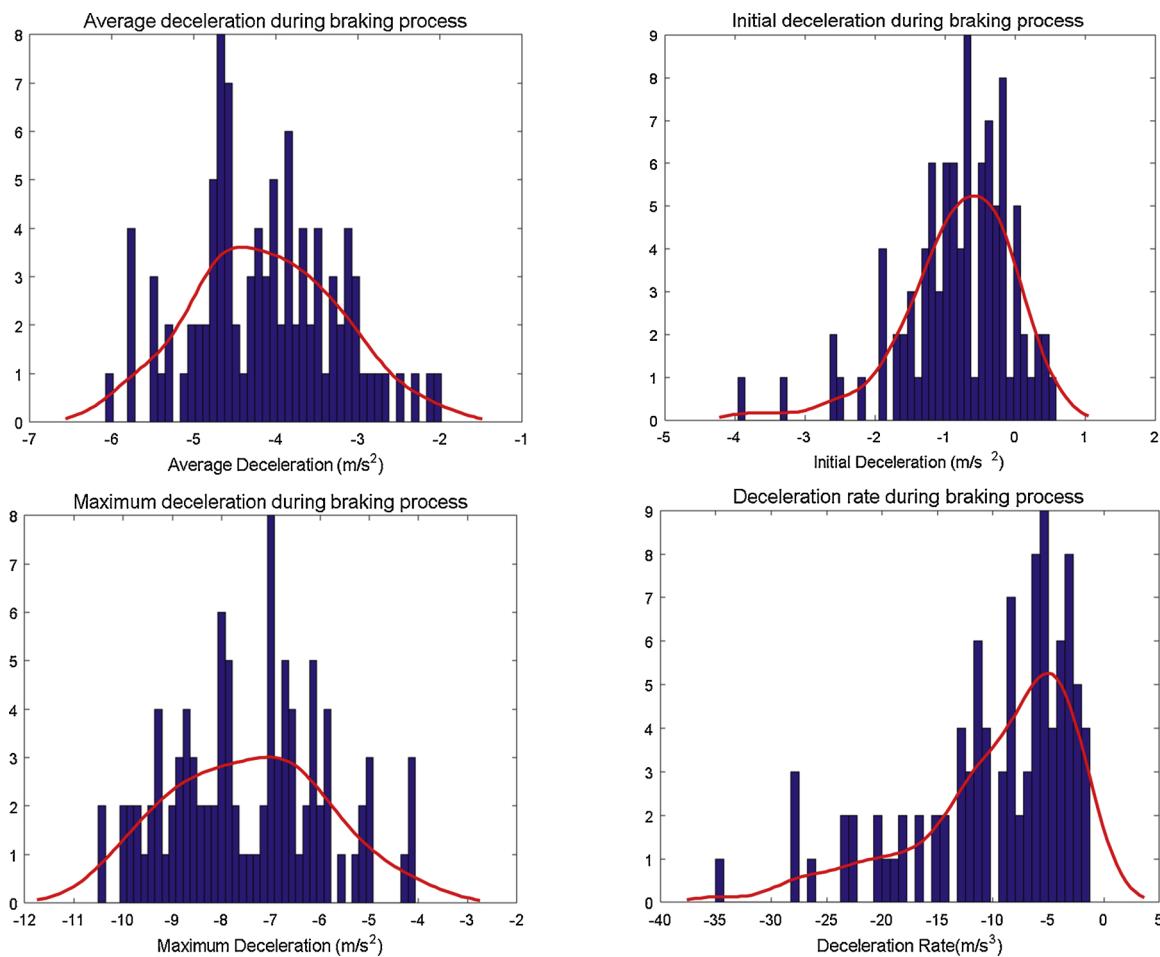


Fig. 4. Statistics of deceleration profiles extracted from “100-car” near-crashes.

Fig. 4, generally consistent with the large deceleration threshold defined in literature (Moon et al., 2009), which happens “when the driver really need to apply severe braking to avoid potential collision”), such threshold could be considered as a screening criteria in identifying starting point of “critical event-reaction braking” in a future larger scale study. Also, at this stage, the identification procedure does not distinguish traffic/driver/vehicle characteristics (such as vehicle type, driver state, and vehicle response during braking, etc.) due to limited sample space. More exploration could be made in the future to differentiate the possible effects of these characteristics/variables on critical braking event identification.

3. Deceleration profile clustering

As discussed in **Fig. 2**, deceleration profiles extracted in Section 2 were clustered into different risk levels first for later safety indicator-based risk level classification rule study. Since the traditional clustering technique of *K-means* cannot directly average time series when updating cluster centroids in the iterative calculation, feature variables of the time series (such as mean/maximum statistics) are usually required for *K-means* clustering (Liao, 2005). However, such feature extraction process may lose some effective information within the complete sequence of time series, which affects the correctness of the clustering results. *Spectral* clustering, as another widely used clustering technique fits into the scope of the study (clustering of deceleration profile/curves), and it only requires similarity information between data points (rather than their centroids) (Luxburg, 2007). Thus spectral clustering was selected for deceleration time series clustering analysis.

3.1. Spectral clustering methodology

Spectral clustering is an algorithm based on graph theory (Luxburg, 2007) and has been widely used in many fields (Bach and Jordan, 2006; Chang and Yeung, 2005; Valgren et al., 2007). *Spectral* clustering treats all data as points in space, and assumes all of these data points are connected by edges assigned with weights (i.e., a connected graph). The edge weight is higher when the similarity between the two points is larger, and vice versa. The edge weights between all n data points constitute the adjacency matrix $W_{n \times n}$ of the graph, with the element in the i -th row and j -th column recording the weight value between point i and point j . The purpose of *Spectral* clustering is to find a partition of the graph such that the weight sum of edges between different groups is small while that of edges within a group is high (Luxburg, 2007).

For the deceleration profile clustering problem, as a higher overall level of deceleration (the deceleration curve being farther away from the deceleration = 0 straight line) and a faster rate of change of deceleration curve generally indicate a higher level of risk perceived by the driver at the start of critical event-reaction braking, dissimilarity of deceleration curves (i.e., edge weight) can be measured in two respects including range/distance metric (dtw) and difference in rate of curve change metric ($\Delta jerk$), which are explained below.

1) Range distance metric (dtw)

As the duration of critical event-reaction braking from its start to reaching its maximum varies from case to case, Dynamic Time Warping (DTW) algorithm was employed to calculate the range distances of the extracted deceleration time series. DTW distance of two time series is calculated by identifying a distance path (repeating/deleting certain points within the time series, i.e., shortening/extending the sequence of

time series) where the sum of pairwise point distances along the path is minimized. Take two K-dimensional signal time series \mathbf{X} and \mathbf{Y} as an example (whereas the deceleration profile in our case is one-dimensional signal time series, i.e., $K = 1$):

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,M} \\ x_{2,1} & x_{2,2} & \dots & x_{2,M} \\ \dots & \dots & \dots & \dots \\ x_{K,1} & x_{K,2} & \dots & x_{K,M} \end{bmatrix} \quad (1)$$

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & y_{1,2} & \dots & y_{1,N} \\ y_{2,1} & y_{2,2} & \dots & y_{2,N} \\ \dots & \dots & \dots & \dots \\ y_{K,1} & y_{K,2} & \dots & y_{K,N} \end{bmatrix} \quad (2)$$

That is, the two time series \mathbf{X} and \mathbf{Y} have M and N sampling points (sample points) respectively. Define the distance between the m th sampling point of \mathbf{X} (i.e., the m th column of \mathbf{X} , $m = 1, 2, \dots, M$) and the n th sampling point of \mathbf{Y} (i.e., the n th column of \mathbf{Y} , $n = 1, 2, \dots, N$) as a Euclidean distance $d_{mn}(\mathbf{X}, \mathbf{Y})$:

$$d_{mn}(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{k=1}^K (x_{k,m} - y_{k,n})^2} \quad (3)$$

The DTW algorithm looks for two path sequences \mathbf{ix} and \mathbf{iy} (of the same length) in the distance lattice d_{mn} , $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$ (as shown in Fig. 5), which minimizes the sum of the distances between the sampling points along the distance path $dist$ (formed by \mathbf{ix} and \mathbf{iy} , i.e., the gray polyline in Fig. 5 (Matlab 2018R, 2018)).

$$\mathbf{ix}^*, \mathbf{iy}^* = \arg \min_{\mathbf{ix}, \mathbf{iy}} \sum_{m \in \mathbf{ix}, n \in \mathbf{iy}} d_{mn}(\mathbf{X}, \mathbf{Y}) \quad (4)$$

The distance path $dist$ is required to start from $d_{11}(\mathbf{X}, \mathbf{Y})$ and end at $d_{MN}(\mathbf{X}, \mathbf{Y})$, which proceeds following the rules below (Fig. 5-b):

- ① lateral movement: $(m, n) \rightarrow (m + 1, n)$
- ② longitudinal movement: $(m, n) \rightarrow (m, n + 1)$
- ③ diagonal movement: $(m, n) \rightarrow (m + 1, n + 1)$

Finally, the DTW distance between the two time series \mathbf{X} and \mathbf{Y} is:

$$DTW(D) = \sum_{m \in \mathbf{ix}^*, n \in \mathbf{iy}^*} d_{mn}(\mathbf{X}, \mathbf{Y}) \quad (5)$$

More details of DTW calculation could be found in (Luxburg, 2007). Such DTW distances do not require feature variables extracted from time series (like their mean, maximum, and minimum, etc.) and could retain the complete deceleration level information within the whole sequence of time series.

2) Difference in rate of change metric ($\Delta jerk$)

As discussed above, DTW algorithm evaluates the range/shape similarity of the deceleration curves by stretching or compressing their segments, but such transformation operation can lose the change rate information of the deceleration curves. However, such change rate information could also be an important indicator for risk level

assessment. For example, for two deceleration profiles with maximum deceleration being the same, one would usually expect the profile decelerating at a higher rate is of higher risk level. Thus, the average deceleration change rate over the extracted deceleration profile, usually referred to as jerk (the ratio of increase in deceleration magnitude to time duration, or average ramp-up), was also employed as another dissimilarity metric in addition to DTW distance.

Based on the range distance metric dtw (in m/s^2) and difference in rate of change metric $\Delta jerk$ (in m/s^3) defined above, adjacency matrix $W_{n \times n} = \{w_{ij}\}$, $i, j = 1, \dots, n$, an $n \times n$ matrix recording pairwise similarities between the n data points (one data point refers to one deceleration curve/time series), can be defined using Gauss Radial basis function kernel (a typical similarity representation metric) (Schölkopf et al., 2004) as below:

$$w_{ij} = \exp(-\text{gamma} \cdot \|x_i - x_j\|) = \exp\left(-\frac{1}{n_m} \cdot \sqrt{(dtw_{ij})^2 + (\Delta jerk_{ij})^2}\right) \quad (6)$$

where w_{ij} measures the similarity between data point x_i and x_j , gamma is dimension parameter defined as $\frac{1}{n_m}$, where n_m represents the number of employed dissimilarity metrics and was set to two in our case (including dtw and $\Delta jerk$). Dissimilarity metrics dtw and $\Delta jerk$ were normalized to 0–1 range in the study based on minimum–maximum bounds to eliminate possible dimension/scaling effects on similarity measurement.

3.2. Clustering results

Spectral clustering was performed for the 97 deceleration profiles extracted in Section 2 using Matlab (Matlab 2018R, 2018). In addition to *Spectral* clustering, *K-means* clustering based on features including average deceleration, maximum deceleration, and jerk was also performed as baseline for comparison. Considering the limited sample size, $K = 2$ clusters were selected for risk level identification, with all clustering results presented in Table 1. Results show that Cluster-1 by *Spectral* clustering includes all the 21 cases assigned to that by *K-means*. By including another 14 cases into Cluster-1, *Spectral* clustering yields higher average maximum deceleration and average mean deceleration but relatively lower mean jerk level. This means compared with *K-means* clustering results, cases with $decel_max > -8.498$, $decel_mean > -3.822$, and $-6.597 < jerk_mean < -21.186$ are transferred from Cluster-2 to Cluster-1 by *Spectral* clustering. As one would expect a braking process that features higher average and maximum deceleration but with a relatively lower jerk is more dangerous than that which has lower average and maximum deceleration with a higher jerk, the *Spectral* clustering results are more reasonable than *K-means*. The *Spectral* clustering technique can better identify the high magnitude and fast changing rate pattern in deceleration profiles compared to *K-means*.

As Cluster-1 generally features higher maximum and mean deceleration level as well as faster average deceleration rate compared to Cluster-2, deceleration profiles in Cluster-1 were categorized into high

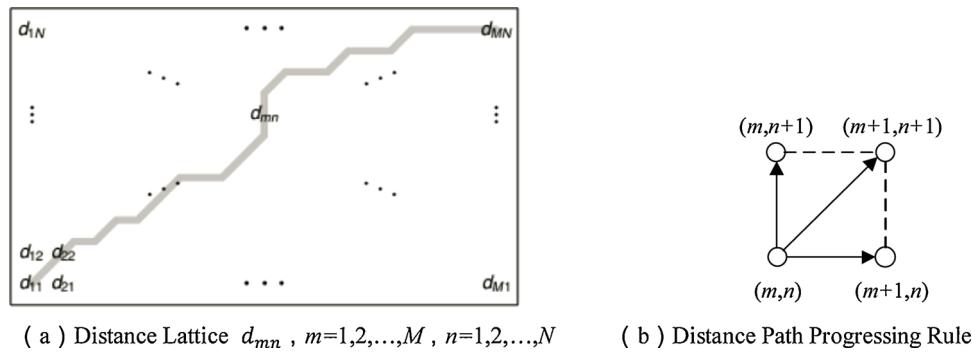


Fig. 5. Diagram of DTW distance calculation (Matlab 2018R, 2018).

Table 1

Spectral vs. K-means clustering results of critical braking deceleration profiles.

		decel_max (m/s ²)	decel_mean(m/s ²)	jerk_mean(m/s ³)
Spectral clustering	Cluster-1(35 cases)	-8.672	-3.925	-17.125
	Cluster-2 (62 cases)	-6.771	-2.982	-5.594
K-means clustering	Cluster-1(21 cases)	-8.498	-3.822	-21.186
	Cluster-2(76cases)	-7.169	-3.185	-6.597

risk level (i.e., their corresponding initial states of vehicle interaction at the start of critical event-reaction braking were also classified to high risk), while those in Cluster-2 were categorized into sub-high/medium risk level (as all sample deceleration profiles were extracted from “100-car” near-crash records, the extracted profiles should not be assigned to low risk).

4. Risk level rule extraction

In this section, given the identified risk levels of the deceleration profiles in Section 3, online risk level classification rules are explored using safety indicators at the onset of critical event-reaction braking.

4.1. Safety indicator selection

As discussed in Section 1, different drawbacks may exist for one single safety indicator due to their assumption limitations. To overcome such assumption limitations, the paper considers three typical critical driving scenarios (based on possible/uncertain combinations of driving conditions of SV and POV that could lead to critical scenarios) and selects three safety indicators to characterize each scenario respectively (summarized in Table 2). Risk level classification rules would then be explored based on the joint set of these safety indicators in the next section.

As shown in Table 2, the typical safety indicator that matches with the assumption of Scenario 1 is TTC, which characterizes the remaining time for SV and POV to collide if they continue driving at current speed along current trajectory (with SV at higher speed than POV); similarly, the typical safety indicator selected for Scenario-2 is THW, which characterizes the remaining time for the two vehicles to collide if POV suddenly stops (an extreme condition) while SV continues driving at current speed along current trajectory; finally, potential indicator of collision with urgent deceleration (PICUD) (Beinum et al., 2016) was selected for Scenario-3, which characterizes the final distance between SV and POV if both of them decelerate under maximum braking force. Fig. 6 shows the pattern of the three safety indicators over the whole driving scenario space (each relative distance-relative speed combination set represents a specific driving scenario), assuming $v_n = 60 \text{ km/h}$ (16.67 m/s), $a_{min} = -8 \text{ m/s}^2$, $t_h = 1 \text{ s}$ (Mullakkal-Babu et al., 2017) (where color changing from blue to red indicates risk level evolving from lower to higher; iTTC is used instead of TTC here for better visualization, as the range of TTC could be very large). It can be seen that

TTC (iTTC), THW, and PICUD show different changing patterns in risk level (featuring radial, strip-shaped, and arc-shaped patterns respectively) when the relative distance between SV and POV increases, indicating the combination of the three safety indicators could capture a linear-nonlinear integrated relationship between risk level and range/relative distance criteria.

4.2. Classification rule extraction based on safety indicators

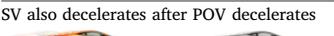
Based on the risk level clustering results in Section 3, statistical measures of the selected safety indicators at critical event-reaction braking onsets (by different risk level clusters) are presented in Fig. 7. It could be seen that at the start of critical event-reaction braking, the distribution of TTC, THW, and PICUD are centered near 2.0 s (iTTC = 0.5 s), 0.7 s, and -6 m respectively, which is generally consistent with the medium to high risk thresholds defined for these indicators in the literature (Mullakkal-Babu et al., 2017; Beinum et al., 2016; Li et al., 2017; Lin et al., 2012), indicating the deceleration profiles extracted in the paper are effective in characterizing driving situations of higher risk. Also, it is noted that Cluster-1 generally features smaller TTC, THW and PICUD than Cluster-2, which validates the conclusion that Cluster-1 represents high risk level while Cluster-2 represents medium risk level based on deceleration profile characteristics (Section 3.2).

The joint distribution of the safety indicator set {TTC, THW, PICUD} at critical event-reaction braking onset is shown in Fig. 8 (note that also for better visualization iTTC is displayed here instead of TTC). It could be seen that no clear thresholds of TTC, THW, or PICUD could be identified to distinguish different risk clusters and more sophisticated classification rules should be explored. As driving risk level (typically categorized into high, medium and low risk levels) is actually a fuzzy concept that could be reasoned from safety indicators (Vicente et al., 2012), fuzzy logic fits into the scope well and was selected in the paper. Fuzzy logic generally has good classification performance in the presence of complex/uncertain situations (like the uncertain critical scenarios discussed in Section 4.1, i.e., the three critical driving scenarios) (George and Bo, 1994), and was employed here to extract risk level classification rules using the safety indicator set {TTC, THW, PICUD}.

As the rule extraction problem could be regarded as a classification performance optimization problem (which is usually to improve true positives [higher risk correctly predicted as higher level] while reducing false positives [lower risk wrongly predicted as higher level] for risk level classification), two confusion matrices were defined for

Table 2

Three critical driving scenarios and corresponding safety indicators.

NO	Critical Driving Scenario	Safety Indicator
1	SV and POV continue driving at current state, with SV at higher speed than POV 	$\text{TTC} = \frac{s_n}{v_n - v_{n-1}}$
2	POV decelerates while SV continues driving at current state 	$\text{THW} = \frac{s_n}{v_n}$
3	SV also decelerates after POV decelerates 	$\text{PICUD} = s_n + \frac{v_n^2 - v_{n-1}^2}{2 \cdot a_{min} } - t_h v_n$

Note: v_n and v_{n-1} represent the current driving speed of SV and POV respectively, with v'_n and v'_{n-1} representing their future speed in the critical scenario; a_n and a_{n-1} represent the deceleration of SV and POV in the critical scenario respectively, with $a_{n,min}$ and $a_{n-1,min}$ representing their maximum value under maximum braking force (assuming $a_{n,min} = a_{n-1,min} = a_{min}$ for simplicity); s_n represents the current distance between SV and POV; t_h represents the driver reaction time.

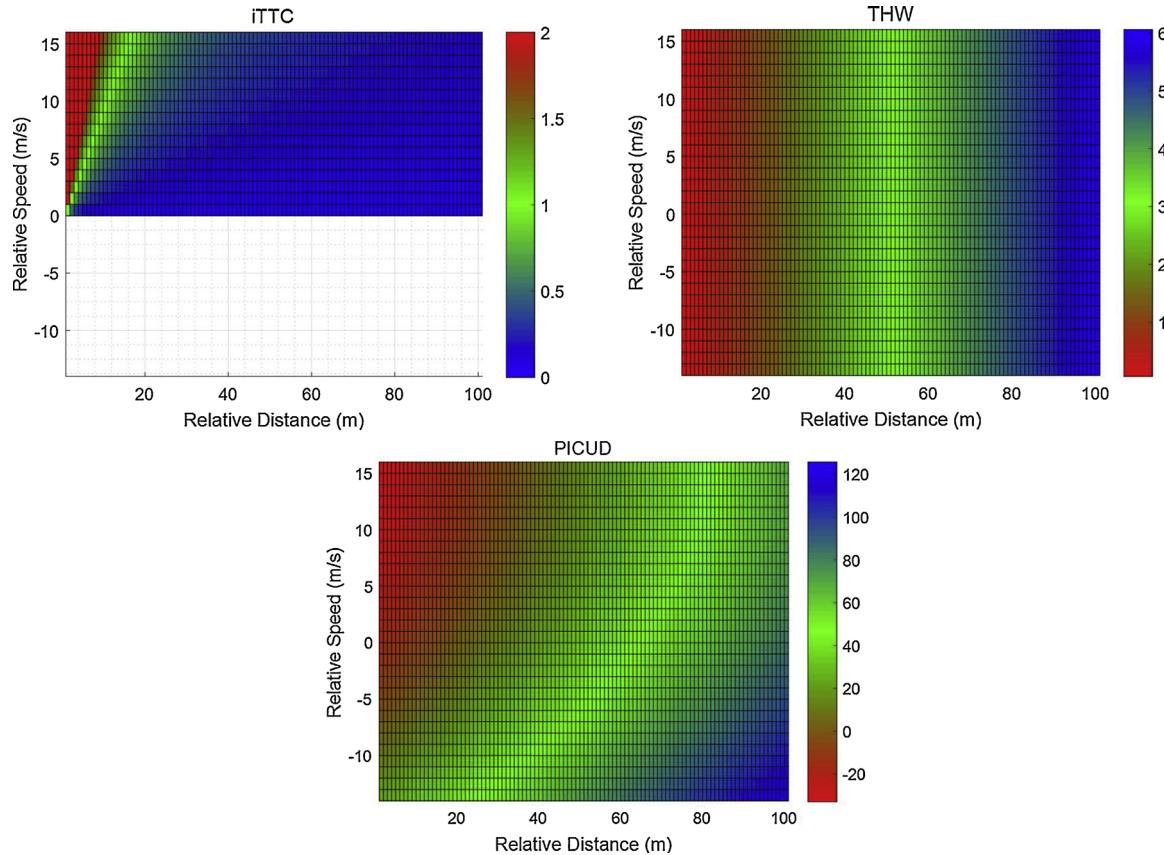


Fig. 6. Distribution of iTTC (1/s), THW (s) and PICUD (m) in relative distance-relative velocity plane (Mullakkal-Babu et al., 2017).

classification performance evaluation (note that risk level includes three categories in the paper, which could lead to two combinations of higher vs. lower risk level notations and result in two matrices) as shown in Table 3.

As shown in Table 3, for each deceleration profile sample, the observed risk level refers to the assigned risk level cluster based on deceleration curves in Section 3; whereas the classified risk level refers to the output of fuzzy logic by feeding observed {TTC, THW, PICUD} at critical event-reaction braking onset, where the output was classified into low, medium, and high risk levels in accordance with the risk level clustering results. Classified samples were labeled as True Negative (TN), False Negative (FN), False Positive (FP), and True Positive (TP) (McNicol, 2004). It should be noted that though no low risk level samples are actually observed in the extracted deceleration profile sample space, low risk level was still defined for fuzzy logic rule output and was utilized for classification performance evaluation. For example, if any profile sample is classified into low risk level by fuzzy logic, then its classification output would be labeled as false negative (FN) according to the confusion matrices in Table 3, which would directly influence the results of classification performance evaluation. Specifically, classification performance index PI was defined using geometric mean of True Positive Rate (TPR) and Precision Rate (Pr) of the two matrices (geometric mean is usually employed when multiplication of ratios, i.e., TPR and Pr , are included in the formula), where both TPR and Pr were selected to balance the objectives of improving threatening situation detection and avoiding false alarms.

$$PI = \sqrt{\sqrt{TPR_1 \cdot Pr_1} \cdot \sqrt{TPR_2 \cdot Pr_2}} \quad (7)$$

$$TPR_i = \frac{TP_i}{FN_i + TP_i} \quad (8)$$

$$Pr_i = \frac{TP_i}{FP_i + TP_i} \quad (9)$$

where TP_i , FN_i , FP_i ($i = 1, 2$) represent the number of classified samples labeled with TP_i , FN_i , FP_i respectively (based on the confusion matrices defined in Table 3).

A fuzzy logic structure typically includes input/output membership functions and fuzzy output rules (with more technical details in (George and Bo, 1994)). Different fuzzy logic structures for risk level classification based on safety indicator set {TTC, THW, PICUD} were examined and the structure defined in Table 4 to Table 6 shows best classification performance by trials.

Z-shaped, S-shaped, and Triangular-shaped membership functions in Table 4–5 were defined following the Matlab built-in formats (Matlab 2018R, 2018). Mean of maximum (MOM) defuzzification method was employed to generate the final output of the fuzzy logic (a continuous value within 0–1 interval), with output value in range [0, 0.25] classified as low risk level, (0.25, 0.75] classified as medium risk level, and (0.75, 1] classified as high risk level. Note that gap1, gap2, and gap3 feature the membership segmentations in TTC, THW, and PICUD respectively (see Table 4). Take gap1 as an example, gap1 could be considered as the time difference between critical (more dangerous) state and soft (safer) state in TTC dimension. Under such definition, when TTC is observed to be lower than TTC_{C1} , the corresponding input variable TTC would be labeled as “Critical” at 100% probability; and when TTC is observed to be larger than $TTC_{C1} + gap1$, a “Soft” label would be assigned at 100% probability; and finally when TTC is observed between the gap, “Critical” and “Soft” label would be assigned with the probabilities following their respective membership functions (the total of the two probabilities is 1). Following the same framework, gap2 and gap3 could be considered as the time and distance difference between critical and soft state in THW and PICUD dimension respectively.

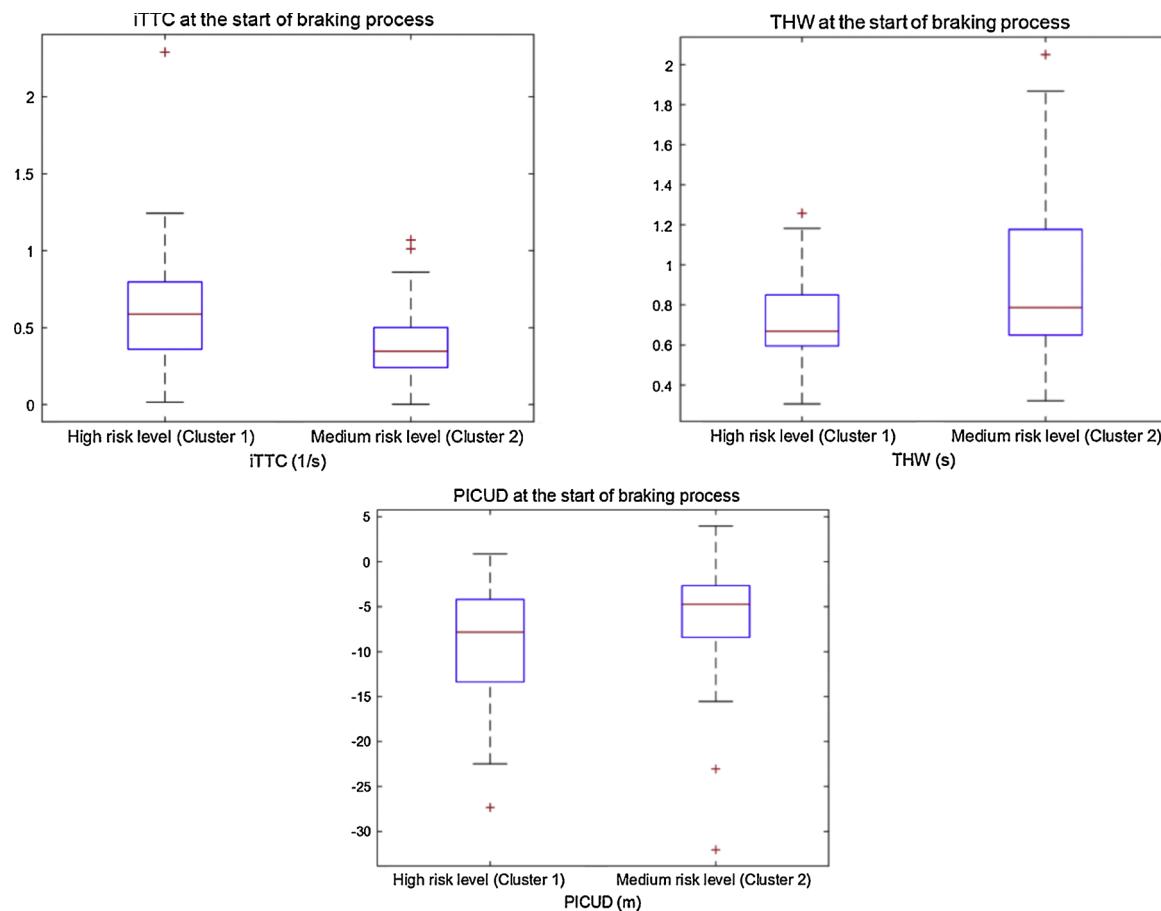


Fig. 7. Statistics of TTC, THW, PICUD at the start of critical event-reaction braking.

As discussed in Eqn. 7, PI is the target value (performance indicator) that needs to be optimized given the observed samples ($\{\text{TTC}, \text{THW}, \text{PICUD}\}$ - risk level cluster observation pairs), while different input parameter set $x = \{\text{TTC}_1, \text{gap}_1, \text{THW}_1, \text{gap}_2, \text{PICUD}_1, \text{gap}_3\}$ (Table 4) could lead to different PI values. Thus the tuning of these parameters is in fact a parameter optimization problem that could be solved by Particle Swarm Optimization (PSO) algorithm, which optimizes a problem by moving a population of candidate solutions (i.e., particles) towards the best known position in the search-space (with no requirement in differentiability of the optimization problem) (Kennedy and Eberhart, 2002). However, instead of using PI as the fitness function directly, $-PI^4 = -TPR_1 \cdot Pr_1 \cdot TPR_2 \cdot Pr_2$ (the same denotation with Eqn. 7) is utilized to speed up calculation considering the complexity in the function form of PI (specifically, the fourth power could counteract the double rooting operation within PI that could affect the performance of PSO algorithm; and adding a negative sign is a routine operation in forming fitness function as the PSO algorithm always finds minimum). The lower/upper bounds of the parameter set x (search-space) were defined as $lb = [0.5, 1, 0.5, 1, -35, 0]$ and $ub = [2, 6, 1, 3, 0, 20]$ according to the distribution of TTC, THW, and PICUD at the start of critical event-reaction braking (see Fig. 7), and with the swarm population size set at 60. Finally, the optimal solution of parameter set x was obtained by PSO algorithm based on Matlab as shown in Table 7 (with PSO iteration process shown in Fig. 9), where the optimal value of fitness function is $-PI^4 = -0.588$, which is equivalent to performance indicator $PI = 0.876$.

According to the optimal fuzzy parameters obtained above, the output surfaces of the fuzzy logic with PICUD fixed at $a = -14$ m and $b = -8$ m (see parameters defined in Table 4) are shown in Fig. 10. It could be seen that the TTC-THW output surface shows an overall downward trend when the fixed value of PICUD increases, which is in

line with one would expect (the overall risk level is expected to be lower as PICUD increases, with TTC and THW remaining the same). The result also indicates that using only one single safety indicator (i.e., with only one critical scenario considered) would be difficult to fully reflect the linear/non-linear characteristics in risk level over the driving space. Adopting such fuzzy logic-based rules could be necessary in improving the accuracy of risk level classification by integrating safety indicators featuring different uncertain critical driving scenarios.

From the previous analysis, two main assumptions have been made in establishing the algorithm:

Ass1: A more critical braking deceleration profile indicates the initial state of the driving situation at higher risk level (Section 1);

Ass2: Critical event-reaction braking of highest risk level in a near-crash represents the final “successful” last-second braking by driver (Section 2);

Specifically, Ass1 is the theoretical basis for fuzzy logic-based risk level classification, where deceleration curve (profile) clusters represent different risk levels at the start of the recorded braking. Ass2 corresponds to the timing of braking control in CAS: in accordance with the final “successful” last-second braking timing perceived by driver, when the output of the fuzzy logic at time t (by feeding observed safety indicator set $\{\text{TTC}_t, \text{THW}_t, \text{PICUD}_t\}$ at time t) is high risk level (corresponding to the moment when the subject driver is expected to take a critical event-reaction braking of highest risk level), an automatic emergency braking should be implemented immediately by CAS. In other words, in the proposed forward collision avoidance algorithm, braking control would be initiated when the current driving situation (represented by $\{\text{TTC}, \text{THW}, \text{PICUD}\}$) is classified as high risk level by fuzzy logic. AEB tests would be employed to validate the proposed fuzzy logic-based braking control algorithm in the next section.

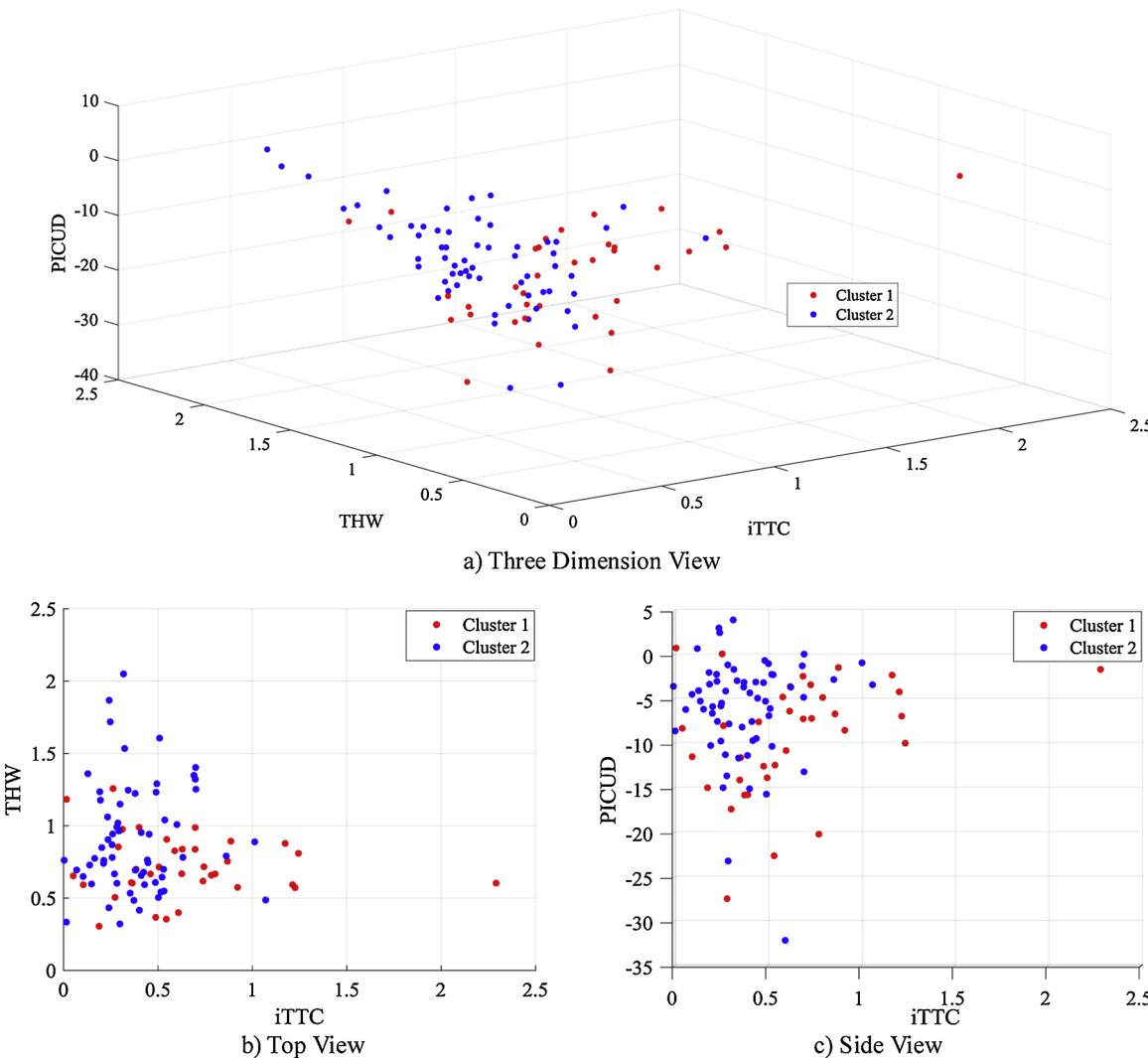


Fig. 8. Cluster distribution of TTC, THW, PICUD at the start of critical event-reaction braking.

Table 3
Classification rule output confusion matrices.

a) Confusion matrix-1

Classified Risk Level	Observed Risk Level		
	Low	Med + High	Med + High
Low	TN_1	FP_1	FN_1
Med + High			TP_1

b) Confusion matrix -2

Classified Risk Level	Observed Risk Level		
	Low + Med	High	High
Low + Med	TN_2	FP_2	FN_2
High			TP_2

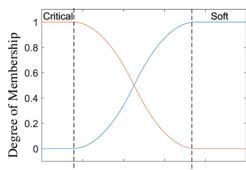
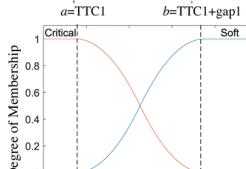
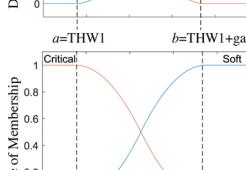
5. Simulation test and results

The AEB test scenarios proposed by European New Car Assessment Programme (Euro-NCAP) (Euro-NCAP, 2013) were selected to validate the proposed collision avoidance algorithm. Test scenarios could be categorized into three main types according to the movement status of POV, including Car-to-Car Rear Standing (CCRs), Car-to-Car Rear Moving (CCRm), and Car-to-Car Rear braking (CCRb) as shown in

Fig. 11 (with SV shown in red and POV shown in white).

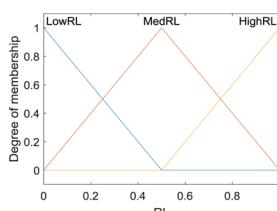
In order to evaluate the collision avoidance algorithm proposed in the paper, classic AEB algorithms including Berkley, Mazda, Honda and SeungwukMoon (Fujita et al., 1995; Doi et al., 1994; Seiler et al., 1998; Moon and Yi, 2008) were also tested for comparison. All tests were implemented via the Prescan Simulation Platform (Prescan, 2018), with vehicle maximum deceleration set at 0.8 g (to account for possible unfavorable driving situations, where generally a maximum

Table 4
Input variable membership function.

Input Variable	Label	Membership Function (Slope changes at a and b)	Graph of Membership Function
TTC	Critical	Z-shaped form $\begin{cases} a = \text{TTC1} \\ b = \text{TTC1} + \text{gap1} \end{cases}$	
	Soft	S-shaped form $\begin{cases} a = \text{TTC1} \\ b = \text{TTC1} + \text{gap1} \end{cases}$	
THW	Critical	Z-shaped form $\begin{cases} a = \text{THW1} \\ b = \text{THW1} + \text{gap2} \end{cases}$	
	Soft	S-shaped form $\begin{cases} a = \text{THW1} \\ b = \text{THW1} + \text{gap2} \end{cases}$	
PICUD	Critical	Z-shaped form $\begin{cases} a = \text{PICUD1} \\ b = \text{PICUD1} + \text{gap3} \end{cases}$	
	Soft	S-shaped form $\begin{cases} a = \text{PICTUD1} \\ b = \text{PICTUD1} + \text{gap3} \end{cases}$	

Note: Label “Critical” represents a more dangerous state, while “Soft” represents a safer state; the x -axis of the membership function graphs represent the input variable TTC, THW, and PICUD respectively.

Table 5
Output variable membership function.

Output Variable	Label	Meaning	Membership Function (Slope changes at a and b)	Graph of Membership Function
Risk Level	HighRL	Driving Situation at high risk level	Line form $\begin{cases} a = 0.5 \\ b = 1 \end{cases}$	
	MedRL	Driving Situation at medium risk level		
	LowRL	Driving Situation at low risk level		

deceleration of 10 m/s^2 and 8 m/s^2 is assumed for dry and wet roads respectively for AEB systems (Jeppsson et al., 2018)) and road friction coefficient set at 0.8, and using Prescan built-in 2D simple vehicle dynamics model and radar sensor model (“2D simple” refers to “a model that is capable of simulating a car’s longitudinal, lateral and roll

Table 6
Output variable fuzzy rule.

Variable	Fuzzy Input					Fuzzy Output	
	TTC	THW	PICUD	Risk Level			
Label	Soft	Critical	Soft	Critical	Soft	Critical	
Rule1	1	0	1	0	1	0	LowRL
Rule2	0	1	1	0	1	0	MedRL
Rule3	1	0	0	1	1	0	MedRL
Rule4	1	0	1	0	0	1	MedRL
Rule5	0	1	0	1	1	0	HighRL
Rule6	0	1	1	0	0	1	HighRL
Rule7	1	0	0	1	0	1	HighRL
Rule8	0	1	0	1	0	1	HighRL

Table 7
Optimal fuzzy parameters.

Parameter	TTC1	gap1	THW1	gap2	PICUD1	gap3
Optimal Value	0.558	2.471	0.756	2.997	-14.488	6.498

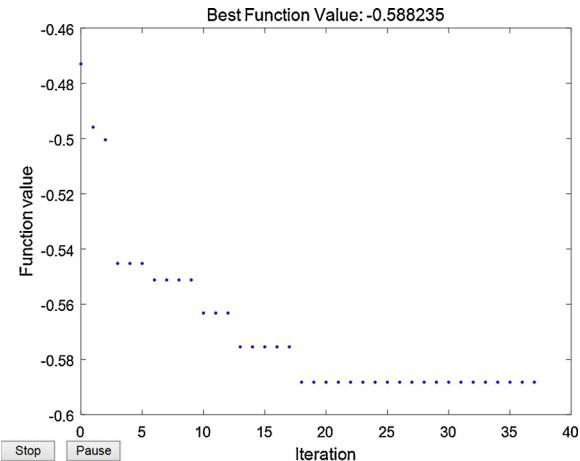


Fig. 9. PSO optimization iterative process output from Matlab.

motion”, but neglects vehicle movement in the vertical direction and assumes the front/rear two tires have the same angle and speed like a bicycle, also referred to as bicycle models (Prescan, 2018)). Finally, simulation results of each algorithm under each test scenario were

obtained and shown in Fig. 12.

In the paper, the performance of the algorithms was evaluated based on the final distance from the front of SV to the rear end of POV when SV stops, i.e., the final relative distance d_{end} . The distribution of d_{end} across different simulation scenarios could indicate the conservativeness/aggressiveness of the algorithm, where a larger average d_{end} generally implies a more conservative system while a smaller average d_{end} indicates a more aggressive one. Generally, a system being too conservative could interrupt driver’s normal driving, while being too aggressive could lead to failures in collision avoidance, and thus a more balanced (between conservativeness and aggressiveness) system is preferred. It could be seen from the simulation results that both the proposed fuzzy logic algorithm and the Mazda algorithm could avoid collision under all Euro-NCAP test scenarios ($d_{end} > 0$). However, the final relative distances by the fuzzy logic algorithm are more concentrated ($d_{end} \in 2\text{--}12 \text{ m}$) compared to Mazda algorithm (d_{end} tends to be large especially in CCRm scenarios, where d_{end} could be larger than 20 m when SV is at speed over 80 km/h), indicating the timing of AEB by the fuzzy logic algorithm is generally more balanced (not being too conservative/early which may interrupt the driver’s normal driving, nor

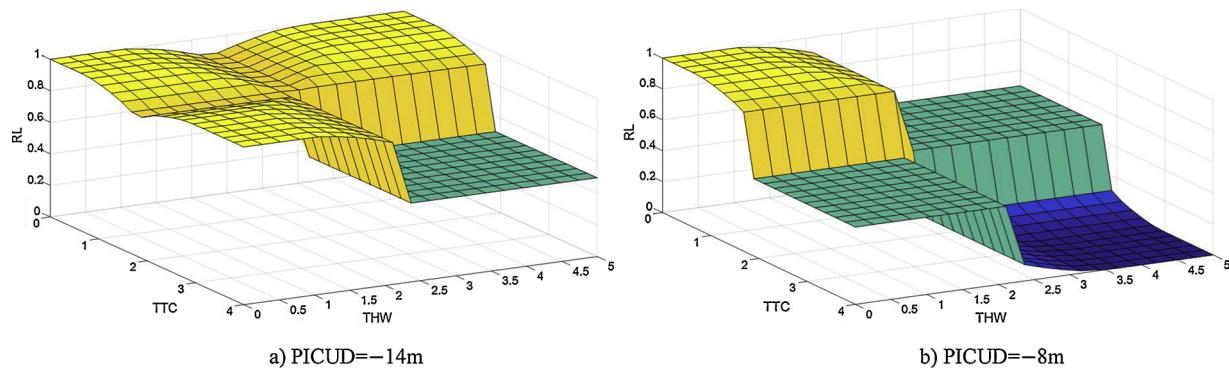


Fig. 10. Optimal fuzzy logic output surface.

being too aggressive/late which may fail in collision avoidance) than the Mazda algorithm. At the same time, the Berkley, Honda and SeungwukMoon algorithms could not avoid collisions in several CCRs and CCRb scenarios (especially those with higher approaching speed of SV and larger deceleration of POV), and the final relative distances in several successful trials (i.e., successfully avoiding collision) by these algorithms could be lower than 2 m, indicating these algorithms are more radical compared to the proposed fuzzy logic algorithm and could be prone to situations where collisions cannot be avoided. In sum, the proposed {TTC, THW, PICUD}-based fuzzy logic algorithm shows promising results in balancing the objectives of avoiding collision and reducing the possibility of interrupting driver's normal driving compared with other algorithms. In future research, human-in-the-loop study could be further implemented to validate/assess the algorithm's interference with driver's normal driving concretely.

It should be noted that the final distance by the proposed fuzzy logic algorithm increases when the speed of the subject vehicle reaches over 70 km/h in the CCRm scenario, which is not in accordance with one would expect. This could be due to the limited deceleration observations with speed ≥ 70 km/h (see the histogram of the speed at critical event-reaction braking initiation in the sampled deceleration profiles in Fig. 13). In other words, the proposed algorithm may fit better in scenarios with the SV speed below 70 km/h.

For those scenarios with SV's speed ≥ 70 m/h, given the same fuzzy logic developed based on lower speed samples, to maintain similar TTC/THW/PICUD level for brake initiation under the same speed of the leading vehicle (POV), SV with higher speed generally needs to brake at a larger relative distance. For example, under $TTC = \Delta d / \Delta v$, same TTC given a higher Δv (higher-speed SV) means a higher Δd (i.e., critical braking distance) at brake initiation (see the relative distances at the moment when fuzzy logic output changes to 1 in the history data shown in Fig. 14-17 below). And such larger critical braking distances could end in larger final relative distances when the increase in critical braking distance exceeds the increase in stopping distance of SV (i.e.,

the unexpected increasing trend in d_{end} for ≥ 70 km/h CCRm scenarios as discussed above). By including more higher-speed observations in developing the fuzzy logic, such limitation is expected to be avoided/improved in the future.

6. Conclusion

This study proposed a forward collision avoidance algorithm where the critical moment for initiating braking control was identified using fuzzy logic rules, which were extracted from the clustering results of driver deceleration profiles (assuming that online driving risk levels could be identified based on offline deceleration profiles). In the study, deceleration curves (time series) featuring “critical event-reaction braking” were extracted from Virginia Tech “100-car” near-crash records. The concept of critical event-reaction braking was proposed (vs. normal event-reaction braking) to capture the moment when the driver deems current driving unsafe and takes evasive actions, and was employed as a reference to CAS timing consistent with a driver’s “successful” decision/manipulation of last-second braking (CAS should implement AEB at the start of critical event-reaction braking of highest risk level). Risk levels of critical event-reaction braking were classified by clustering the resultant deceleration curves using *Spectral* clustering technique (based on curve range distance and change rate of curve characteristics). A fuzzy logic was finally developed to represent the clustered risk levels (based on offline deceleration profiles) by using a set of online safety indicators {TTC, THW, and PICUD} (representing three uncertain critical driving scenarios respectively), which could serve as an online risk classification algorithm for CAS. Euro-NCAP test scenarios were simulated for algorithm validation based on Prescan, and results show that the proposed fuzzy logic-based algorithm generally outperforms other classical AEB algorithms on timing of intervention (balancing the objectives of avoiding collision and reducing interference with driver’s normal driving).

The core of contribution of the paper lies in the combination of

Euro NCAP Test Methods for AEB Systems	CCRs: Approach to stationary target	CCRm: Approach to slower target	CCRb: Approach to braking target		
Scenario Parameter Setting	Speed of SV ①30 km/h ②40 km/h ③45 km/h ④50 km/h ⑤55 km/h ⑥60 km/h ⑦65 km/h ⑧70 km/h ⑨75 km/h ⑩80 km/h	Speed of POV 0km/h	Speed of SV ①30 km/h ②40 km/h ③50 km/h ④60 km/h ⑤70 km/h ⑥75 km/h ⑦80 km/h	Speed of POV 20 km/h	Speed of SV ①50 km/h ②50 km/h ③50 km/h ④50 km/h 50 km/h, $a=-2m/s^2(d_0=12m)$ 50 km/h, $a=-2m/s^2(d_0=40m)$ 50 km/h, $a=-6m/s^2(d_0=12m)$ 50 km/h, $a=-6m/s^2(d_0=40m)$

Fig. 11. Euro-NCAP AEB test scenarios and parameter settings (Euro-NCAP, 2013).

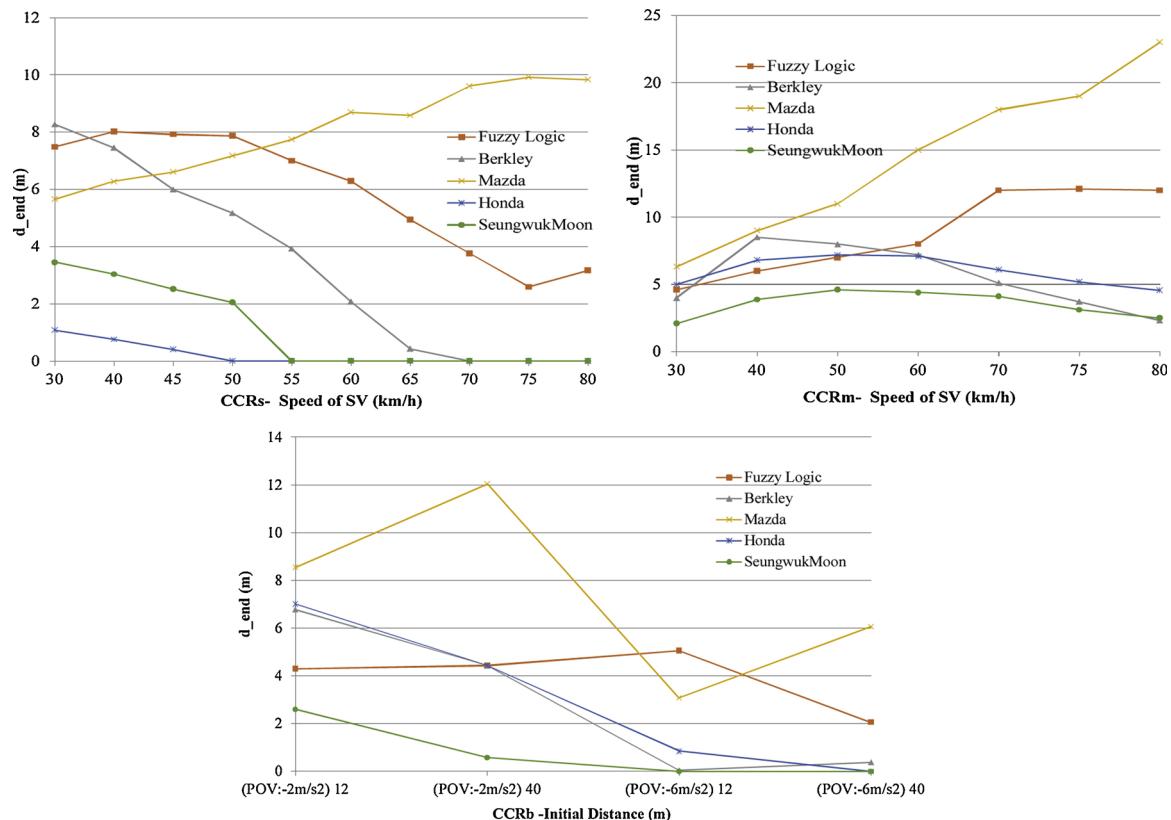


Fig. 12. Relative distances between two vehicles at the end of tests.

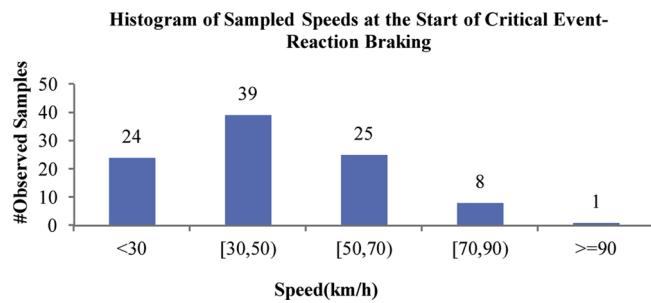


Fig. 13. Histogram of sampled speeds at the start of critical event-reaction braking.

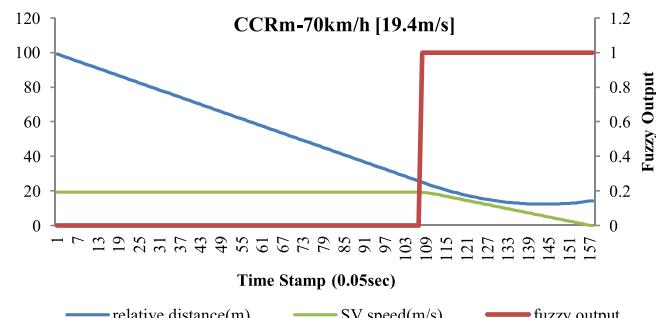


Fig. 15. Simulation history data for CCRm-70 km/h scenario.

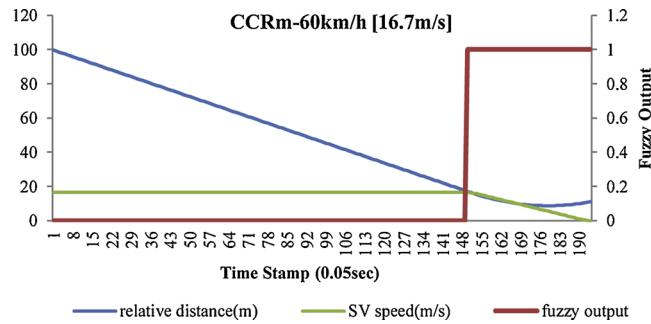


Fig. 14. Simulation history data for CCRm-60 km/h scenario.

offline deceleration profile segmentation and online risk level classification, that is, to examine braking control algorithm based on clustered deceleration profiles. Specifically, deceleration profiles have not been researched from the perspective of curve clustering to the best of the authors' knowledge. Few attempts have been made in exploring the connection between offline deceleration curve and online braking

control, however their deceleration profile segmentation was based on predefined thresholds and lacked validation. The proposed framework provides a new perspective on real-time risk level classification and collision avoidance system development. However, since there is a limitation in data sample size representing critical event-reaction braking, more deceleration profiles (by collecting more near-crash

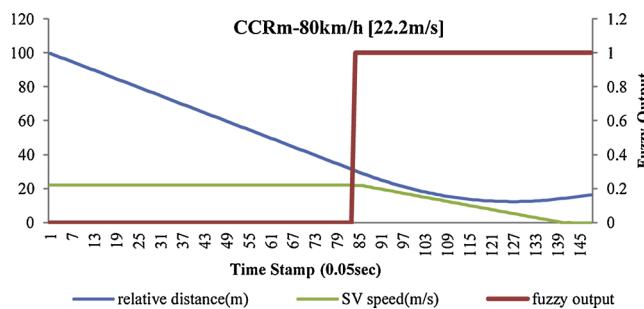


Fig. 17. Simulation history data for CCRm-80 km/h scenario.

records from other resources) should be explored to improve parameter tuning of the proposed fuzzy logic in the future (especially obtaining more higher-speed observations to overcome the current limitation in the algorithm for higher-speed braking scenarios). Besides, traffic/driver/vehicle characteristics (such as vehicle type, driver state, and vehicle response during braking, etc.) need to be investigated in future research concerning their possible effects on timing of critical braking. Also, variations of safety indicators representing uncertain critical driving scenarios could be further considered, and the probability of the uncertain scenarios could also be explored (by predicting accelerating/decelerating behaviors of SV and POV in V2V environments) and introduced into the fuzzy logic (by assigning probability-based weights of fuzzy rules) to improve its risk level classification performance. In addition, other machine learning classification algorithms (such as Support Vector Machine) instead of fuzzy logic could be explored to learn the effective representation of risk levels derived from offline deceleration profiles in further study.

Acknowledgements

This work has been supported by the National Key Research and Development Program of China (2017YFB0102603), National Natural Science Foundation of China (U1564201, U1762264, 51875255, 61601203), Key Research and Development Program of Jiangsu Province (BE2016149), and Natural Science Foundation of Jiangsu Province (BK20180100). The authors would also thank Virginia Tech for providing the “100-car” naturalistic driving data, and Marco Dozza from Chalmers – University of Technology for providing the Natware visualization tool.

References

- Bach, F.R., Jordan, M.I., 2006. Learning spectral clustering, with application to speech separation. *J. Mach. Learn. Res.* 7, 1963–2001.
- Beinum, A.V., Farah, H., Wegman, F., et al., 2016. Critical assessment of methodologies for operations and safety evaluations of freeway turbulence. *Transp. Res. Record J. Transp. Res. Board* 2556, 39–48.
- Brannstrom, M., Sjöberg, J., Coelingh, E., 2008. A situation and threat assessment algorithm for a rear-end collision avoidance system. *IEEE Intelligent Vehicles Symposium* 265, 102–107.
- Brännström, M., Coelingh, E., Sjöberg, J., 2014. Decision-making on when to brake and when to steer to avoid a collision. *Int. J. Veh. Saf.* 7 (1), 87–106.
- Cabrera, A., Gowal, S., Martinoli, A., 2012. A new collision warning system for lead vehicles in rear-end collisions. *IEEE Intelligent Vehicles Symposium*, pp. 674–679.
- Chang, H., Yeung, D.Y., 2005. Robust path-based spectral clustering with application to image segmentation. *10th IEEE International Conference on Computer Vision* 1, 278–285.
- Crick, J.H., McKenna, F.P., 1992. Hazard perception: can it be trained? *Proceedings of Manchester University Seminar: Behavioural Research in Road Safety II*.
- Doi, A., Butsuen, T., Niibe, T., et al., 1994. Development of a rear-end collision avoidance system with automatic brake control. *JSAE Rev.* 15 (4), 335–340.
- Dozza, M., 2013. What factors influence drivers’ response time for evasive maneuvers in real traffic. *Accid. Anal. Prev.* 58, 299–308.
- Euro-NCAP, 2013. Test Protocol-AEB Systems.
- Evans, T., Macdonald, W., 2002. Novice driver situation awareness and hazard perception: an exploratory study. *Road Safety Research and Policing and Education Conference*.
- Fujita, Y., Akuzawa, K., Sato, M., 1995. Radar brake system. *JSAE Rev.* 16 (1), 113.
- George, J.K., Bo, Y., 1994. Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Goodrich, M.A., Boer, E.R., Inoue, H., 1999. A characterization of dynamic human braking behavior with implications for ACC design. *IEEE/IEE/JSAI International Conference on Intelligent Transportation Systems, Proceedings*, pp. 964–969.
- Halim, Z., Kalsoom, R., Bashir, S., et al., 2016. Artificial intelligence techniques for driving safety and vehicle crash prediction. *Artif. Intell. Rev.* 46 (3), 1–37.
- Haworth, N., Mulvihill, C., Symmons, M., 2005. Hazard Perception and Responding by Motorcyclists: Background and Literature Review. Report No. 235. Monash University Accident Research Centre, Melbourne.
- Hayward, J.C., 1972. Near-miss determination through use of a scale of danger. *Highway Res. Record* 384, 24–34.
- Hillenbrand, J., Spieker, A.M., Kroschel, K., 2006. A multilevel collision mitigation approach—its situation assessment, decision making, and performance tradeoffs. *IEEE Trans. Intell. Transp. Syst.* 7 (4), 528–540.
- Jeppsson, H., Östling, M., Lubbe, N., 2018. Real life safety benefits of increasing brake deceleration in car-to-pedestrian accidents: simulation of vacuum emergency braking. *Accid. Anal. Prev.* 111, 311–320.
- Keller, C.G., Dang, T., Fritz, H., et al., 2011. Active pedestrian safety by automatic braking and evasive steering. *IEEE Trans. Intell. Transp. Syst.* 12 (4), 1292–1304.
- Kennedy, J., Eberhart, R., 2002. Particle swarm optimization. *IEEE International Conference on Neural Networks* 4, 1942–1948.
- Kluger, R., Smith, B.L., Park, H., et al., 2016. Identification of safety-critical events using kinematic vehicle data and the discrete fourier transform. *Accid. Anal. Prevent.* 96, 162–168.
- Lee, D., Yeo, H., 2016. Real-time rear-end collision-warning system using a multilayer perceptron neural network. *IEEE Trans. Intell. Transp. Syst.* 17 (11), 3087–3097.
- Li, Y., Lu, J., Xu, K., 2017. Crash risk prediction model of lane-change behavior on approaching intersections. *Discrete Dyn. Nat. Soc.* 2017, 1–12.
- Liao, T.W., 2005. Clustering time series data — a survey. *Pattern Recognit.* 38 (11), 1857–1874.
- Lin, Q., Cheng, B., Qu, X., Zhang, B., et al., 2012. Vehicle rear-end collision warning algorithm based on driver’s braking behavior. *Automotive Eng.* 34 (3), 232–235.
- Liu, Y.C., Wu, T.J., 2009. Fatigued driver’s driving behavior and cognitive task performance: effects of road environments and road environment changes. *Saf. Sci.* 47 (8), 1083–1089.
- Luxburg, U., 2007. A Tutorial on Spectral Clustering. Kluwer Academic Publishers.
- MacCall, J.C., Trivedi, M.M., 2007. Driver behavior and situation aware brake assistance for intelligent vehicles. *Proceedings of the IEEE* 95 (2), 374–387.
- Matlab 2018R, 2018. Documentation. Mathworks Inc.
- McNicol, D., 2004. A Primer of Signal Detection Theory. Lawrence Erlbaum, Hillsdale, NJ.
- Moon, S., Yi, K., 2008. Human driving data-based design of a vehicle adaptive cruise control algorithm. *Veh. Syst. Dyn.* 46 (8), 661–690.
- Moon, S., Moon, I., Yi, K., 2009. Design, tuning, and evaluation of a full-range adaptive cruise control system with collision avoidance. *Control Eng. Pract.* 17 (4), 442–455.
- Mullakkal-Babu, F.A., Wang, M., Farah, H., et al., 2017. Comparative assessment of safety indicators for vehicle trajectories on highways. *Transp. Res. Record: J. Transp. Res. Board* 2659, 127–136.
- Prescan, 2018. TNO. Available at: <http://www.tassinternational.com/prescan-overview>.
- Schölkopf, B., Tsuda, K., Vert, J., 2004. A Primer on KernelMethods. Kernel Methods in Computational. MIT Press, Massachusetts, USA.
- Seiler, P., Song, B., Hedrick, J.K., 1998. Development of a collision avoidance system. *Neurosurgery* 42 (2) 492–492.
- Shai, B.D., Eyal, K., Yishay, M., 1997. Online learning versus offline learning. *Mach. Learn.* 29 (1), 45–63.
- Valgren, C., Duckett, T., Lilienthal, A., 2007. Incremental spectral clustering and its application to topological mapping. *IEEE International Conference on Robotics and Automation*, pp. 4283–4288.
- Vicente, M., Josué, P., Godoy, J., et al., 2012. A fuzzy aid rear-end collision warning/avoidance system. *Expert Syst. Appl.* 39 (10), 9097–9107.
- Virginia Tech Transportation Institute, 2014. VTTI Data Warehouse. Available at: <http://forums.vtti.vt.edu/index.php?/files/category/3-100-car-data/>.
- Wada, T., Doi, S., Tsuru, N., et al., 2010. Characterization of expert drivers’ last-second braking and its application to a collision avoidance system. *IEEE Trans. Intell. Transp. Syst.* 11 (2), 413–422.
- Wang, J., Zhang, L., Zhang, D., et al., 2013. An adaptive longitudinal driving assistance system based on driver characteristics. *IEEE Trans. Intell. Transp. Syst.* 14 (1), 1–12.
- Wang, J., Wu, J., Li, Y., 2015a. The driving safety field based on driver–vehicle–road interactions. *IEEE Trans. Intell. Transp. Syst.* 16 (4), 2203–2214.
- Wang, J., Zheng, Y., Li, X., et al., 2015b. Driving risk assessment using near-crash database through data mining of tree-based model. *Accid. Anal. Prev.* 84, 54–64.
- World Health Organization, 2016. Road Traffic Injuries. Available at: <http://www.who.int/mediacentre/factsheets/fs358/en/>.
- Zhang, Y., Antonsson, E.K., Grote, K., 2006. A new threat assessment measure for collision avoidance systems. *IEEE Intelligent Transportation Systems Conference, Proceedings*, pp. 968–975.