Capturing Driver Evasive Manoeuvres In Pre-crash Phases From Large-scale Real-world Critical Scene Videos

Shun Gan, Quan Li, Qingfan Wang, Taisong Cui, Bingbing Nie

I. INTRODUCTION

Accurate assessment of drivers' active behaviours in pre-crash phases contributes to an enhanced understanding of injury prevention and mechanisms in real-world crashes [1]. Active responses include the conscious or subconscious evasive manoeuvres of an alert driver in an impending crash event, such as braking or swerving. In unavoidable collisions, driver injury risk and severities are significantly associated with posture change and muscle activation during such manoeuvres [2-3]. Previous studies investigated drivers' evasive manoeuvres through the event data recorder (EDR) from real-world traffic accidents and from driving simulation experiments [4]. Due to the high cost of virtually reconstructing traffic scenes, driving simulation experiments have been limited to a few specific critical traffic scenes [5]. To obtain a large-scale dataset on driver responses, this study proposes an in-lab experimental framework for efficiently capturing evasive manoeuvres when stimulated by real-world critical scene videos (Fig. 1). The evasive manoeuvre data from different real-world critical scenes are beneficial to behavioural pattern recognition. Thus, the behaviour pattern of collision avoidance was analyzed and discussed using the preliminary results of this study.

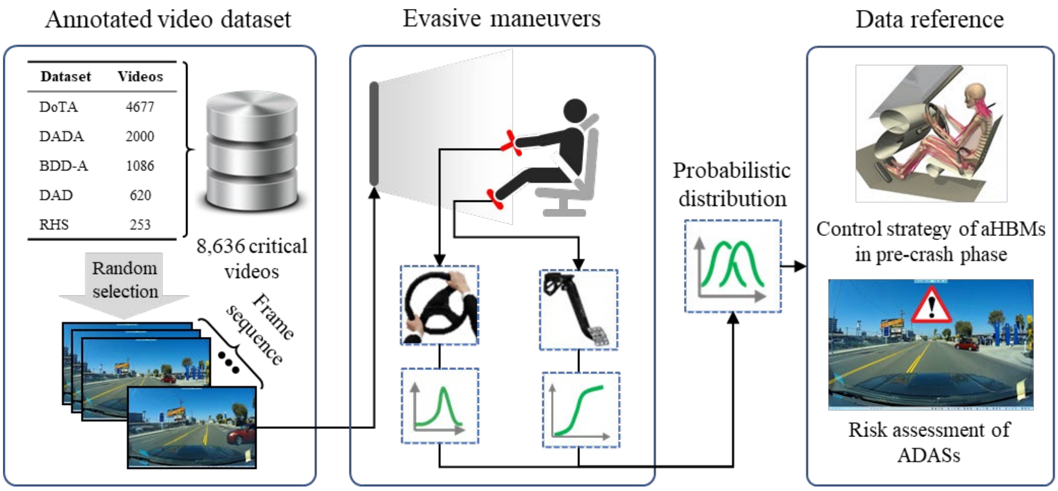

Fig. 1. The overall experimental design. The stimuli videos were selected from the 8,636 real-world critical scenes with detailed annotations; evasive manoeuvres were recorded by the steering-wheel angle and brake pedal position; the statistic results provided data reference for the control strategy of active human body models (aHBMs) and risk assessment of advanced driver assistance systems (ADASs) in critical scenes.

II. METHODS

Stimuli videos

Drivers view stimuli videos of dynamic real-world critical scenes in a driving simulator and carry out manoeuvres to evade the hazard. Driver evasive manoeuvres were recorded by the steering-wheel angle and the brake pedal position. The stimuli video dataset comprises 8,636 real-world critical traffic scene videos captured by the ego vehicle dashcam from five public datasets, i.e. DoTA [6], DADA [7], BDD-A [8], DAD [9] and RHS [10]. As a preliminary study, only RHS was implemented in the current experiment. The RHS published by MIT Agelab contains 250 normal traffic scene videos and 253 critical scene videos with detailed temporal annotations. Each

B. Nie (e-mail: nbb@tsinghua.edu.cn; tel: +86-10-6278-8689) is an Associate Professor and S. Gan, Q. Li and Q. Wang are PhD students in Mechanical Engineering, in the School of Vehicle and Mobility at Tsinghua University. T. Cui is a researcher at Chongqing Changan Automobile Company Limited. Affiliations: State Key Laboratory of Automotive Safety and Energy, School of Vehicle and Mobility, Tsinghua University (B. Nie, S. Gan, Q. Li, Q. Wang); State Key Laboratory of Vehicle NVH and Safety Technology, Chongqing, China (B. Nie, S. Gan, T. Cui).

video duration is 8 s, involving the critical event leading to a collision or a near-collision. Three professional annotators marked each video that involved a critical scene with the necessary annotated timestamp, e.g. the first visual indication (FVI) of a potential hazard in the critical scene. We randomly selected 50 critical videos and 50 normal videos from the RHS dataset with 30 frames per second (fps) and 1280 x 720 pixel resolution to ensure a uniform format for all of the stimuli videos.

Subjects

All subjects had normal (or corrected to normal) vision, hearing and driving behaviour, with no disability or heart disease. The experimental procedures were approved by the Institutional Review Board (IRB) of Tsinghua University. The procedures were performed following the approved guidelines.

A total of n=30 subjects (all males) were recruited, all of whom met the inclusion criteria. Eight subjects were excluded from the final analysis: five of them aborted during the experiment due to strong motion sickness, while the data of three subjects were not entirely recorded due to equipment failure. Ten experienced drivers (average age 49; driving experience 21.5 years; driving frequency 28.5 times per month; driving distance 44,000 km/year) were recruited via DiDi, the largest online taxi-hailing and private car-hailing service in China. Twelve novice drivers (average age 25; driving experience 3.5 years; driving frequency 2.0 times per month; driving distance 1,308 Km/year) were recruited among the graduate student population.

Procedure

The stimuli videos were presented on a 27-inch Philips monitor with 1920 x 1080 pixel resolution and a 60 Hz refresh rate. The subject was seated approximately 1.5 m away from the monitor and had a visual field of 24.31° vertically and 13.86° horizontally. Steering-wheel angle and brake pedal position were recorded per Logitech G29 driving force steering wheel and pedals, coupling with Simulink sampling 300 fps.

Before conducting the formal experiments, subjects were required to participate in pre-tests until they were familiar with the test procedure. In pre-tests, subjects were informed that they were driving a semi-autonomous vehicle with ACC (adaptive cruise control) function and were required to take over when potential hazards appeared in traffic scenes. Subjects were told to stop at any time during the experiment process if they experienced discomfort caused by motion sickness.

III. INITIAL FINDINGS

We defined two metrics, i.e. evasive reaction time (ERT) and hazard prediction event (HPE), to quantitatively evaluate the temporal features of evasive manoeuvres in the given scenarios (Fig. 2). ERT is defined as the time difference between the annotated FVI (i.e. the first visual indication of a potential hazard in the critical scene) and the driver's first reaction time within a 5 s response time window. HPE is defined as the situation in which the evasive reaction happens before FVI, which denotes that the driver had anticipated the potential hazard in the scene and taken the evasive manoeuvres in advance of collisions.

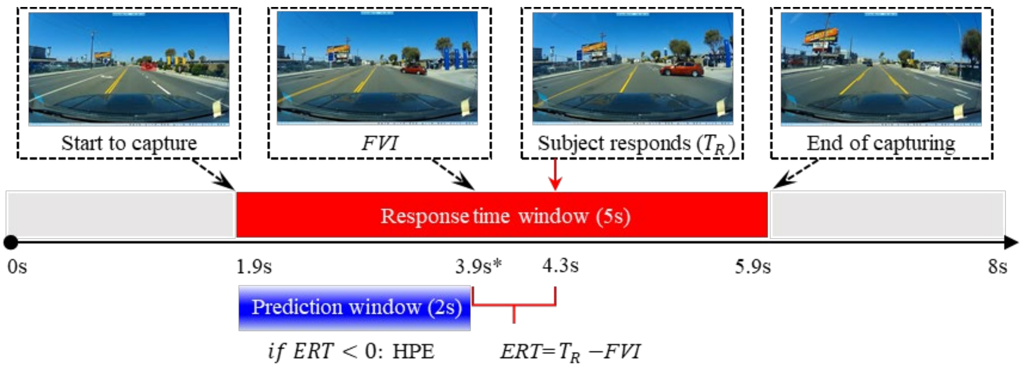

Fig. 2. Illustration of temporal annotations and metrics. * denotes the FVI timestamp; 1.9 s and 5.9 s denote start and end of response time window; the mean ERT is 4.3 s in the above critical scene (shown: No. 013 critical scene from the RHS dataset).

Initial results showed that the subjects actively responded to most of the real-world critical videos (89.5%) within the pre-defined response window. Significantly, the "swerving first" (30.9%) and "braking only" (28.6%) were the most common evasive patterns (Fig. 3). The probability density curves of ERT were fitted through Gaussian kernel density estimation (Fig. 4). ERT of the highest density for both braking and swerving was about 0.6 s. Compared with the ERT of braking, the ERT of swerving occurred more frequently in the time interval from

-2 to 0 s. Furthermore, the mean ERT of swerving (0.46±1.18 s) is shorter than that for braking (0.59±0.90 s).

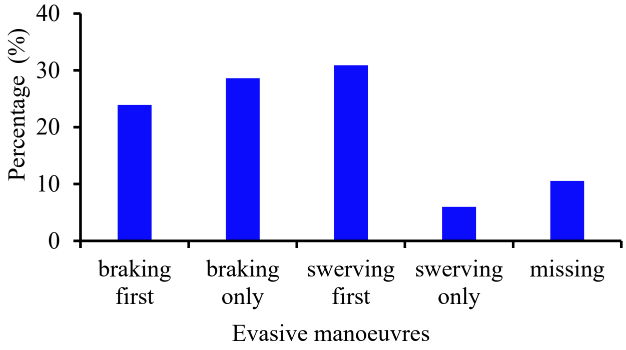

Fig. 3. The proportions of evasive manoeuvres in critical scenes.

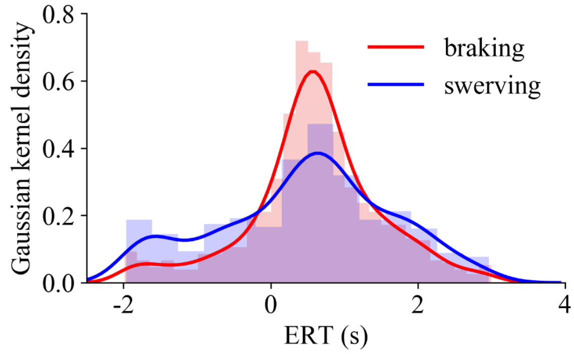

Fig. 4. The distribution of the braking and swerving ERT. Time zero denotes FVI.

In descending order by ERT, 50 critical scenes revealed that human drivers have different evasive patterns to respond to diverse critical events (Fig. 5(a)). The cases with relatively shorter ERT usually correspond to some common critical scenes that are predictable for human drivers, while the cases with relatively longer ERT feature some unpredictable and complicated scenarios involving multiple traffic events. The scenes with shorter ERT tended to have a greater probability of HPE (Fig. 5(b)). A significant positive correlation between the braking and swerving time was affirmed (Spearman coefficient = 0.75, $p < 0.01$). Four representative cases at the two ends of ERT sorting were analyzed (Fig. 5(c)), i.e. Case 26 (ERT=1.80±0.94 s), Case 52 (ERT=1.52±1.11 s), Case 85 (ERT=-0.65±1.42 s), and Case 47 (ERT=-0.94±0.67 s). Case 26 showed two sedans colliding on the highway, with one of them intruding into the driving lane after losing control. Case 52 displayed an excavator crossing the road and reversing into the driving lane. In these two scenes, few subjects (1.25 on average) predicted the critical scene's evolution and they took a long time to make an evasive response. Case 85 and Case 47 demonstrated a similar gradual-onset critical scene in which a sedan broke into the driving lane from the T-junction and from the left roundabout, respectively. In this test, most of the subjects (19.5 on average) anticipated the common hazard and took evasive manoeuvres in advance.

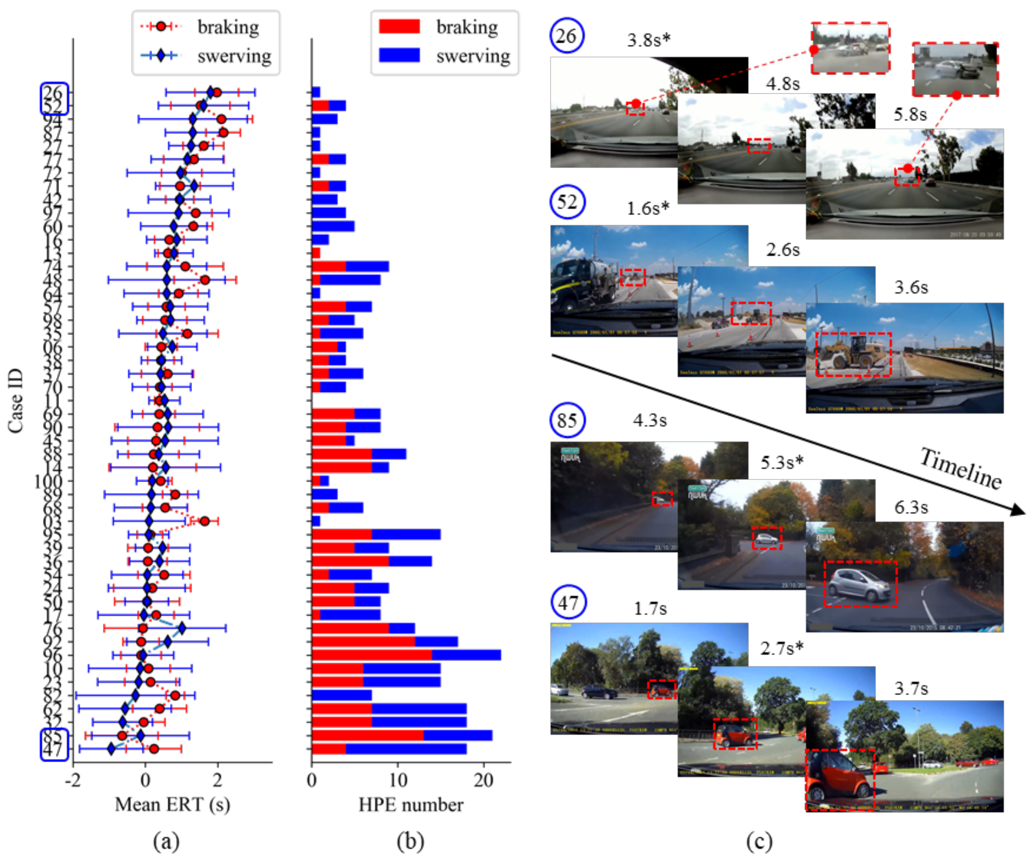

Fig. 5. (a) Mean ERT in each case; (b) the HPE number; (c) four representative cases. * denotes the FVI in the stimuli video timeline.

IV. DISCUSSION

The diverse critical traffic scenes used in this study triggered drivers' to perform various evasive manoeuvres. The present experimental framework demonstrates an efficient approach to capturing drivers' evasive manoeuvres in critical scenes with temporal annotations. Identifying the resultant behavioural patterns will help to establish a generalised representation of drivers' active behaviours, providing data reference on active human body models (aHBMs) for active safety systems (e.g. ADAS).

Several behaviour patterns were identified. On average, the ERT of braking lags behind that of swerving (0.13 s). The "swerving first" (30.9%) and "braking only" (28.6%) proved to be the two most common evasive manoeuvres used to avoid collision in emergency events. The critical scene with similar traffic features could stimulate driver response with similar ERT values. In some typical, therefore more predictable, critical scenes drivers can anticipate potential hazards and take evasive manoeuvres in advance. An important factor was individual driving experience, which significantly influenced each driver's behavioural pattern. Compared with less experienced drivers, experienced drivers have a greater ability to anticipate potential hazards, evidenced by their shorter reaction time (0.39±0.86 s vs. 0.80±0.58 s) and higher probability of hazard prediction (21.5% vs. 6.7%). This means individual behaviour patterns in collision avoidance should be taken into account in evasive manoeuvres data, which can in turn contribute to the control strategy of aHBMs and risk assessment in critical scenes.

Several limitations must be noted. First, we did not record the muscle activity levels or drivers' body kinematic features through the motion capture system in a driving simulator with moving capability. A more comprehensive data acquisition system on active behaviour is needed to provide detailed data on drivers' dynamic responses. Second, only 50 critical scenes were presented as the stimuli sources in the present work. The dataset size should be expanded in our subsequent studies.

V. ACKNOWLEDGEMENTS

This study was in part supported by the National Natural Science Foundation of China (52072216, 51705276) and by the State Key Laboratory of Vehicle NVH and Safety Technology (NVHSKL-2020012).

VI. REFERENCES

[1] Hault, D. A., *et al*., *Accid Anal Prev,* 2011.
[2] Huelke, D. F., et al., *J Trauma Acute Care Surg*, 1995.
[3] Nie, B., *et al*., *Traffic Inj Prev*, 2018.
[4] Riexinger, L. E., *et al*., *IRCOBI*, 2019.
[5] Hu, M., *et al*., *IRCOBI*, 2017.
[6] Yu, Y., *et al*., arxiv.org/abs/2004.03044, 2020.
[7] Fang, J., *et al*., *ITSC*, 2019.
[8] Xia, Y., *et al*., *ACCV*, 2018.
[9] Chan, F., *et al*., *ACCV*, 2016.
[10] Wolfe, B., *et al*., *J Exp Psychol Anim Learn Cogn*, 2019.