Driving patterns in connected environments

A case study of intersection-approaching behavior of professional and non-professional drivers

Zhang, Hailun; Fu, Rui; Wang, Jianqiang; Calvert, Simeon C.; van Lint, Hans

DOI
10.1016/j.trf.2024.04.014

Publication date
2024

Document Version
Final published version

Published in
Transportation Research Part F: Traffic Psychology and Behaviour

Transportation Research Part F: Psychology and Behaviour 103 (2024) 230–259

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

Driving patterns in connected environments: A case study of intersection-approaching behavior of professional and non-professional drivers

Hailun Zhang [a,c,*], Rui Fu [b], Jianqiang Wang [a], Simeon C. Calvert [c], Hans van Lint [c]

[a] *School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China*
[b] *School of Automobile, Chang'an University, Xi'an 710064, China*
[c] *Transport and Planning Department, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft 2628 CN, The Netherlands*

ARTICLE INFO

Keywords:
Connected vehicle
HMI
Intersection
Response mechanism
Driving patterns

ABSTRACT

The in-vehicle communication provides promising opportunities to improve the road safety and traffic efficiency. Previous studies demonstrated that the professional drivers have better driving skills than the non-professional drivers who allocate more attention to secondary tasks. However, they may not be sensitive to the new in-vehicle technology. In addition, these qualitative studies failed to elaborate on the visual and response behavior differences among different driver groups (professional drivers such as taxi, bus, motorcoach, and non-professional drivers), and lacked the quantitative analysis of driving patterns in a new environment. This paper explores the differences in visual interaction, response characteristics, driving performance, and behavior patterns between the professional and non-professional drivers in the connected environment through a case study of intersection-approaching behavior using a driving simulator. More precisely, two driving scenarios (baseline and human–machine interface (HMI)) were designed in the driving simulator, and 65 participants, including 34 professional drivers and 31 non-professional drivers, completed the experiment. In the HMI scenario, the driver was provided with the signal light phase and phase transition remaining time of the current intersection. This paper also proposes a driving pattern extraction model based on the Bayesian non-parametric method combined with a text clustering algorithm to perform a quantitative description of the driving patterns. The results show that the professional drivers tend to interact less with the HMI compared with the non-professional drivers. Moreover, the professional drivers' first gaze at the HMI occurs and responds earlier. The proposed driving model can effectively describe 7 patterns of intersection-approaching behavior. The connected information can significantly improve the efficiency of the intersection traffic and the driving behavior. However, the professional drivers are more responsive and behave more consistently. This study can provide insights into the development of personalized assisted driving systems, as the two driving populations differ in their interactions, responses, and behavioral patterns.

* Corresponding author at: School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China.
E-mail address: Iszhanghailun@outlook.com (H. Zhang).

https://doi.org/10.1016/j.trf.2024.04.014
Received 6 November 2023; Received in revised form 16 April 2024; Accepted 20 April 2024
Available online 25 April 2024
1369-8478/

1. Introduction

1.1. Motivations

Using the vehicle-to-infrastructure (V2I) technology combined with the vehicle-mounted human–machine interface (HMI) to provide drivers with auxiliary information on daily driving tasks, especially the traffic lights information when approaching intersections, can significantly reduce the traffic congestion and accident rates, and improve the driving behavior and intersection efficiency (Zhou et al., 2019; Tielert et al.; Yang et al., 2021). However, the development of the communication and sensor technology changed the space–time dimension of the driver perception information as well as the driver's cognitive process and response characteristics, which are specifically reflected in its perceptual interaction process, response characteristics, attention distribution, driving performance, and behavior patterns (Yang et al., 2017; Liu and Kircher, 2018). Learning and adapting to the driver's original behavior pattern is the prerequisite for the advanced driver assistance system (ADAS) to formulate an acceptable ecological driving strategy. Therefore, exploring the driver's cognitive process and responses characteristics and understanding its behavior pattern in the connected environment, is a key topic for the development of the next-generation ADAS and autonomous driving technologies.

Many studies have demonstrated that the newly developed technologies deployed in vehicles can improve the driver performance and traffic safety, and reduce the traffic congestion, such as the in-vehicle communications and automation (Monsaingeon et al., 2021; Maag et al., 2022). Calvert et al. (Calvert et al., 2020) deduced that the cooperative adaptive cruise control significantly reduces the travel times at intelligent intersections, and platooning occurs at relatively low time headways (Calvert and Arem, 2020). In addition, it has been shown that the drivers make informed and better safe driving decisions in a connected environment. More precisely, they maintain longer time-to-collision during car-following, longer time-to-collision to pedestrians when approaching intersections, reduced deceleration during lane changes to avoid collisions, and slow driving when running yellow lights (Ali et al., 2020). Compared with the traditional environments, drivers can obtain traffic information beyond the visual range in the connected mode, which allows them to perceive and respond to various risk scenarios in advance, increasing the safety and smoothness of maneuvering (Kenney, 2011).

To date, although the in-vehicle communication technology has become more affordable, few quantitative studies investigated the drivers' response mechanism to vehicle-mounted V2I communication devices. The interactions with HMI, response characteristics, and driving patterns may vary among heterogeneous driving groups (e.g., professional versus non-professional drivers). The studies on this topic are crucial, not only because of the considerable proportion of professional drivers, but also to provide a theoretical support for the traffic department to train different types of drivers in the application of new technologies. Moreover, they have the application prospect that the manufacturers of the in-vehicle communication equipment can develop personalized applications according to the characteristics of different drivers. Therefore, this study focuses on the characteristics of driver responses to connected information, such as the traffic signal, and provide a comparative case of intersection-approaching behavior of professional and non-professional drivers.

1.2. Related studies

Many studies have been conducted to understand the impact of the driver attributes (i.e., driving experience) on the driving behavior and its relationship to traffic accidents (Savage et al., 2020; Montgomery et al., 2014; He and Donmez, 2019; Ni et al., 2010). The professional drivers (i.e., taxi, bus and coach drivers) tend to have extensive driving experience and demonstrate a mature mind and high driving performance (i.e., conservative driving characteristics) (Newnam et al., 2020). However, due to the nature of their jobs, the prolonged exposure to traffic is inevitable, which may increase their chances of being involved in an accident (Duke et al., 2010; Islam and Ozkul, 2019; Dorn and Brown, 2003). The non-professional drivers are more likely to underestimate the potential risk of driving (Castellà and Pérez, 2004), and they are at a higher crash risk than the experienced professional drivers for an identical driving time (Evans, 2004). Hong et al. (Hong et al., 2016) demonstrated that the older drivers approach intersections with more erratic behavior and higher speed. Although the age and driving experience are sometimes confounding the risk factors of driving, there is evidence that the impact of the driving experience is independent of the age (Tao et al., 2017; Cheung and McCartt, 2011). More precisely, a satisfactory driving performance could be attributed to self-regulation (Dykstra et al., 2020). Therefore, a specific analysis of the behavior of the professional and non-professional drivers is required for different driving tasks because they may have different perceptions, cognitions, and response characteristics to traffic, which may also lead to differences in driving patterns.

In recent years, researchers have focused on the differences in road behavior between professional and non-professional drivers (Öz et al., 2010; Lal and Craig, 2002; Chen et al., 2021). For instance, Oz et al. (Öz et al., 2010) studied the stress reactions and accident involvement of professional and non-professional drivers. Rosenbloom and Shahar (Rosenbloom and Shahar, 2007) showed differences in reactions and risky behavior between professional and non-professional drivers. In addition, it was deduced that non-professional drivers drove faster on urban roads compared with the drivers of taxis, minibuses, and heavy vehicles, while professional drivers were more prone to stress reactions in traffic and more likely to engage in dangerous traffic behavior. He and Donmez (He and Donmez, 2019) deduced that the experienced drivers were more conservative in engaging in second tasks in automated vehicles than novice drivers. Chen et al. (Chen et al., 2021) explored whether older professional drivers used their driving experience more effectively to reduce accident risks. The results of simulator experiments showed that the professional drivers were more capable of reducing the risk by adopting compensatory strategies. Mahajan and Velaga (Mahajan and Velaga, 2022) compared the performance of young non-professional and professional taxi drivers during fatigue driving. It was observed that the driving experience/skills of the professional drivers did not compensate for the worse effects of partial sleep deprivation. However, it is important to mention that in these studies, the differences in the performance between the professional and non-professional drivers were obtained in a conventional driving environment. Therefore, the driving performances of these two groups in the connected environment remain unclear.

Ali et al. (Ali et al., 2020) designed a connected environment to study the drivers' interactions with pedestrians entering a crosswalk from the sidewalk and their reactions to traffic signal changes (i.e., green light to yellow light) at signalized intersections. Their results showed that the driver maintained a longer time-to-collision to pedestrians and had a lower propensity of running a yellow light in a connected environment. Li et al. (Li et al., 2020) designed a driving experiment for approaching intersections and provided advice and feedback information to drivers through an HMI. Their results demonstrated that the HMI significantly improved the driver's safe driving behaviors without causing excessive mental and visual loads. Kohl et al. (Kohl et al., 2020) studied the driver distraction while receiving in-vehicle information cues by analyzing the driver's glance behavior. It has shown that the drivers don't exhibit longer glance durations towards in-vehicle information systems because they cause distractions. They prefer dynamic illustrations on the HMI rather than static information (Kraft et al., 2020).

To study the impact of the traffic signal information on the driving behavior, the researchers mainly analyzed the collision risk, traffic efficiency, and driving performance at intersections. Yan et al. (Yan et al., 2022) evaluated the short-term and long-term impact of the green signal countdown information (GSCT) on the road safety. Their driving simulator and natural driving experiments showed that GSCT reduced the frequency of red-light running (RLR) in the short term, but there was no change in the long term. In addition, GSCT led to the increase of the rear-end collisions (RAC). As the risks increased, the speed strategies and behavior patterns of the drivers also underwent significant changes. Similarly, Paul et al. (Paul et al., 2022) deduced that GSCT reduced the risk of crossing and RAC caused by RLR and inconsistent stopping behavior. Islam et al. (Islam et al., 2016) deduced that the existence of SCT significantly reduced the headway of the first car in a convoy, reduced the lost time of starting, and improved the traffic efficiency at the intersection. Long et al. (Long et al., 2013) deduced that SCT led to the increase of the frequency of yellow-light running (YLR), and it cannot significantly improve the intersection safety. Ma et al. (Ma et al., 2010) showed that the signal light device can prompt drivers to accelerate during the yellow light phase, which allows them to more smoothly cope with the signal light phase changes, prevent sudden speed changes, and reduce the distressed area at the intersection.

1.3. Objectives, contributions, and hypotheses of the study

It can be deduced from the literature review that: (i) The early efforts mainly provided auxiliary information to the driver through the HMI and focused more on the analysis of the influencing factors, while they failed to quantitatively and descriptively explain the driving behavior characteristics in the connected environment. (ii) The recent studies on the risk behavior of professional and non-

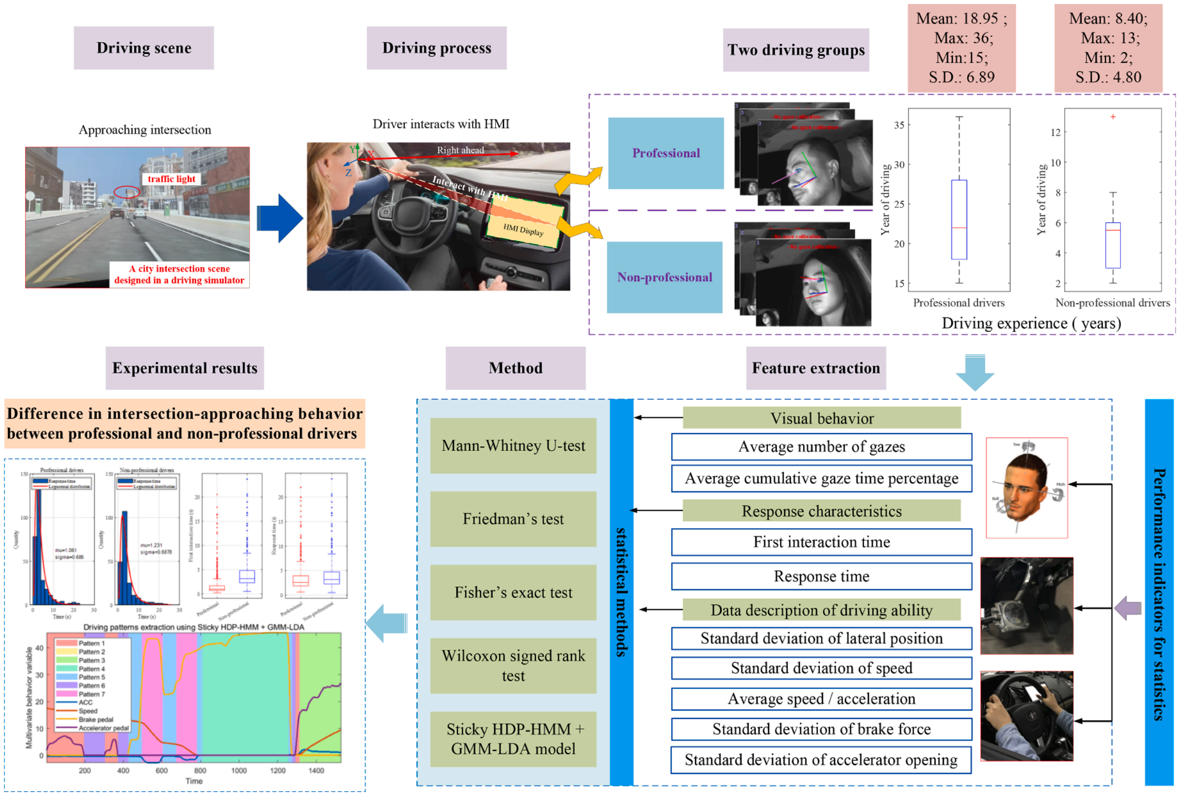


Fig. 1. Diagram of this study.

professional drivers mainly focused on car-following tasks (compensatory strategy), fatigue driving, and stress reactions in the traditional environment. (iii) Most of the connected environment studies were conducted in driving simulator environments. (iv) The professional drivers tend to be more proficient in dealing with unfavorable driving environments, thereby showing higher abilities to mitigate risks, while the inexperienced non-professional drivers may be more adaptable to new vehicle technologies. (v) A traffic signal countdown timer can potentially increase the efficiency and comfort without causing noticeable distractions. In summary, the difference in the intersection-approaching behaviors of professional and non-professional drivers in connected environments is still unclear, including the interactions with HMIs, response characteristics, and behavioral patterns.

The contributions of this paper are summarized as follows: (i) A framework for studying the intersection-approaching behavior is first proposed, and the visual perception process of drivers is visualized using a hierarchical clustering method to perform human–machine interaction feature extraction. (ii) A signal energy detection method based on the wavelet transform is then developed to monitor the driver's response characteristics in real-time. (iii) Finally, a driving pattern extraction model combining Bayesian non-parametric and text clustering algorithms is proposed, which can quantify the driving patterns distribution of intersection-approaching behavior, and non-parametric test methods are used to illustrate the influence of the connected environment on the driving behavior.

It is hypothesized that (1) when approaching an intersection, the professional drivers interact with the HMI less frequently than the non-professional drivers, and (2) after obtaining traffic light information through the HMI, they respond faster, they are less affected by the driving environment, and they maintain a relatively consistent driving performance. Based on the literature review, these hypotheses are proposed considering the possible differences in the driving skills and experience between the professional and non-professional drivers. Fig. 1 shows a schematic of this study.

2. Experimental data

2.1. Simulator and eye-tracker

A realistic intersection in a connected traffic environment was designed using a high-fidelity real-time simulator to collect the data required for this study and ensure the safety of the drivers. Fig. 2 shows the driving simulation platform used in this study. A sedan was mounted on a six-degrees-of-freedom (DOF) motion platform to provide a real sense of control to the participants, including visual, auditory, and motor simulation. In particular, it included the following experimental equipment: a real vehicle cabin, a front-view curved screen display system, a two-channel rear-view display system to observe the traffic through a rear-view mirror, an E2M 6-DOF motion platform, a control system to collect data, a high-performance industrial computer, and a scene projection system with five projectors. The simulator provides almost the same responses as a real car, such as accelerating, steering, and braking. The vehicle position and driving performance data, such as the speed, acceleration/deceleration, steering wheel angle, brake force/pedal displacement, and accelerator pedal displacement, were recorded at 60 Hz. The SmartEye 8.0 eye-tracking system was used to collect the eye movement of the drivers. It included four eye motion-tracking cameras and three infrared light sources, as shown in Fig. 3.a and b. As a non-invasive eye-tracking device, it allows the participants to make free head movements within the active working area of the eye-tracker. The system can track and record the eye movement with high precision. Fig. 3.c shows the positioning and tracking of the eye performed by the system. The sampling frequency of the eye tracker was 60 Hz. Gaussian smoothing was used to filter the fluctuations and errors in the gaze data (Guo et al., 2021).

Fig. 2. Driving simulator and components.

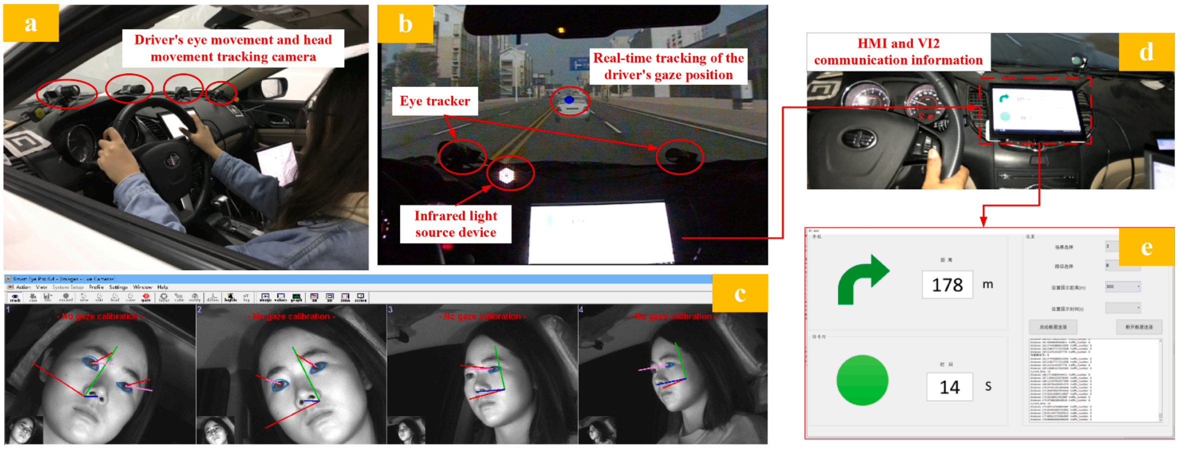


Fig. 3. The eye-tracking system and the HMI display.

2.2. HMI system

The design and development of the HMI prototype follows a human-centered design approach, and it is based on the V2X technology to obtain real-time traffic and road information, such as the vehicle trajectory information, phase and countdown of traffic lights, and distance to intersections (Vaezipour et al., 2017). Strayer et al. (Strayer et al., 2019) deduced that the design of in-vehicle information systems requires to consider which type of interaction should be available to the driver so that he/she receives intuitive and concise information rather than complex instructions. Therefore, the system consists of an 11-inch monitor and real-time communication software. A monitor was used to display connection information. It was mounted on the right side of the dashboard without blocking the view of the road environment and speedometer, as shown in Fig. 3.d. The location of the monitor was similar to that of the other in-vehicle eco-driving displays that are proposed in previous studies (Rouzikhah et al., 2013; Kircher et al., 2014). The communication software was developed in Python, and the connected traffic information was obtained in real-time by connecting the system with the data acquisition system of the simulator, as shown in Fig. 3.e. In the latter, the green arrow refers to turning right at the intersection, 178 refers to a distance of 178 m from the intersection, the green circle refers to the current traffic light phase at the intersection, and 14 indicates 14 s of green light time remaining. The right half of the HMI is primarily used to select functions in the test program. It is designed to help the tester quickly switch scenarios and routes after an experiment in order to avoid order and learning effects for the participants.

2.3. Participants

Table 1 presents the summary of the participants in the driving simulator study. All participants were asked to complete a short questionnaire providing information on age, annual driving distance, occupation, traffic violation and accident involvement records. In this study, professional drivers, including full-time taxi, public bus, and motorcoach drivers, were recruited through advertising and online registration, while the non-professional drivers were Ph.D. students, staff recruited from the university campus, and those whose daily commute is by car. A total of 65 experimental participants were recruited. The sample size for the two driving groups met the minimum sample size requirements (Ahmad and Halim, 2017), which is calculated as:

Table 1
Summary of participants in the driving simulator study.

Variable description	Professional		Non-professional	
	Mean	S.D.	Mean	S.D.
Years holding a full driving license	18.95	6.89	8.40	4.80
Approximate driving distance annually (10^3 km)	52.35	12.21	18.88	7.02
Number of accidents	1.37	2.89	1.12	2.34
Age				
≤25	–		9.7 %	
26–35	5.9 %		58.1 %	
36–45	52.9 %		29.0 %	
46–55	35.3 %		3.2 %	
≥55	5.9 %		–	
Number of participants	34		31	

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times (\sigma_{pro}^2 + \sigma_{nonpro}^2)}{E^2} \quad (1)$$

where α represents the significance level, β represents the power of the test (the value of Z is determined by α and β), σ represents the standard deviation, and E represents the difference between the means of the two groups.

The recruitment criteria included an official (Chinese) driver's license and (self-proclaimed) good health and eyesight. Professional drivers are all road transport practitioners in China, and all have professional qualification certificates. Before the experiment, each participant read and signed a consent form that described the study and its procedures. The consent form also warned the participants of the potential for motion sickness and stated that they could withdraw from the study at any time if they experienced motion sickness or for any other reason. All the participants completed a 5-min to 10-min warm-up drive in the simulator to familiarize themselves with the controls and the virtual driving environment. After completing all the driving scenarios, they were asked to fill a simulated sickness questionnaire. The questionnaire asked them to rate the severity of each symptom of motion sickness on a scale of none, mild, moderate, and severe. One participant reported moderate motion sickness, and therefore it was withdrawn from the study. Each participant was provided with a monetary reward (200 Yuan) for completing the experiment.

2.4. Driving scenario and test procedures

A layout consistent with a real urban environment was designed in the driving simulator to provide a relatively realistic traffic conditions, as shown in Fig. 4.b. The layout contained multiple signalized intersections. Each leg of the intersection had a two-way four-lane with a lane width of 3.75 m. Fig. 4.a illustrates the city scene designed in the simulator and the numbering of each intersection. Fig. 4.b is the perspective of a driver entering an intersection. Fig. 4.c shows the software GUI for designing the traffic scene and traffic flow in the simulator. The software also includes a traffic signal-independent control system, which can make each intersection traffic signal independent and unpredictable. During the experiment, the speed was limited to 60 km/h. Note that the driver's behavior is often the result of interaction with traffic environment information. However, this paper focuses on the impact of the connected information on the driving behavior. A low traffic flow is set in the traffic scene to prevent the driver's interaction with surrounding vehicles from affecting his/her response to connected information (Chen et al., 2021).

Each participant performed the following driving tasks:

- Baseline (Traditional environment): In this scenario, no connected information was provided, and the driver could only perceive the surrounding traffic and infrastructure information through vision and hearing, which is consistent with traditional driving. The system was reset each time the experiment started to avoid possible learning effects. The test was performed three times, and the paths were 1-2-3-14-15-4-7-8-3-2-1, 1-9-10-1-12-13-14-3-8-7-4, and 1-10-9-8-7-4-15-14-3-2-1, as shown in Fig. 4.a.
- HMI Scenario (Connected environment): The driving process was conducted in a connected environment, and the driver obtained traffic light information through the HMI. When the vehicle was 300 m from the intersection, the HMI function was activated, and the connection information was displayed in real-time. The driver only received a voice prompt when the HMI was activated to inform him/her of the traffic light information at the current intersection. In a complete traffic light cycle, the red-light, green-light, and yellow-light durations were 35, 25, and 2 s, respectively. This test was also performed three times, and the driving paths were similar to those in the baseline drive.

The communication channel was free of latency, and the information was received instantaneously upon the vehicle's arrival at the designated location (Yu et al., 2019). It is important to mention that the activation of the V2I communication 300 m before the intersection was determined by the requirements of the People's Republic of China's "Specification for Layout of Urban Road Traffic Signs and Markings" and "Specification for Layout of Highway Traffic Signs and Markings" to implement warning signs at intersections (M. o. T. o. t. P. s. R. o. China, , 2009; M. o. H. a. U.-R. D. o. t. P. s. R. o. China, , 2015). In this study, the influence of the temporal and spatial characteristics of the connection information on the driver's decision and behavior was not considered. The state of the vehicle entering the intersection may have been accelerating, decelerating, or passing at a constant speed, or it may have been stopped. The experiment lasted four weeks, and a total of 65 drivers completed it. A total of 820 (426 for professional and 394 for non-professional drivers) and 860 (440 for professional and 420 for non-professional drivers) valid intersection-approaching maneuvers were extracted

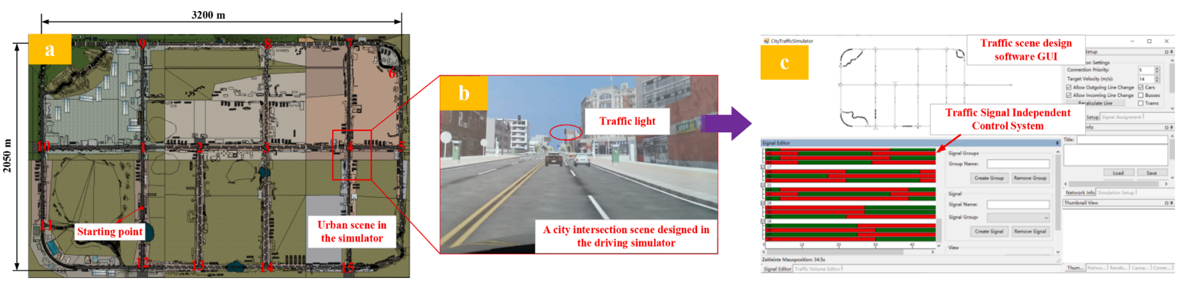


Fig. 4. The top view of the city scene in the simulator (a), one of the intersections (b), the scene design software GUI (c).

in the baseline and HMI scenarios, respectively.

3. Driving performance indicators and statistical methods

Before introducing and defining each variable, it is necessary to analyze the driver's behavior while approaching the intersection, as shown in Fig. 5. There are three stages of the intersection-approaching behavior in a connected environment (cf. the three-color bars). (i) After the HMI has been activated at the designated location, the driver does not immediately interact with it, and there is no behavioral response at this stage (i.e., attention lag). (ii) The driver starts to look at the HMI (gaze response), but there is still no response performance (i.e., perceptual interpretation). Finally, (iii) the driver frequently looks at the HMI and continuously controls the vehicle based on the traffic light information (i.e., full response). Therefore, various driving behavior indicators, such as the visual interaction behavior, response characteristics, driving performance, and behavior patterns, were analyzed to reveal the response mechanism of the professional and non-professional drivers to the connected environment when approaching the intersection.

3.1. Extraction of the visual behavior of the driver

The driver's visual response to the HMI can be recorded by capturing the eye movement by the eye tracker, which is represented by two measurement variables: the average number of gazes (ANG) on the HMI and the average cumulative percentage of the gaze time on the HMI (ACPG) (Maag et al., 2022).

According to the layout of the cabin, the field of vision of the driver is mainly divided into the front road area, HMI, left rearview mirror, right rearview mirror, dashboard, and other areas of interest (AOI), as shown in Fig. 6.a. The gaze behavior of the driver was extracted using a hierarchical clustering method, which has proven to be effective in the previous gaze behavior clustering studies presented in (Iwatsuki et al., 2016; Sato et al.). Note that the latter have also proved the effectiveness of the clustering algorithm. Fig. 6.b shows an example of the eye gaze direction (left in 1st row) and the gaze heading (right in 1st row). In the 2nd row of Fig. 6, cluster 1–5 represent the driver looking at the HMI, observing the right rearview mirror, looking forward, viewing the dashboard, and switching between gazing ahead and observing the right rearview mirror process. The frequency and accumulated time of drivers looking at the HMI were obtained according to the clustering results, as shown in the right figure in the 2nd and 3rd rows of Fig. 6. The inverted triangle in the 3rd row of Fig. 6 is the number of gazes extracted by the *findpeaks* function in MATLAB, and the blue line represents the average cumulative percentage of the gaze time on the HMI.

3.2. Response characteristics

In contrast to the existing studies that provide early warning information to drivers through the HMI (Yan et al.,Feb, 2015), the connection information in this study can be considered as a support for the cooperative driving (Liu and Kircher, 2018; Maag et al., 2022; Kraft et al., 2020). Therefore, the response of the driver cannot be solely represented by the braking response time (Chen et al., 2021;Yan et al.,Feb, 2015). In this study, two measures were defined to reflect the response characteristics of the driver: the first interaction time and the response time. As the name implies, the first interaction time is the time elapsed between the activation of the

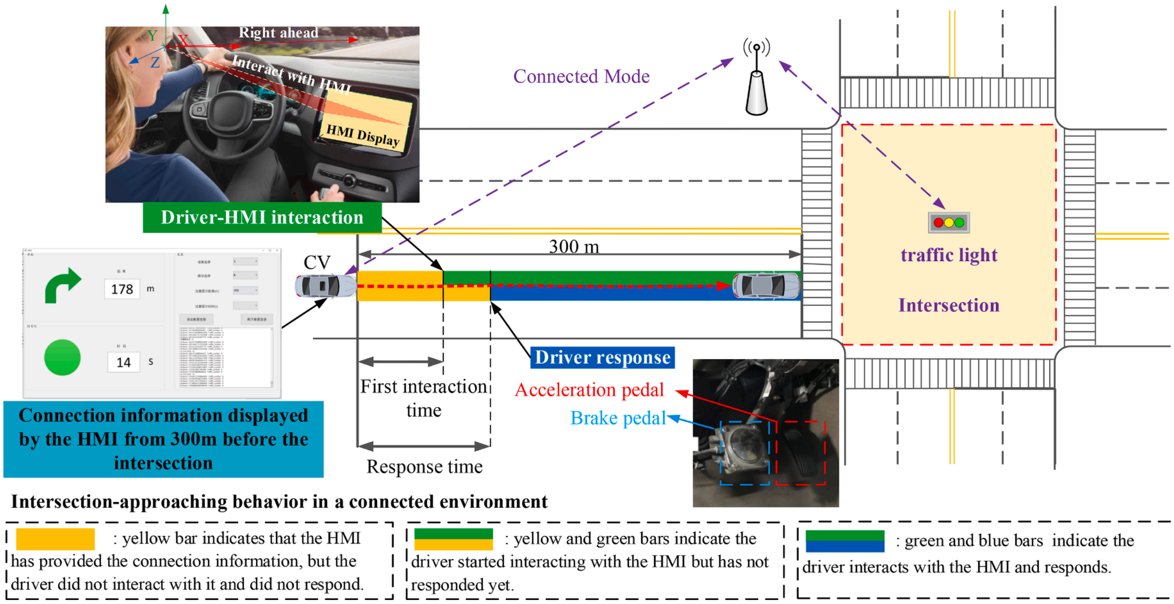


Fig. 5. Schematic diagram of intersection-approaching behavior.

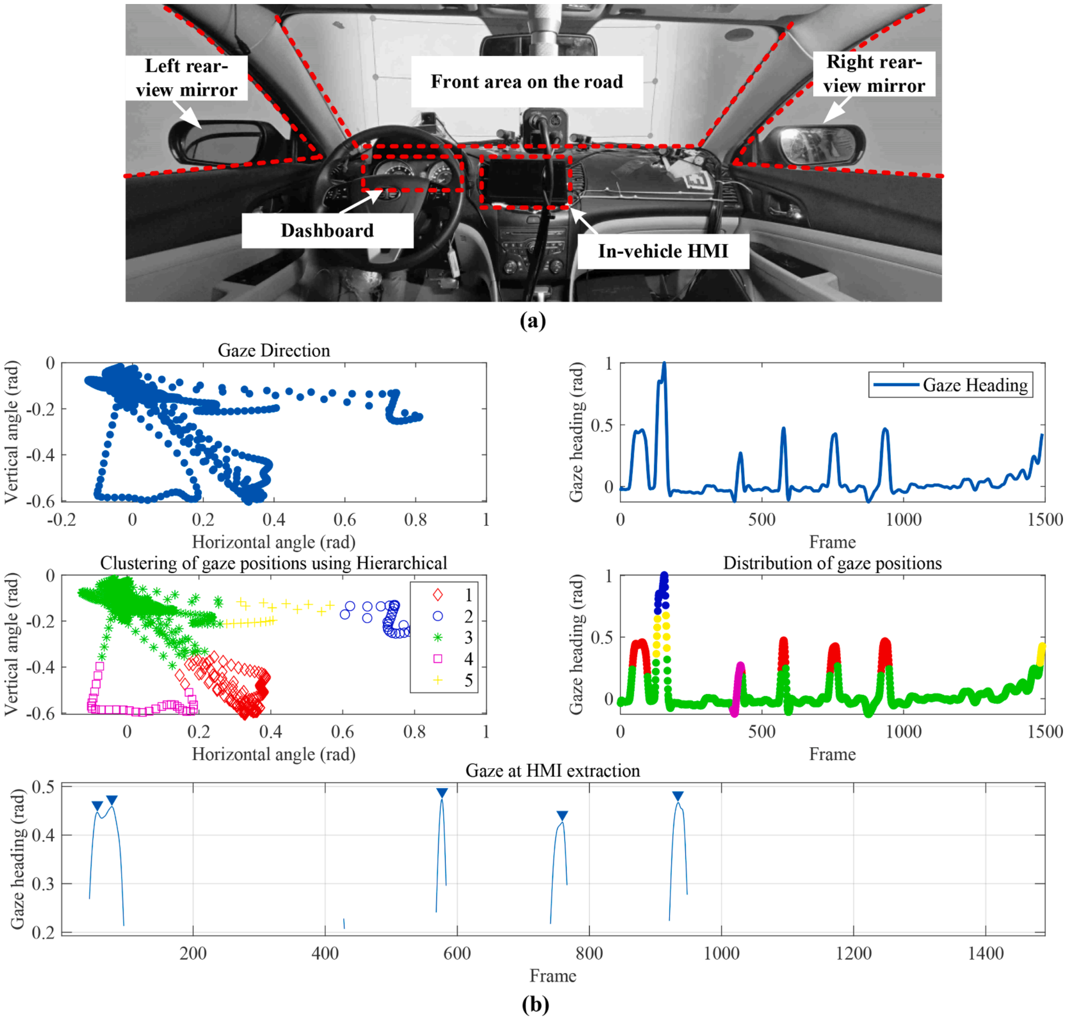


Fig. 6. AOIs of driver's field of view in cabin (a) and example of gaze data extraction (b).

HMI and the first time the driver looks at the HMI, while the response time is defined as the time elapsed between the HMI starting to issue connected information and the driver starting to control the speed of the vehicle (Sharma et al.,Apr, 2019).

The driver's response in traffic is essentially a human response to stimuli. In general, the response time can be estimated by monitoring the lateral and longitudinal movement of the vehicle. For example, Sullivan et al. (Sullivan et al., 2008.) defined the lateral response as $\pm$ 3°/s of the steering movement. In general, the longitudinal response is reflected in the braking behavior (Zhang et al., 2019). When a driver approaches an intersection, the connection information can reduce the unnecessary accelerations and decelerations. Therefore, the speed profile can reflect the driver's response to the connection information. However, the change of the vehicle speed is the result of the driver's foot pressing on the brake pedal or accelerator. Thus, in time, the changes of the vehicle speed lag behind the driver's action of depressing the brake pedal or accelerator. Therefore, to accurately reflect the response characteristics of the driver after receiving the HMI information, the fluctuation of the brake and accelerator pedals should be directly monitored instead of monitoring the speed profile. Zheng and Washington (Zheng and Washington, 2012) comprehensively studied the applicability of different wavelets for analyzing traffic data. They deduced that the Mexican hat wavelets had a satisfactory performance for detecting traffic state change points, starting points of acceleration (deceleration) waves, and discontinuities in the fundamental diagram. Therefore, this study also used the Mexican hat wavelets to detect the significant change points in the brake and throttle profiles in order to estimate the response time.

A wavelet transform coefficient of a continuous signal $x(t)$ is referred to as a continuous wavelet transform, which is expressed as:

$$T(a,b) = w(a)\int_{-\infty}^{\infty} x(t)\psi\frac{t-b}{a}dt \tag{2}$$

where a is a scale parameter that governs the dilation and contraction of the wavelet, b is a translation parameter that governs the movement of the wavelet in the time dimension, and $w(a)$ is a weighting function typically set to $1/\sqrt{a}$ to ensure that the wavelets at all the scales have the same energy.

When $a = 1$ and $b = 0$, $\psi(t)$ is referred to as the mother wavelet. The Mexican hat wavelet is defined as:

$$\psi\frac{t-b}{a} = \frac{2}{\sqrt{3}\pi^{1/4}}\left[1-\left(\frac{t-b}{a}\right)^2\right]e^{-\frac{\left(\frac{t-b}{a}\right)^2}{2}} \tag{3}$$

In this study, the speed time series $v(t)$ is a continuous signal function. The wavelet transform coefficient of $v(t)$ can be obtained by substituting Eq. (3) into Eq. (2):

$$T(a,b) = \frac{2}{\sqrt{3a}\pi^{1/4}}\int_{-\infty}^{\infty} v(t)\left[1-\left(\frac{t-b}{a}\right)^2\right]e^{-\frac{\left(\frac{t-b}{a}\right)^2}{2}}dt \tag{4}$$

Zheng et al. (Zheng et al., 2011) demonstrated that the temporal distribution of the wavelet-based energy E_b can be used to identify significant signal changes. The average wavelet-based energy at b is computed based on the wavelet transform coefficients for different scales as follows:

$$E_b = \frac{1}{\max(a)}\int_0^{\infty}|T(a,b)|^2 da \tag{5}$$

Eq. (5) indicates that an abrupt change in the signal can cause a rapid increase in the time distribution of the wavelet-based energy. Therefore, the energy distribution can be used to identify significant signal changes attributed to the onset and clearance of the queues and to the arrivals of the oscillation waves.

It is important to mention that when the driver observes the HMI for the first time, the following three situations exist in its foot-pressing behavior: (i) his/her right foot is just on the accelerator; (ii) on the brake pedal; and (iii) on neither the accelerator nor the brake pedal. For the first two cases (which make up the majority of the sample), it is efficient to monitor the point of significant change in the accelerator or brake pedal after the first look at the HMI, as shown in the examples of Fig. 8.a and b). Various hypotheses may exist for the third case. For example, the driver may need a second observation to make a reasonable decision (Caird et al., 2007), or the current speed is already as expected without further response, or the driver may have a relatively long decision time. In these cases, an accurate response time cannot be estimated due to the inability to obtain an observable activity. The third case accounted for 3.72 % of the total sample (32 out of 860), which occurred 14 times among non-professional drivers and 18 times among professional drivers. These samples will be removed from the response characteristic analysis. Therefore, in this study, the response time is defined as the time when the first peak appears in the wavelet-based energy distribution calculated based on the brake and throttle profiles after the first interaction time. Fig. 7 presents two examples corresponding to the first time the driver looks at the HMI with his/her foot on the accelerator (a) and on the brake pedal (b).

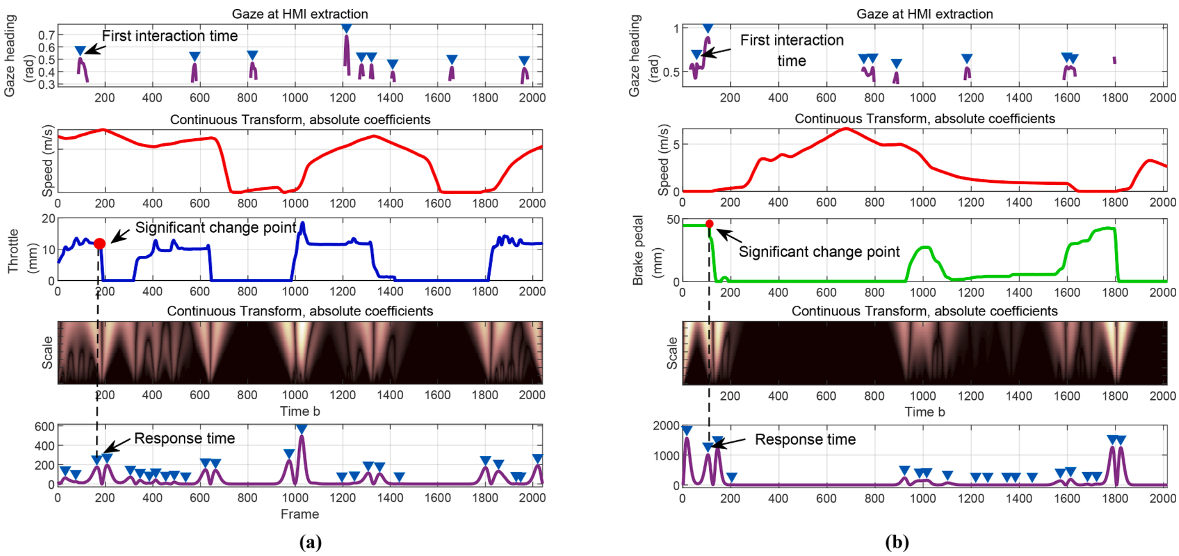


Fig. 7. Examples of the drivers' first interaction times (top of Fig. a and b) and response times on throttle (a) and on brake pedal (b). Response times were localized to the time of significant change in foot-pedaling behavior after the first interaction with the HMI.

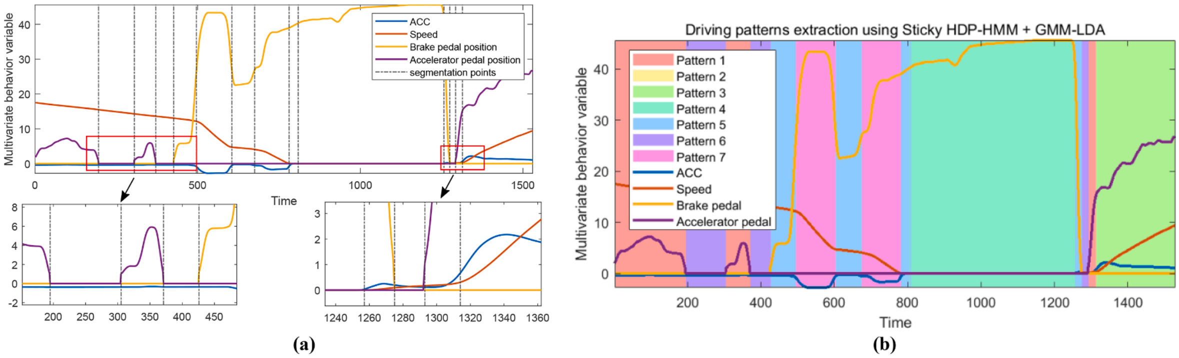


Fig. 8. Example of time series segmentation of intersection-approaching behavior (a). Example of driving pattern clustering using GMM-LDA (b).

3.3. Driving performance response and nonparametric statistical methods

Referring to the definition of the driving performance provided by Evans (Evans, 1991), this study uses the standard deviation of lateral position (SDLP), standard deviation of speed (SDS), average speed (AS), average acceleration (AA), standard deviation of braking force (SDBF), and average and standard deviation of accelerator pedal position (ATO and SDTO) to reflect the driving performance. The SDLP and SDS were used to assess the lateral and longitudinal controls, respectively (Li et al., 2016). Moreover, the AS was used to reflect the speed strategy of the driver when entering the intersection. AA, SDBF, ATO, and SDTO were used to evaluate the effectiveness of the HMI in improving the drivers' deceleration and acceleration performance while approaching the intersections. The difference between the driving performances of the two types of drivers was determined by comparing these parameters.

Various statistical analysis techniques were used to analyze the influence of the connection information on the driving behavior of the professional and non-professional drivers and to compare their differences. Before the analysis of the variance, the assumptions of parametric tests of the data were checked, and it was deduced that all the driving and visual behavior variables violate the normal distribution assumption. Therefore, Friedman tests were conducted to test the significance of the difference between the means of the groups for different traffic light phases. Mann-Whitney U-tests were also conducted for the pairwise comparisons of a parameter between two driver types. The Fisher's exact test has no restrictions on the minimum number of frequencies of a variable, and it can be used to study the statistical associations between two categorical variables. Wilcoxon signed-rank tests were conducted for the pairwise comparison of driving performance variables between the two driving environments. The Kruskal-Wallis test method was also adopted to test the probability distribution of the behavior variables of different driving patterns. Note that a 5 % significance level was assumed in these statistical analyses.

3.4. Driving pattern extraction model

The effect analysis on the driving performance can only initially obtain the qualitative results of the impact of the connected environment on the driving behavior. However, it cannot ensure the quantitative impact of the environmental changes on the driving behavior. The descriptive extraction of driving patterns can provide a deep understanding of the behavior characteristics of the driver in the two scenarios and clarify its response mechanism in the connected environment.

Chen et al. (Li et al., 2021; Chen et al., 2021) employed an unsupervised learning approach to extract behavioral patterns of lane-changing and car-following. Multivariate behavior sequences were first segmented using a Bayesian approach, and then the latter were clustered using an extended latent Dirichlet allocation (LDA) model to obtain the spatiotemporal distribution of driving patterns. The study presented in (Zhang and Wang, 2019) also proved that the Bayesian non-parametric method can be used to effectively segment the vehicle trajectory in order to obtain the basic behavioral motion unit. Based on this, this paper proposes a multivariate behavior sequence segmentation method relying on the Sticky Hierarchical Dirichlet process-hidden Markov model (Sticky HDP-HMM). This method can consider the uncertainty of multivariate observations, and automatically decompose the sequence into behavioral units without prior knowledge.

The sticky HDP-HMM generative process can be written as:

$$\begin{aligned} &\beta|\gamma \sim \mathrm{GEM}(\gamma), \\ &\pi_i|\alpha,\beta,\gamma \overset{iid}{\sim} \mathrm{DP}\left(\alpha+\kappa, \frac{\alpha\beta+\kappa\delta_i}{\alpha+\kappa}\right), i=1,2,...k,..., \\ &\theta_i \overset{iid}{\sim} H, i=1,2,...k,... \\ &z_t \sim \pi_{z_{t-1}} \\ &y_t \sim f(\theta_{x_t}), t=1,2,\dots,T \end{aligned} \tag{6}$$

where GEM denotes a sticky breaking process, H is the base probability measure, $\alpha > 0$ represents the hyperparameter for the prior distributions on state-transition parameters, β represents the weight of the given probability measure, γ represents the concentration parameter, $\mathrm{DP}(\gamma, H)$ denotes a Dirichlet process, f denotes an observation distribution parameterized by draws from H, δ_i denotes an indicator function, $\kappa \in [0,1]$ is the sticky parameter, π_k is the transition distribution such that $\sum_{k=1}^{\infty}\pi_k = 1$, λ is the distribution parameter, θ represents independently distributed random variables, $z_1, z_2, z_3, \cdots, z_T$ represent each hidden state, and $y_1, y_2, y_3, \cdots, y_T$ are observations of different states.

In HMM, the number of states is limited and should be determined in advance. HDP-HMM solves the limitation of the state in HMM. The sticky HDP-HMM augments the HDP-HMM with an additional parameter $\kappa > 0$, that biases the process toward self-transitions, which provides a way to encourage longer state durations (Johnson and Willsky, 2013). More details on the theory and modeling process are presented in (Fox et al., 2011). In this paper, to accurately reflect the intersection-approaching behavior pattern of the driver, the input variables of the driving pattern extraction model are longitudinal behavior characteristics, which are the acceleration, speed, and position of the brake pedal and accelerator pedal.

Fig. 8 (a) shows an example of using Sticky HDP-HMM to segment a multivariate time series of an intersection-approaching process, where the dotted lines indicate the split positions. It can be seen from the partially enlarged picture that the model has a good effect on the time series segmentation, and every state transition of behavior can be detected.

In this paper, each segment, behavior pattern, and a sample of the intersection-approaching behavior sequence are considered as documents, topics, and words of LDA, respectively. However, there is a discrepancy between the behavioral samples of continuous vectors on the timeline and raw LDA model inputs (discrete words). Therefore, after obtaining behavior sequence fragments, a driving pattern clustering method which combines the Gaussian mixture model (GMM) with LDA is proposed to cluster behavior samples in order to find the behavior patterns that exist in each sequence fragment. The detailed process of theoretical modeling of GMM-LDA can be found in (Prabhudesai et al.). The model first uses GMM to cluster the behavior observation samples in the sequence fragments, and clusters the vector samples into discrete behavior words. LDA then analyzes the topics (i.e., driving patterns) in each behavior segment to obtain the distribution of driving patterns in the entire sequence.

Fig. 8 (b) shows the clustering effect of GMM-LDA on the driving patterns when seven behavior patterns are defined. Different colors are used to represent different behavior patterns. More precisely, the red, yellow, green light, blue-green, dark blue, and magenta colors represent driving patterns 1–6, respectively. It can be seen from Fig. 8 (b) that LDA can effectively extract different pattern distributions in time series based on the clustering results of GMM. It is shown that the proposed behavior sequence segmentation and pattern clustering method can perform the quantitative extraction and description of driving patterns, and they can also accurately obtain the temporal distribution of each behavior pattern.

It is important to mention that the above sequence segmentation and pattern clustering are just examples. In order to accurately describe the driving patterns of the intersection approaching process, it is necessary to use quantitative indicators to evaluate the model for selecting the optimal number of patterns. In this paper, the sequence fragments obtained by sticky HDP-HMM are used as the basic unit of behavioral pattern analysis, and it is expected that the use of GMM-LDA can ensure that the pattern in each unit is as single as possible for clear description. Therefore, the entropy and perplexity, that are the most commonly adopted quantitative indicators, are used to evaluate the model performance. In information theory, the entropy is used to assess the inherent uncertainty in the possible output of a probabilistic model. More precisely, a lower entropy indicates a higher model performance. The perplexity is used to evaluate the prediction effect of the probability model for the sample. A lower perplexity indicates that the clustering algorithm is superior (Chatzisavvas et al., 2005). The entropy and perplexity are calculated as:

$$Entropy(\mathscr{D}) = \sum_{i=1}^{m}\left(-\sum_{k=1}^{K} p(k)\log p(k)\right) \tag{7}$$

$$Perplexity(\mathscr{D}) = \exp\left(-\sum_{i=1}^{m}\log p(w_i) / \sum_{i=1}^{m} N_i\right) \tag{8}$$

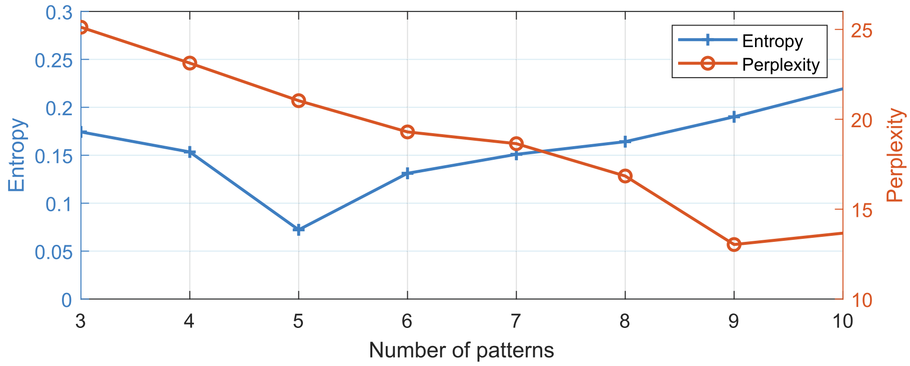


Fig. 9. Entropy and perplexity values under different numbers of patterns.

where $Entropy(\mathscr{D})$ is the entropy in sequence $\mathscr{D}$, $p(k)$ is the probability distribution of pattern k, $Perplexity(\mathscr{D})$ represents the perplexity of $\mathscr{D}$, and $\log p(w_i)$ denotes the log-likelihood of behavior segment i which indicates the number of samples in it.

The number of behavior patterns (k) is set in the range of 3–10, and tests are performed to select its optimal value. Different k values are considered for training on the behavior samples of the two driving environments, as shown in Fig. 9, where each point represents the average value of the two indicators under the same number of patterns. Finally, based on the trade-off between the two metrics, k is set to 7.

4. Results

4.1. Descriptive statistics of the driver's visual behavior

During the HMI driving scenario, the phase of the traffic light seen by the driver through the HMI may be red, green, or yellow, which could cause the drivers to behave differently. Therefore, it is necessary to analyze the driver interaction process and response behavior under different traffic light phases. In this study, in a completed traffic light cycle, the durations of the red, green, and yellow lights are 35, 25, and 2 s, respectively. For convenient statistical analysis, a time interval of 5 s was set. Thus, the red and green lights were respectively divided into 7 and 5 intervals, and the yellow light had only one interval. The interaction characteristics of the drivers were analyzed with the HMI in different traffic light phases and for different countdown times. In Fig. 10(a) and (b), 'Red 1-5′ on the x-axis indicates that the red light is displayed when the HMI is activated and there are 1 s to 5 s remaining from the phase transition. Similarly, 'Green 1–5′ indicates that the current traffic light is green and there are 1 s to 5 s left before the phase transition. Since there is only one interval for the yellow light, 'Yellow' indicates that the HMI displays a yellow light when it is activated. Fig. 10 (a) and (b) respectively show the average number of gazes on the HMI and the average cumulative percentage of the gaze time on the HMI of the professional and non-professional drivers for different traffic light phases and remaining time. Fig. 10(c) and (d) show the within-group and between-group tests of ANG and ACPG under different traffic light phases, where 'Red (i.e., Yellow, Green) pro' and

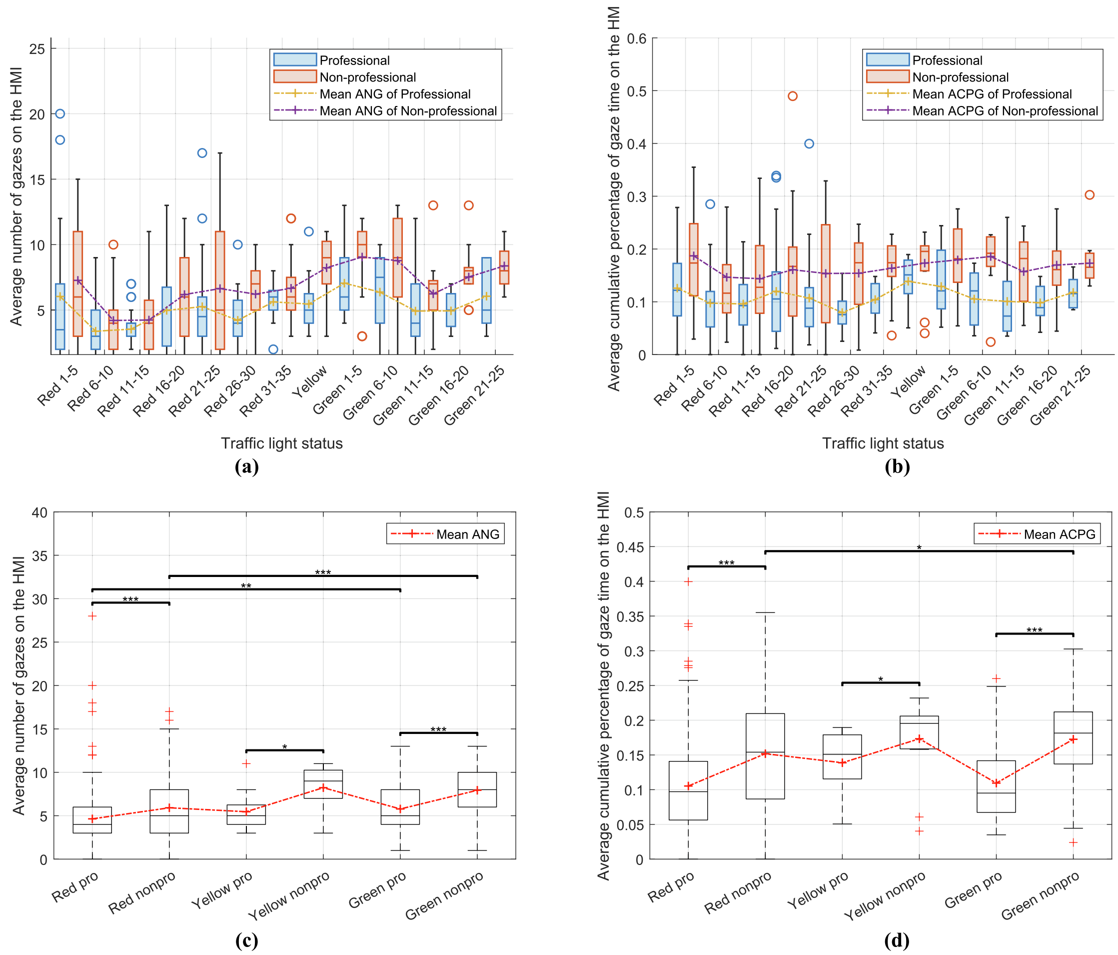


Fig. 10. Average number of gazes on the HMI (a) and average cumulative percentage of gaze time on the HMI (b), between- and within-group testing of ANG (c) and ACPG (d) under different traffic light phases. The dotted line in the figure represents the mean value.

’Red (i.e., Yellow, Green) nonpro’ represent the visual performance statistics of professional and non-professional drivers at the red (i.e., Yellow, Green) light phase, respectively. Table 2 shows descriptive statistics of ANG and ACPG in each time interval for the two driver groups and comparisons within and between groups using the non-parametric Friedman and Mann-Whitney U- tests. There is no driver-HMI interaction process in the baseline driving scnario, and thus no visual behavior statistics are performed. In Table 2, the ’Red light phase’ and ’Green light phase ’ respectively represent the ANG and ACPG under the red and green phases, and ’All average’ indicates the ANG and ACPG in connected driving scenarios for the two groups of drivers.

Cohen’s d is used as the effect size indicator, which is calculated as:

$$d = \frac{m_1 - m_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2}}} \tag{9}$$

where m_1 and m_2 respectively represent the average value of the two groups of data, n_1 and n_2 are the sample sizes of the two data, s_1 and s_2 are the standard deviations of the two groups of samples.

It can be seen from Fig. 10(a) and (c) and Table 2 that the ANG values of the professional and non-professional drivers in the HMI scene were respectively 5.000 and 6.584, which presents a significant difference ($p < 0.001$). There was a significant difference in the ANG between the professional and non-professional drivers in all the three traffic light phases ($p_{red} < 0.001$, $p_{yellow} = 0.016$, $p_{green} < 0.001$), while the largest difference (Cohen’s d = 1.125) is in the yellow phase. In addition, it can be observed from the results of the within-group test that the professional and non-professional drivers had higher ANG during the green phase compared with the red phase (Fig. 10 (c)). This indicates that drivers interact more frequently when the HMI displays a green light. The ANG values of the professional and non-professional drivers in the yellow light phase were respectively 5.462 and 8.231, and no significant difference was observed between it and the other two phases. It can also be seen from the ACPG statistics and test results in Fig. 10(b) and (d) and Table 2 that there were significant differences between different traffic phases and intervals ($p = 0.023$, $\chi^2 = 23.62$ for professional drivers and $p = 0.034$, $\chi^2 = 22.27$ for non-professional drivers). The ACPG values of the professional and non-professional drivers were respectively 0.108 and 0.162, which presents a significant difference ($p < 0.001$, Cohen’s d = 0.692). In addition, under different traffic phases and intervals, the ACPG of the non-professional drivers was significantly higher than that of the professional drivers. This indicates that the non-professional drivers allocate more attention to HMI in HMI scenarios.

4.2. Response characteristics

The response time and first interaction time are important variables that describe the driver’s cognition of assisted driving information in a connected environment. Fig. 11 and Table 3 present the histograms and descriptive statistics of the response time and first interaction time of the professional and non-professional drivers. The distribution of response time and first interaction time was also fitted to obtain its distribution law, as shown in Fig. 11.

In addition, the driver’s response time follows the logarithmic normal distribution ($\mu = 1.064$, $\sigma = 0.691$ for professional drivers, $\mu = 1.197$, $\sigma = 0.674$ for non-professional drivers). The first interaction time follows an exponential distribution ($\mu = 2.120$ for professional drivers and $\mu = 2.829$ for non-professional drivers). Moreover, the response rates of the professional and non-professional drivers were 98.15 % and 98.01 %, respectively. A correlation analysis shows that the earlier the driver interacts with the HMI, the

Table 2
Descriptive statistics of ANG and ACPG, Mann-Whiney *U* test between the driver groups and Friedman test under different traffic light phases within the group.

Traffic light phase	ANG				Mann-Whitney *U* test		ACPG				Mann-Whitney *U* test	
	Professional		Non-professional		p	Cohen’s d	Professional		Non-professional		p	Cohen’s d
	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Red 1–5	6.038	6.654	7.269	6.654	0.010	0.437	0.126	0.071	0.187	0.071	0.003	0.959
Red 6–10	3.385	2.099	4.206	2.099	0.004	0.816	0.097	0.066	0.147	0.066	0.036	0.474
Red 11–15	3.543	1.516	4.240	1.516	0.046	0.633	0.096	0.050	0.144	0.050	0.002	0.822
Red 16–20	4.977	2.849	6.186	2.849	0.002	0.696	0.120	0.082	0.161	0.082	0.005	0.598
Red 21–25	5.262	3.108	6.643	3.108	0.041	0.580	0.107	0.073	0.154	0.073	0.025	0.630
Red 26–30	4.217	2.066	6.217	2.066	<0.001	1.213	0.080	0.032	0.154	0.032	<0.001	1.462
Red 31–35	5.625	1.455	6.667	1.455	0.011	0.965	0.104	0.032	0.164	0.032	0.001	1.507
Red light phase	4.640	3.276	5.916	3.719	<0.001	0.364	0.105	0.065	0.157	0.103	<0.001	0.600
Yellow	5.462	2.184	8.231	2.713	0.016	1.125	0.139	0.045	0.173	0.058	0.012	0.662
Green 1–5	7.056	3.115	9.056	2.623	0.001	1.462	0.129	0.060	0.179	0.067	0.001	1.315
Green 6–10	6.375	2.895	8.778	3.370	0.002	1.429	0.105	0.051	0.186	0.048	<0.001	2.299
Green 11–15	4.926	2.934	6.238	2.468	0.001	1.074	0.101	0.072	0.157	0.065	<0.001	1.254
Green 16–20	4.941	1.519	7.529	2.478	<0.001	1.990	0.098	0.031	0.170	0.062	<0.001	1.936
Green 21–25	6.059	2.331	8.375	1.668	<0.001	2.108	0.118	0.027	0.173	0.041	<0.001	2.387
Green light phase	5.779	2.741	7.933	2.751	<0.001	0.785	0.110	0.054	0.173	0.057	<0.001	1.133
All average	5.000	3.130	6.584	3.559	<0.001	0.472	0.108	0.061	0.162	0.092	<0.001	0.692
Friedman test	df	12	12	–	12	12	–					
	χ^2	56.51	54.79		23.62	22.27						
	p	**<0.001**	**<0.001**		**0.023**	**0.034**						

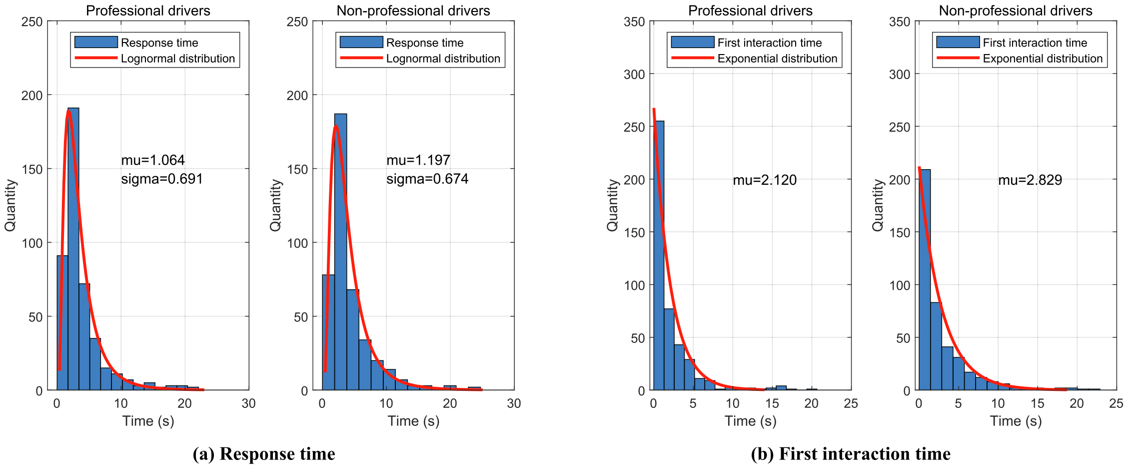


Fig. 11. Histograms and distribution law of the driver response time (a) and first interaction time (b) for professional and non-professional drivers.

Table 3
Descriptive statistics of the response time and first interaction time (s) and Mann-Whiney *U* test between driver groups.

Statistical indicators	Response time		Mann-Whitney *U* test		First interaction time		Mann-Whitney *U* test	
	Professional	Non-professional	P	Cohen's d	Professional	Non-professional	P	Cohen's d
Mean	3.787	4.223	0.001	0.126	2.120	2.829	<0.001	0.226
Standard Deviation	3.439	3.496			2.789	3.475		
Variance	11.825	12.222			7.780	12.074		
Median	2.667	3.072			1.050	1.442		
Mode	2.033	2.100			0.617	1.500		
Maximum	22.017	24.363			20.550	22.867		
Minimum	0.461	0.500			0.010	0.026		

earlier the response occurs ($R = 0.92$, $p < 0.001$ for professional drivers and $R = 0.96$, $p < 0.001$ for non-professional drivers).

It can be deduced from the statistical results in Table 3 that the average time for the professional drivers to gaze at the HMI for the first time was 2.12 s (median of 1.05 s and standard deviation of 2.789 s), and the average response time was 3.787 s (median of 2.667 s and standard deviation of 3.439 s). As for the non-professional drivers, the first interaction time was 2.829 s (median of 1.442 s and standard deviation of 3.475 s), and the average response time was 4.223 s (median of 3.072 s and standard deviation of 3.496 s). The Mann-Whitney *U* test showed significant differences in the response time and first interaction time between the professional and non-professional drivers ($p = 0.001$, Cohen's $d = 0.126$ for professional drivers and $p < 0.001$, Cohen's $d = 0.226$ for non-professional drivers). In addition, Friedman tests were performed on the response time and first interaction time under different signal light phases, and the obtained results showed no significant difference ($p > 0.1$, $\chi^2 = 6.31$ for professional drivers, $p > 0.1$, $\chi^2 = 4.28$ for non-professional drivers). This indicates that the time for the driver to interact with the HMI and generate the response behavior is less affected by the traffic light information.

4.3. Driving performance

4.3.1. Statistics of drivers passing the intersections without stopping

The connected environment should be able to reduce the unnecessary stops at intersections (Ilgin Guler et al., 2014). Table 4 presents the statistics of two groups of drivers passing the intersection without stopping (PIWS) in the two driving scenarios, and the results of the Fisher's exact test for professional and non-professional drivers' PIWS.

It can be seen from Table 4 that there were 444 passes and 376 stops in the Baseline group, with 233 passes and 193 stops for

Table 4
Frequency and Fisher's exact test of two groups of drivers passing intersections without stopping in HMI and baseline scenarios.

Scenario	Professional		Fisher's exact test	Non-professional		Fisher's exact test	Total		Fisher's exact test
	Pass	Stop		Pass	Stop		Pass	Stop	
HMI	293	147	$p < 0.001$	249	171	$p = 0.104$	542	318	$p < 0.001$
Baseline	233	193		211	183		444	376	

Table 5
Descriptive statistics of SDLP and SDS of the intersection-approaching behavior when passing an intersection.

	Scenario	Standard deviation of lateral position (m)				Mann-Whitney *U* test		Standard deviation of speed (m/s)				Mann-Whitney *U* test	
		Professional		Non-professional		p	Cohen's d	Professional		Non-professional		p	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	0.439	0.445	0.523	0.513	0.823	0.175	3.544	2.217	6.141	2.003	<0.001	1.229
	HMI	0.441	0.460	0.427	0.465	0.187	0.031	3.866	2.373	2.572	1.858	<0.001	0.607
Wilcoxon signed rank test	p	0.413		0.815				0.161		<0.001			
	Cohen's d	0.058		0.231				0.139		1.910			

Table 6
Descriptive statistics of SDBF and SDTO of the intersection-approaching behavior when passing an intersection.

	Scenario	Standard deviation of braking force (N)				Mann-Whitney *U* test		Standard deviation of accelerator pedal position (mm)				Mann-Whitney *U* test	
		Professional		Non-professional		p	Cohen's d	Professional		Non-professional		P	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	11.24	11.707	15.112	11.544	0.001	0.333	8.306	3.581	11.612	3.43	$<$0.001	0.943
	HMI	10.53	13.73	10.41	10.696	0.161	0.167	7.974	3.679	6.805	2.718	$<$0.001	0.509
Wilcoxon signed rank test	p	0.385		$<$0.001				0.417		$<$0.001			
	Cohen's d	0.075		0.515				0.076		1.690			

Table 7
Descriptive statistics of AS and ATO of the intersection-approaching behavior when passing an intersection.

	Scenario	Average speed (m/s)				Mann-Whitney *U* test		Average accelerator pedal position (mm)				Mann-Whitney *U* test	
		Professional		Non-professional		p	Cohen's d	Professional		Non-professional		p	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	11.559	4.452	13.746	5.705	0.007	0.427	7.264	4.091	6.334	3.238	0.033	0.323
	HMI	10.932	4.89	10.393	4.653	0.223	0.113	7.048	4.434	7.252	3.656	0.656	0.106
Wilcoxon signed rank test	p	0.012		<0.001				0.524		0.044			
	Cohen's d	0.358		0.701				0.072		0.373			

Table 8
Descriptive statistics of SDLP and SDS of the intersection-approaching behavior when stopping before an intersection.

	Scenario	Standard deviation of later position (m)				Mann-Whitney *U* test		Standard deviation of speed (m/s)				Mann-Whitney *U* test	
		Professional		Non-professional		p	Cohen's d	Professional		Non-professional		p	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	0.705	0.499	0.923	0.584	<0.001	0.401	5.025	2.329	5.199	1.578	0.155	0.087
	HMI	0.631	0.493	0.658	0.533	0.382	0.054	5.694	1.716	5.592	2.39	0.531	0.049
Wilcoxon signed rank test	p	0.331	0.005		<0.001	0.395							
	Cohen's d	0.139	0.412		0.697	0.169							

Table 9
Descriptive statistics of SDBF and SDTO of the intersection-approaching behavior when stopping before an intersection.

	Scenario	Standard deviation of braking force (N)				Mann-Whitney *U* test		Standard deviation of accelerator pedal position (mm)				Mann-Whitney *U* test	
		Professional		Non-professional		p	Cohen's d	Professional		Non-professional		p	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	20.45	12.481	26.493	3.471	<0.001	0.531	6.915	2.489	13.308	3.826	<0.001	1.981
	HMI	17.034	10.529	11.153	16.006	<0.001	0.580	6.262	2.658	6.508	2.39	0.041	0.098
Wilcoxon signed rank test	p	0.785		<0.001			0.42		<0.001				
	Cohen's d	0.086		0.981			0.042		2.180				

Table 10

Descriptive statistics of AA and ATO of the intersection-approaching behavior when stopping before an intersection.

		Average acceleration (m/s^2)				Mann-Whitney *U* test		Average accelerator pedal position (mm)				Mann-Whitney *U* test	
	Scenario	Professional		Non-professional		p	Cohen's d	Professional		Non-professional		p	Cohen's d
		Mean	SD	Mean	SD			Mean	SD	Mean	SD		
	Baseline	−0.139	0.265	−0.382	0.561	0.001	0.524	5.048	2.563	8.798	4.179	<0.001	1.082
	HMI	−0.247	0.144	−0.257	0.212	0.477	0.053	3.793	2.127	4.644	2.029	<0.001	0.41
Wilcoxon signed rank test	p	<0.001		0.012				0.001		<0.001			
	Cohen's d	0.834		0.673				0.424		1.265			

professional drivers, and 211 passes and 183 stops for non-professional drivers. In the HMI scenario, there are 542 passes and 318 stops, with 293 passes and 149 stops for professional drivers, and 249 passes and 171 stops for non-professional drivers. It was also deduced that the PIWS of all the drivers were significantly different between the two driving environments, as shown by the Fisher's exact test ($p < 0.001$). The results demonstrate that providing traffic light information to drivers in a connected environment is effective in prompting traffic efficiency at intersections. However, it can be observed from Table 4 that the PIWS of the professional drivers significantly changes from the Baseline to the HMI scenarios ($p < 0.001$), while that of the non-professional drivers was not significantly different ($p = 0.104$), as shown by the Fisher's exact test results. These results indicate that the traffic light information provided by the HMI before the intersection is more helpful for professional drivers than for non-professional drivers to efficiently pass the intersection.

4.3.2. *Difference analysis of the driving behavior*

The intersection-approach behavior in the two states of stopping when arriving at the intersection or passing through it has different driving performances. Therefore, the performance indicators of intersection-approach behavior in these two states were separately counted, and the Wilcoxon signed-rank test was conducted to analyze the performance in the two driving environments, and the Cohen's d was used as the effect size. The post-hoc analysis and inspection results of different performance indicators in the two driving environments are shown in Tables 5-10. Tables 5-7 show descriptive statistics for various driving performance variables of the intersection-approach behavior for professional and non-professional drivers while passing through the intersection. Tables 8-10 present the analysis of the intersection-approach behavior when stopping before the intersection, as well as comparisons results of each variable between the two driver groups using Mann-Whitney U tests and within the driver groups between the two environments using Wilcoxon signed-rank tests.

The lateral and longitudinal stability of the vehicles were measured by the SDLP and SDS. It can be seen from Tables 5 and 8 that there was no significant difference in SDLP of the intersection-approach behavior between the professional and non-professional drivers when passing through an intersection, nor did the Wilcoxon post-hoc test. However, from the baseline to the HMI scene, the SDLP of the intersection-approaching of non-professional drivers significantly decreased ($p = 0.005$, Cohen's $d = 0.412$) when the vehicle stopped before the intersection, while the lateral control of professional drivers was relatively stable ($p = 0.331$, Cohen's $d = 0.139$). It can be deduced from the SDS results that the professional and non-professional drivers have significantly different intersection-approach performances when driving through or stopping before intersections, as shown in Tables 5 and 8. The professional drivers' SDS from the baseline to the HMI scenarios did not show a significant difference when passing the intersection but increased by 0.67 m/s when stopping before it ($p < 0.001$, Cohen's $d = 0.697$). There was a significant decrease by 3.57 m/s in the SDS of the intersection-approaching for non-professional drivers from the baseline to the HMI ($p < 0.001$, Cohen's $d = 1.91$) when passing the intersection, but no significant difference was observed in the SDS for the non-professional drivers from the baseline to the HMI when stopping before an intersection. In addition, the Mann-Whitney U independent test results showed that the SDS of the professional and non-professional drivers had significant differences in the two driving scenarios when passing intersections ($p < 0.001$).

The dynamic characteristics of the driver's response were measured by the standard deviation of braking force and the standard deviation of accelerator pedal position, as shown in Tables 6 and 9. Whether passing or stopping before an intersection, SDBF and SDTO of the intersection-approach behavior for non-professional drivers significantly vary from the baseline to the HMI scenarios ($p < 0.001$). However, although the SDBF and SDTO of the professional driver from the baseline to the HMI scene decreased, they did not show a significant difference. More precisely, when the vehicle passed through the intersection, the SDBF of the non-professional driver significantly decreased by 4.7 N ($p < 0.001$, Cohen's $d = 0.515$), and the SDTO significantly decreased by 4.8 mm ($p < 0.001$, Cohen's $d = 1.69$), as shown in Table 6. Moreover, when the vehicle stopped before the intersection, the SDBF of non-the professional drivers significantly decreased by 15.3 N ($p < 0.001$, Cohen's $d = 0.981$), and the SDTO significantly decreased by 6.8 mm ($p < 0.001$, Cohen's $d = 2.18$), as shown in Table 9. In addition, the results of the Mann-Whitney U between-group test showed that the SDTO of the professional and non-professional drivers was significantly different in the two driving scenarios when passing or stopping before the intersections. However, the comparison results of SDBF between the two driver groups show that this parameter was only significantly different in the baseline drive ($p = 0.001$, Cohen's $d = 0.333$), but not in the HMI scenarios ($p = 0.161$, Cohen's $d = 0.167$) when passing intersections.

Tables 7 and Table 10 show the descriptive statistics of average speed, average acceleration, and average accelerator pedal position, which can reflect the speed strategy and acceleration characteristics of the driver's intersection-approach behavior. It can be seen from Table 7 that the AS of the professional and non-professional drivers significantly decreased from the baseline to the HMI scenarios when passing intersections, with the non-professional driver showing an even greater decrease of 3.4 m/s (Cohen's $d = 0.701$). There was no significant difference in the ATO between the HMI and the baseline scenarios for professional drivers when passing intersections, but the ATO of the non-professional drivers significantly increased from the baseline to the HMI scenarios ($p = 0.044$, Cohen's $d = 0.373$). It can also be seen from Table 10 that, from the baseline to the HMI scenario, the average deceleration of the non-professional drivers was reduced by 0.112 m/s^2 when the vehicle stopped before an intersection, while that of the professional drivers was significantly increased by 0.108 m/s^2 ($p < 0.001$). However, there was a significant difference in the AA between professional and non-professional drivers in the baseline scenario ($p = 0.001$) but not in the HMI scenario ($p = 0.477$). Furthermore, when stopping before intersections, the ATO in the HMI scenario was significantly reduced from the baseline to the HMI drive for the professional ($p = 0.001$) and non-professional ($p < 0.001$) drivers, as shown in Table 10. In addition, the ATO of the professional drivers was significantly smaller than that of the non-professional drivers ($p < 0.001$) in the two scenarios.

4.3.3. Quantitative analysis of driving patterns

The driving patterns were further extracted to quantitatively analyze the impact of the connected environment on the driving behavior and to compare the differences in driving patterns between professional and non-professional drivers. The extracted driving patterns should be descriptive to intuitively describe their switching process and distribution characteristics. Descriptive statistics and non-parametric Kruskal-Wallis test were then performed on the speed, acceleration, and pedaling behavior of the seven driving patterns, as shown in Table 11 and Fig. 12. The obtained results showed significant differences in the behavioral characteristics under different driving patterns. In addition, to further demonstrate the state switching of driving patterns in different driving environments, some visualization results are shown in Figs. 13 and 14.

It can be deduced from Table 11 and Figs. 12-14 that pattern 1 has a relatively large accelerator pedal stepping behavior, but no braking behavior, and the speed is relatively high with extremely small acceleration. This shows that driving pattern 1 is in a relatively stable state. Pattern 2 has relatively large brake-pedaling behavior and relatively small accelerator-pedal fluctuations. The speed is relatively low with small acceleration, which indicates that it is a low-speed pattern under braking behavior. Pattern 3 has significant accelerator-pedaling behavior, but no braking behavior, and it has a relatively large speed and maximum acceleration, which indicates that it is an acceleration pattern. In pattern 4, there is no acceleration behavior, the position of the brake pedal is the largest, the fluctuation is small, and the speed and acceleration are both null, which indicates that it is in the stopped state. The acceleration pedals of patterns 5 and 7 are both null, and the average values of the brake pedals are 15.910 mm and 37.868 mm, respectively. Pattern 7 has the largest deceleration, and pattern 5 has a relatively small deceleration, which indicates that they are both braking patterns. However, pattern 5 is a light braking pattern, while pattern 7 is a strong braking pattern. Pattern 6 has neither acceleration nor braking behavior, and the speed is relatively high, with a small deceleration, which indicates that the current vehicle is in the coasting pattern of slow deceleration. It is important to mention that the small accelerator pedaling behavior in pattern 2 is due to fluctuations and outliers in the data, which did not affect the driving pattern description.

It can also be seen from Fig. 12 that there were significant differences in speed ($p < 0.001$) and acceleration ($p < 0.001$) under the seven driving patterns. The significant differences were also found in the accelerator pedal position ($p < 0.001$) between patterns 1 and 3 and brake pedal position between patterns 4 and 6 ($p < 0.001$).

Fig. 13(a)-(c) show the driving pattern extraction cases of 6 different remaining time and initial speed states under the green light phase. Fig. 13 (d) and Fig. 14(a) and (b) show the visualization results of driving pattern extraction at different times under the red light phase. Fig. 14(c) and (d) show the driving pattern extraction results under the yellow light phase and the baseline group, respectively. It can be seen from Figs. 13 and 14 that the driving pattern extraction model can effectively cluster intersection-approaching behaviors into seven driving patterns and cover all the intersection-approaching behavior sequences based on the four driving behavior variables. The descriptive statistics results and the visualization results of Figs. 13 and 14 show that the clustering results of the seven driving patterns are consistent with the operation of the two types of drivers.

In addition, it can be seen from Figs. 14 and 15 that when the remaining time of the green phase is less than 5 s, patterns 4 and 6 are dominant, and the switching between the driving patterns is frequent. When the remaining time of the green light phase is relatively long (greater than 20 s), the proportion of driving pattern 1 is the largest. When the remaining time of the red light phase is less than 10 s, since the driver can predict that there would be a green light when arriving at the intersection, the relatively stable driving pattern would also prevail. When the remaining time of the red light phase is greater than 25 s, the proportion of the brake pattern is significantly increased. Due to the short duration of the yellow light and the average first interaction time of the professional and non-professional drivers exceeding 2 s, drivers miss the yellow light phase and mainly observe the red light. Therefore, the driving pattern distribution at the yellow light phase is similar to that encountered at the red light. It can be seen from Fig. 14(d) that the distribution and proportion of pattern 3 increase. This is because the driver cannot obtain the current traffic information at the intersection, which results in more unnecessary acceleration behaviors.

The distribution probabilities of professional and non-professional drivers in different driving environments were further analyzed based on the characteristics and differences of the seven driver patterns, as shown in Fig. 15 and Table 12. The abscissas 'HMI Average' and 'Baseline' in Fig. 15 represent the average pattern distribution of the HMI scenario and the driving pattern distribution in the baseline group, respectively.

It can be seen from Fig. 15 and Table 12 that driving patterns 1 (stable driving), 4 (stopped state), 5 (light braking), 6 (coasting pattern), and 7 (strong braking) are the main driving patterns in the connected environment, while the probability distributions for

Table 11
Behavioral performance descriptive statistics and Kruskal-Wallis test for different driving patterns.

	Speed		Acceleration		Accelerator pedal position		Brake pedal position	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pattern 1	10.669	5.696	−0.020	0.212	10.365	5.580	0.000	0.000
Pattern 2	3.917	4.990	−0.014	0.960	5.209	6.223	30.407	31.027
Pattern 3	6.728	5.887	0.886	0.607	19.182	9.579	0.000	0.000
Pattern 4	0.000	0.001	0.000	0.002	0.000	0.000	46.084	7.845
Pattern 5	5.815	6.059	−0.206	0.212	0.000	0.000	15.910	11.362
Pattern 6	9.268	6.333	−0.260	0.157	0.000	0.000	0.000	0.000
Pattern 7	8.762	5.887	−1.983	1.298	0.000	0.000	37.868	7.829
Kruskal Wallis test	$p < 0.001$		$p < 0.001$		$p < 0.001$		$p < 0.001$	

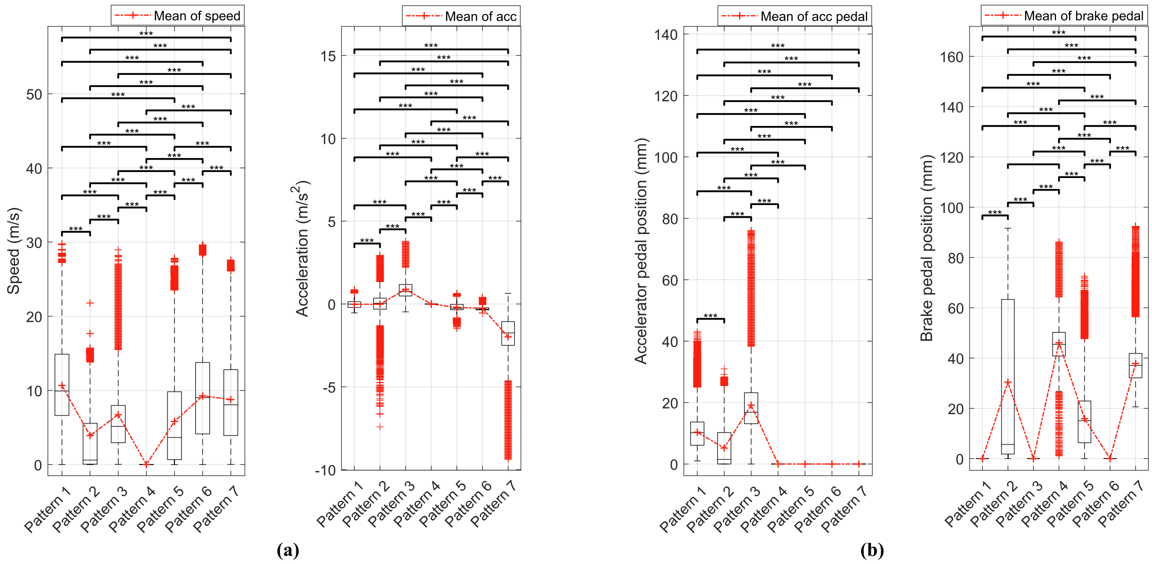


Fig. 12. Speed and acceleration boxplots (a), acceleration and brake pedal boxplots (b) and Kruskal-Wallis test results under different driving patterns.

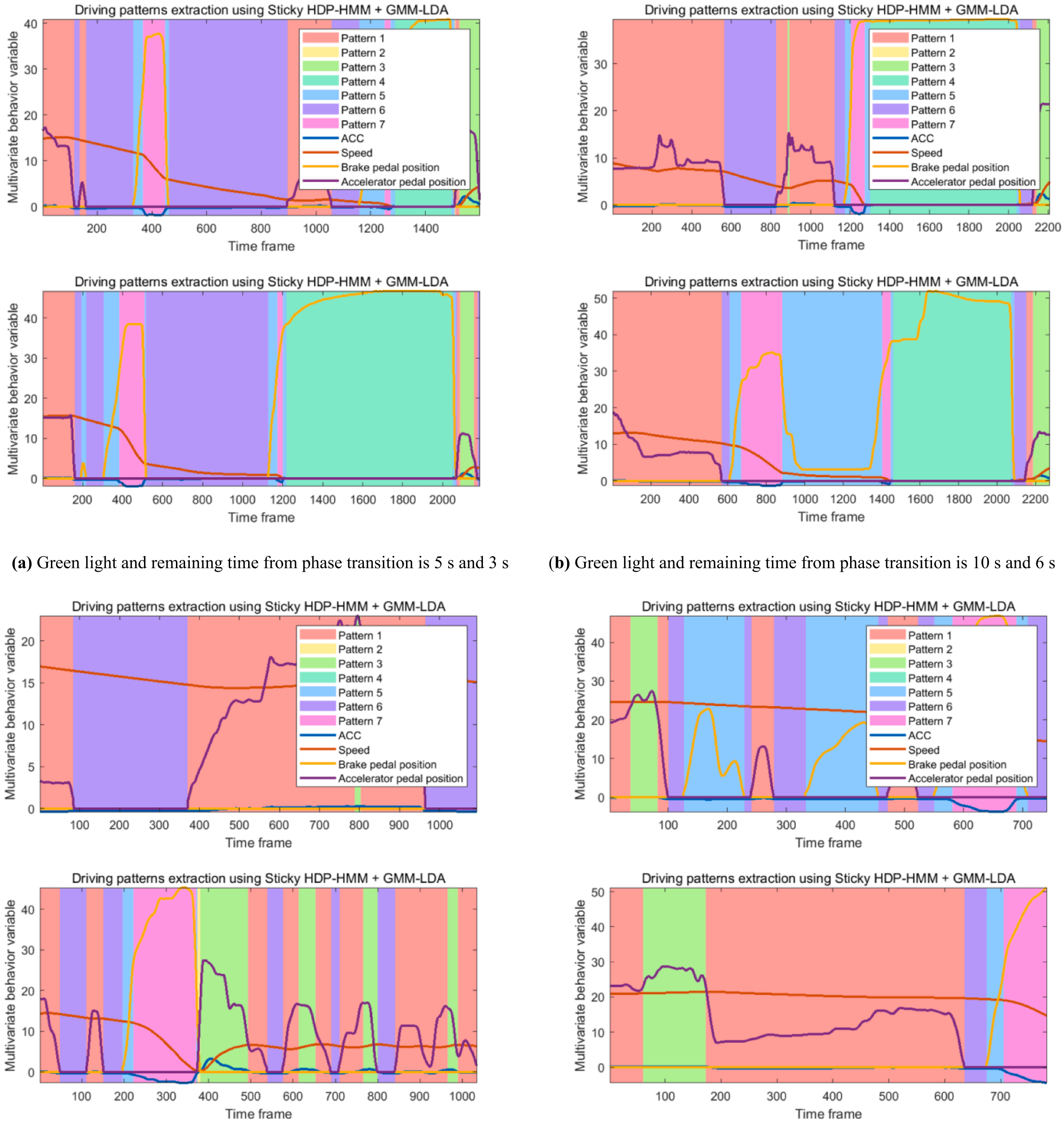
driving patterns 2 (low-speed pattern under braking) and 3 (significant acceleration) are relatively small. The average distribution probability of pattern 3 was 0.102, which indicates that there was less significant acceleration before intersections in the connected environment. However, in the traditional environment, the distribution probability of driving pattern 3 was 0.131, and the distribution probability of driving pattern 1 was 0.256, while it was 0.329 in the HMI scenario. More precisely, in the green light phase, when the remaining time increased, the distribution probability of pattern 1 gradually increased, the distribution probability of braking patterns showed a downward trend (patterns 5 and 7), and pattern 6 also showed a downward trend for the professional and non-profession drivers. In the red light phase, the opposite trend occurs. When the remaining time of the red light increased, the probability of stable driving (pattern 1) showed a downward trend and the proportion of braking (i.e., patterns 5 and 7) and coasting pattern (pattern 6) increased.

Table 12 also shows that from the baseline to the HMI scene, the probability distribution of pattern 1 increased by 0.073 (28.52 %), where the professional and non-professional drivers accounted for 0.019 (7.36 %) and 0.126 (49.61 %) of the increase, respectively. The probability of pattern 3 decreased by 0.029 (22.14 %) from the baseline group to the HMI, where the professional and non-professional drivers accounted for 0.028 (19.44 %) and 0.031 (25.83 %) of the decrease, respectively. On the contrary, compared with the baseline driving scenario, the probability distribution of pattern 6 increased in the HMI scenario by 0.035 (20.59 %): 0.023 (11.22 %) for the professional drivers and 0.04 (28.17 %) for the non-professional drivers.

5. Discussion

5.1. Frequency of interaction with HMI

This study deduced that the professional drivers looked at the HMI less frequently than the non-professional drivers when approaching an intersection in the connected environment. Correspondingly, the accumulated time of looking at the HMI was shorter for the professional drivers. This result confirms the first hypothesis of this study. More precisely, the ANG and ACPG of the non-professional drivers were 31.68 % and 50 % higher than those of the professional drivers, respectively. Under the three phases, there were significant differences in the visual behavior of the professional and non-professional drivers. The non-professional drivers allocated more attention to the interaction with the HMI. A recent study reported that older drivers performed fewer head and eye scan actions and, on average, had shorter gaze times than younger drivers. In addition, older drivers did not compensate for the head-movement component by making greater eye movements (Savage et al., 2020). Similar results were also presented in (Romoser et al., 2013) and (Romoser and Fisher, 2009). However, none of these studies investigated the visual behavior of drivers interacting with HMIs as they approach intersections in a connected environment, where information reception is the key for drivers to make informed decisions. Therefore, the findings of this study further contribute to the literature on the utilization of in-vehicle technology among different driver groups from a visual behavioral perspective. The study presented in (Charlton et al., 2013) investigated the distraction activities of elderly drivers before intersections. The results showed that elderly drivers would selectively participate in secondary activities according to road or driving conditions. When the current driving task is challenging, such as stably and safely entering an intersection, the driver would be less involved in secondary tasks. Previous studies indicate that older drivers might be less active when interacting with new in-vehicle technologies. Despite a considerable proportion of middle-aged individuals among the



(a) Green light and remaining time from phase transition is 5 s and 3 s

(b) Green light and remaining time from phase transition is 10 s and 6 s

(c) Green light and remaining time from phase transition is 24 s and 21 s

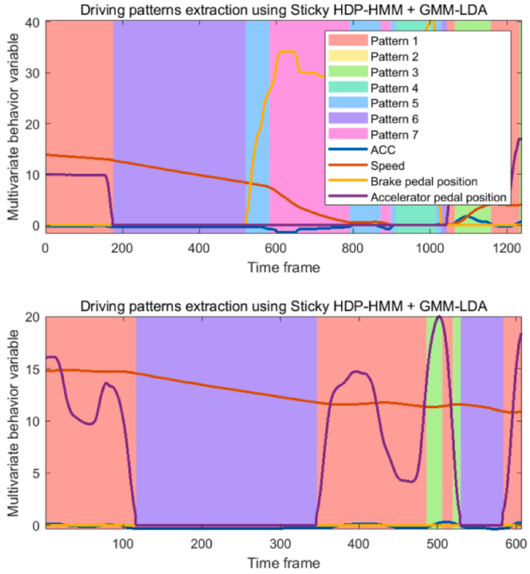
(d) Red light and remaining time from phase transition is 5 s and 3 s

Fig. 13. Example of driving pattern extraction under different traffic light phases.

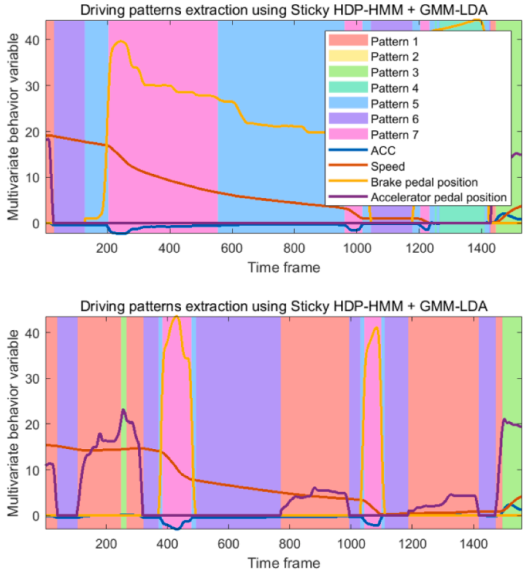
professional drivers in this study and a predominance of relatively younger non-professional drivers in the control group, this disparity may indicate that heightened driving experience is associated with decreased susceptibility to new technologies. Therefore, this study also supports the results of previous studies. Unfortunately, this study did not reveal any significant correlation between age, experience and the driver's visual interaction pattern. Hence, it is important to exercise caution when interpreting the results of this study, as the sample of professional drivers does not encompass younger drivers, and non-professional drivers do not include elderly drivers.

5.2. First interaction and response time

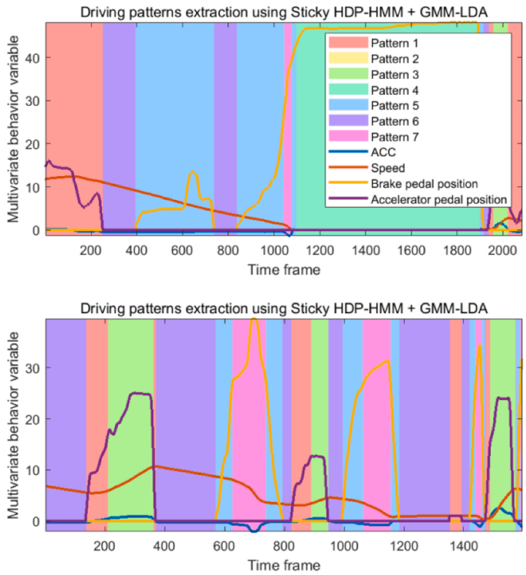
Although the professional drivers tended to interact less with the HMI, this study also deduced that their first interaction occurred earlier compared with the non-professional drivers, as did their response behavior. The first interaction time of drivers is highly correlated with the response time, which indicates that the earlier the HMI is observed, the earlier the response behavior will be. Many studies have shown that professional drivers are generally considered to be experienced and skilled (Damm et al., 2011). In particular,



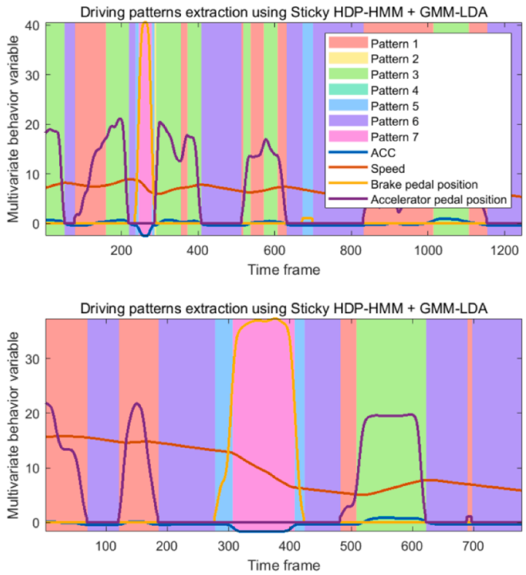
(a) Red light and remaining time from phase transition is 9 s and 7 s



(b) Red light and remaining time from phase transition is 28 s and 26 s



(c) Yellow light



(d) Baseline Group

Fig. 14. Example of driving pattern extraction under different traffic light phases and baseline driving scenarios.

they are better at anticipating risk scenarios than non-professional drivers (Borowsky and Oron-Gilad, 2013). Chen et al. (Chen et al., 2021) observed that professional drivers had shorter braking reaction times than non-professionals, which indicates that the perception-decision duration of the professional drivers is shorter than that of the non-professional drivers. Similarly, this study deduced a significant difference in the response time between the professional and non-professional drivers, and the professional drivers also observed the HMI less frequently than the non-professional drivers. Therefore, they are more likely to positively respond after obtaining the connection information and rely less on this information, which may reflect their confidence in their driving skills and their greater adaptability to the connected environment.

In addition, the response times had a lognormal distribution, and the first interaction time with the HMI had an exponential distribution for the professional and non-professional drivers. Moreover, the driver's first interaction time with the HMI and the response time was independent of the traffic light phase and the remaining time. This demonstrates that the driver's response to the HMI technology has no correlation with the prompted information, but may be related to age, experience, habits, and gender. Therefore, in the extended work, it may be worth exploring the relationship between the types of HMI information (e.g., safety warning

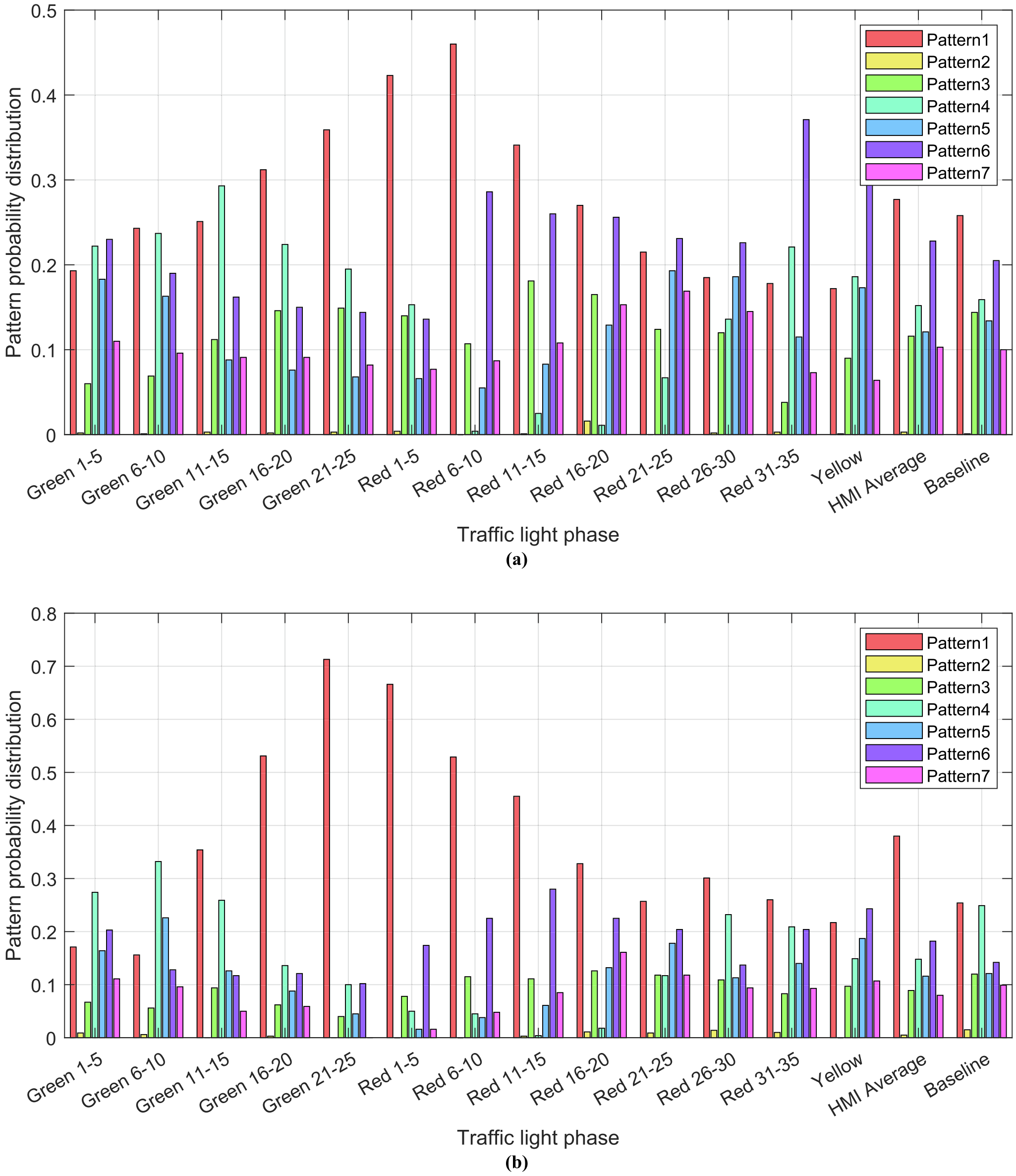


Fig. 15. Probability distribution of driving patterns in two environments for professional (a) non-professional drivers (b).

Table 12
Probability of seven driving patterns of professional and non-professional drivers in HMI and baseline scenarios.

		Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5	Pattern 6	Pattern 7
All driver	HMI Average	0.329	0.004	0.102	0.150	0.119	0.205	0.092
	Baseline	0.256	0.008	0.131	0.208	0.127	0.170	0.099
Professional	HMI Average	0.277	0.003	0.116	0.152	0.121	0.228	0.103
	Baseline	0.258	0.001	0.144	0.159	0.134	0.205	0.100
Non-professional	HMI Average	0.380	0.005	0.089	0.148	0.116	0.182	0.080
	Baseline	0.254	0.015	0.120	0.249	0.121	0.142	0.099

information and assisted driving information) and the driver's response time. The driver's interaction and response to assisted driving information in a connected environment have been previously studied (Kramer et al., 2007). Ali et al. (Ali et al., 2020) reported the driver interaction and response time to lane-change requests in a connected environment. They noted that older drivers responded to requests earlier than younger drivers. Although the lane changing and intersection approaching behavior are two distinct driving tasks, the response time is an essential component of the two. Similarly, Caird et al. (Caird et al., 2008) reported that older drivers benefitted more and responded more quickly than younger drivers after receiving signalized intersection representations through in-vehicle information systems. The findings of this study contribute to the literature on the response time of professional and non-professional drivers.

5.3. Driving performance and behavioral patterns

The PIWS results should be interpreted with caution because each participant only drove three times in each driving scenario. Although the overall PIWS showed that from the baseline to the HMI scenario, the frequency of drivers passing through the intersection significantly increased, only the PIWS of the professional drivers had a significant difference. This shows that the professional drivers can pass the intersection more efficiently with the provision of traffic signal information. The non-professional drivers may not be as efficient in using traffic light information as professional drivers. The studies on PIWS at intersections in connected environments for professional and non-professional drivers are few. Although no direct evidence is available to support this finding, it can be deduced from other perspectives. Long et al. (Long et al., 2013) showed that the countdown timers could assist the drivers in decision-making, which allows to reduce the risky driving maneuvers during phase transitions. Their results showed that the countdown timer led to more vehicles crossing the intersection during the yellow phase. Ma et al. (Ma et al., 2010) also obtained similar results in their study. In addition, the study conducted by Islam et al. (Islam et al., 2016) showed that real-time traffic signal timers at signalized intersections significantly reduced the headway of the first vehicle and the starting time, and improved the traffic efficiency of intersections. These results demonstrate that the traffic light information can improve the traffic efficiency of intersections, and the professional drivers may benefit even more.

A qualitative analysis of the driving performance demonstrated that the effect size (i.e., Cohen's d) of each driving performance indicator of the professional drivers between two driving scenarios is smaller than that of the non-professional drivers, as shown in the post-hoc paired tests when passing through intersections. In this case, the connected environment can significantly reduce the speed fluctuation and average speed of the non-professional drivers, which indicates that the longitudinal stability of the non-professional drivers is significantly improved in the connected environment. Similarly, the non-professional drivers can also reduce the unnecessary acceleration and deceleration in the connected environment, which has a positive significance for improving the driving comfort and reducing the energy consumption. When the driver fails to pass through the intersection, the connected environment can inhibit the acceleration behavior, and it has a significant impact on the braking behavior of the non-professional drivers. This study also deduced that, compared with the baseline scenario, the driving behavior of the non-professional drivers in the HMI scenario is relatively stable, and some performances are not significantly different from those of the professional drivers. This shows that a connected environment can benefit inexperienced young drivers. The driving pattern extraction model can well describe the seven behavior patterns in the intersection-approaching process. The quantitative analysis demonstrated that the connected information can greatly improve the probability distribution of stable driving, reduce unnecessary acceleration behavior, increase the proportion of coasting pattern, and reduce the probability of strong braking behavior. However, it was also deduced that the professional drivers benefited relatively less than the non-professional drivers. Both the qualitative and quantitative analysis showed that the professional drivers are less affected by changes in the driving environment compared with the non-professional drivers, and their driving performance is relatively stable, which is consistent with the second hypothesis of this study that professional drivers use the information provided by the HMI faster and demonstrate more consistent driving performance.

In the existing studies, there is no quantitative analysis of driving patterns in the connected environment. However, some studies have compared the behavioral differences between professional and non-professional drivers from different perspectives, such as compensation strategies (Chen et al., 2021), mental workload (Yared and Patterson, 2020), partial sleep deprivation (Mahajan and Velaga, 2022), risky driving (Öz et al., 2010), and performing secondary tasks (Choudhary and Velaga, 2019). It can be deduced that the non-professional drivers often have relatively large fluctuations in driving ability under restricted conditions, while the professional drivers often adopt compensation strategies to overcome unfavorable factors in order to maintain relatively stable driving in different driving environments due to their higher experience. The findings of this paper reinforce this conclusion. Although the connected environment provides drivers with favorable information, the results of this paper show that the professional drivers try to maintain stable driving while using new technologies.

5.4. Study limitations and future work

The findings of this study should be interpreted in the context of its limitations. The first is the effectiveness of apparatus and driving scenarios. The simulator studies do not entirely reflect the real-world driving conditions. However, many studies have demonstrated the absolute and relative validity of simulation experiments (Wynne et al., 2019; Saifuzzaman et al., 2015). In particular, Lobjois et al. (Lobjois et al., 2021) reported that the speed measurements were relatively robust in simulator studies. In addition, due to the complexity of the connected driving environment, it is impossible to accurately simulate the real connected traffic environment. Therefore, it is reasonable to consider only one connected driving scenario. The impact of the auxiliary information on the collaborative driving of different driver groups in a connected environment is also an interesting topic for future study. Furthermore, the reason for setting the connection information prompt at 300 m from the intersection was the lack of relevant standards and reports for establishing a connected traffic environment. The researchers use different experimental variables and scenarios based on their understanding of the connected driving environment (Kraft et al., 2019; Faas et al., 2020).

Secondly, the sample of professional drivers in this study did not include younger drivers (i.e., the youngest being 31 years old), while non-professional drivers also did not include elderly drivers (i.e., the oldest being 47 years old, categorized as middle-aged). Therefore, conclusions drawn from this study should be interpreted within this context. Validating whether the behaviors of driver groups across different genders and age brackets align with the current findings represents another focus for future research. Despite efforts to recruit more female drivers, the lack of female professional drivers was a prevalent issue globally (O. f. N. Statistics, 2024). Other potential confounding factors, such as risk-taking and anticipatory psychology, may also have an impact on driver interaction and response behaviors, and are therefore a focus of future expansion of this study.

Third, the consideration of setting low traffic flow density in this paper consists in exploring the real response of drivers in the connected environment (under the speed limit conditions of urban scenes) as much as possible, rather than being affected by other vehicles in the traffic flow. To avoid the driver's decision-making being the coupling result of the cognition of connected information and the interaction with the surrounding vehicles, this study set a low traffic flow density. However, the response behavior under different traffic densities will be the focus of future work. In addition, this study is only a manifestation of driver visual interaction and response behavior in a connected environment, and other new technologies (e.g., voice prompts or vibration feedback) require further research.

6. Conclusion

This study investigates the visual interaction, response characteristics, driving performance, and behavior patterns differences in the intersection-approach process of professional and non-professional drivers. More precisely, an effect analysis was conducted on the driving performance, and a driving pattern extraction model was developed to analyze the impact of the connected information on the driving behavior in a qualitative and quantitative manner. The obtained results showed that professional drivers look at the HMI less frequently than non-professional drivers. Moreover, they interacted with the HMI earlier, and their response behavior occurred earlier. In addition, the connected information reduced the unnecessary acceleration and deceleration behaviors, and increased the proportion of stable driving. However, the professional and non-professional drivers were differently affected by this information. After obtaining the traffic lights information, the professional drivers more quickly responded to this information, which resulted in relatively stable driving performance. The non-professional drivers greatly increased the driving stability from receiving connection information. However, their efficiency through intersections did not increase.

To the authors' knowledge, this is the first study to quantify driving behavior in a connected environment. The findings of this study provide preliminary results showing differences in the visual interaction and behavioral patterns of professional and non-professional drivers when exposed to new technologies in vehicles. Exploring the driver's utilization of connected information and the differences in behavior patterns between traditional and connected environments can provide a reference for the in-vehicle communication equipment manufacturers to provide personalized services. It could also provide a theoretical basis for the study of ecological driving strategies. Furthermore, this paper contributes to the field of human–machine interaction interface design and personalized decision-making assistance systems for heterogeneous drivers in a connected environment. For example, professional drivers may prefer simplified traffic information when using HMI, while non-professional young drivers may prefer rich information to help make better decisions. Organizations have the capability to deliver training in various skills, attuned to distinct anticipatory learning proficiencies among disparate groups of drivers. This tailored approach aims to foster heightened adaptation to the progressively prevalent advancements in in-vehicle technologies.

CRediT authorship contribution statement

Hailun Zhang: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Rui Fu:** Writing – review & editing, Visualization, Supervision, Formal analysis, Conceptualization. **Jianqiang Wang:** Writing – review & editing, Visualization, Funding acquisition. **Simeon C. Calvert:** Writing – review & editing, Visualization, Supervision, Formal analysis. **Hans van Lint:** .

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grant 52131201, the National Funded Postdoctoral Researcher Program of China (No. GZB20230355), the Beijing Natural Science Foundation (No. 3244031).

References

Ahmad, H., & Halim, H. (2017). Determining sample size for research activities. *Selangor Business Review*, 20–34.

Ali, Y., Bliemer, M. C. J., Zheng, Z., & Haque, M. M. (2020). Cooperate or not? Exploring drivers' interactions and response times to a lane-changing request in a connected environment. *Transportation Research Part C: Emerging Technologies, 120*, Article 102816, 2020/11/01.

Ali, Y., Sharma, A., Haque, M. M., Zheng, Z., & Saifuzzaman, M. (2020). The impact of the connected environment on driving behavior and safety: A driving simulator study. *Accident Analysis & Prevention, 144*, Article 105643, 2020/09/01.

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accident Analysis & Prevention, 59*, 548–565, 2013/10/01.

Caird, J. K., Chisholm, S. L., Edwards, C. J., & Creaser, J. I. (2007). The effect of yellow light onset time on older and younger drivers' perception response time (PRT) and intersection behavior. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(5), 383–396, 2007/09/01.

Caird, J. K., Chisholm, S. L., & Lockhart, J. (2008). Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. *International Journal of Human-Computer Studies, 66*(3), 132–144, 2008/03/01.

Calvert, S. C., & Arem, B.v. (2020). Cooperative adaptive cruise control and intelligent traffic signal interaction: A field operational test with platooning on a suburban arterial in real traffic. *Iet Intelligent Transport Systems, 14*(12), 1665–1672, 2020/10/28.

Calvert, S. C., Klunder, G., Steendijk, J. L. L., & Snelder, M. (2020). "The impact and potential of cooperative and automated driving for intelligent traffic signal corridors: A field-operational-test and simulation experiment". *Case Studies on Transport Policy, 8*(3), 901–919, 2020/09/01.

Castellà, J., & Pérez, J. (2004). Sensitivity to punishment and sensitivity to reward and traffic violations. *Accident Analysis & Prevention, 36*(6), 947–952, 2004/11/01.

Charlton, J. L., Catchlove, M., Scully, M., Koppel, S., & Newstead, S. (2013). Older driver distraction: A naturalistic study of behaviour at intersections. *Accident Analysis & Prevention, 58*, 271–278, 2013/09/01.

Chatzisavvas, K. C., Moustakidis, C. C., Panos, C. P. (2005). "Information entropy, information distances, and complexity in atoms," *The Journal of Chemical Physics,* vol. 123, no. 17, 2005.

Chen, Y., Li, G., Li, S., Wang, W., Li, S. E., & Cheng, B. (2021). Exploring Behavioral Patterns of Lane Change Maneuvers for Human-Like Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*, 1–14.

Chen, T., Sze, N. N., Newnam, S., & Bai, L. (2021). Effectiveness of the compensatory strategy adopted by older drivers: Difference between professional and non-professional drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 77*, 168–180, 2021/02/01.

Cheung, I., & McCartt, A. T. (2011). Declines in fatal crashes of older drivers: Changes in crash risk and survivability. *Accident Analysis & Prevention, 43*(3), 666–674, 2011/05/01.

Choudhary, P., & Velaga, N. R. (2019). Effects of phone use on driving performance: A comparative analysis of young and professional drivers. *Safety Science, 111*, 179–187, 2019/01/01.

Damm, L., Nachtergaële, C., Meskali, M., & Berthelon, C. (2011). The evaluation of traditional and early driver training with simulated accident scenarios. *Human Factors, 53*(4), 323–337, 2011/08/01.

Dorn, L., & Brown, B. (2003). Making sense of invulnerability at work—a qualitative study of police drivers. *Safety Science, 41*(10), 837–859, 2003/12/01.

Duke, J., Guest, M., & Boggess, M. (2010). Age-related safety in professional heavy vehicle drivers: A literature review. *Accident Analysis & Prevention, 42*(2), 364–371, 2010/03/01.

Dykstra, C., Davis, J. J., & Conlon, E. G. (2020). Tactical and strategic driving behaviour in older drivers: The importance of readiness to change. *Accident Analysis & Prevention, 141*, Article 105519, 2020/06/01.

Evans, L. (1991). *Traffic safety and the driver*. Science Serving Society.

Evans, L. (2004). *Traffic safety*.

Faas, S. M., Mathis, L.-A., & Baumann, M. (2020). External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F: Traffic Psychology and Behaviour, 68*, 171–186, 2020/01/01.

Fox, E. B., Sudderth, E. B., Jordan, M. I., & Willsky, A. S. (2011). A Sticky HDP-HMM with application to speaker diarization. *The Annals of Applied Statistics, 5*(2A), 1020–1056.

Guo, Y., Zhang, H., Wang, C., Sun, Q., & Li, W. (2021). Driver lane change intention recognition in the connected environment. *Physica A: Statistical Mechanics and its Applications, 575*, Article 126057, 2021/08/01.

He, D., & Donmez, B. (2019). Influence of driving experience on distraction engagement in automated vehicles. *Transportation Research Record, 2673*(9), 142–151, 2019/09/01.

Hong, S., Min, B., Doi, S., & Suzuki, K. (2016). Approaching and stopping behaviors to the intersections of aged drivers compared with young drivers. *International Journal of Industrial Ergonomics, 54*, 32–41.

Ilgin Guler, S., Menendez, M., & Meier, L. (2014). Using connected vehicle technology to improve the efficiency of intersections. *Transportation Research Part C: Emerging Technologies, 46*, 121–131, 2014/09/01.

Islam, M. R., Hurwitz, D. S., & Macuga, K. L. (2016). Improved driver responses at intersections with red signal countdown timers. *Transportation Research Part C: Emerging Technologies, 63*, 207–221, 2016/02/01.

Islam, M., & Ozkul, S. (2019). Identifying fatality risk factors for the commercial vehicle driver population. *Transportation Research Record, 2673*(9), 297–310, 2019/09/01.

Iwatsuki, A., Hirayama, T., Morita, J., Mase, K., 2016. Skilled gaze behavior extraction based on dependency analysis of gaze patterns on video scenes. In: Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, Charleston, South Carolina, 2016, pp. 299–302.

Johnson, M.J., Willsky, A. S., "Bayesian nonparametric hidden semi-Markov models," 2013.

Kenney, J. B. (2011). Dedicated Short-Range Communications (DSRC) Standards in the United States. *Proceedings of the IEEE, 99*(7), 1162–1182.

Kircher, K., Fors, C., & Ahlstrom, C. (2014). Continuous versus intermittent presentation of visual eco-driving advice. *Transportation Research Part F: Traffic Psychology and Behaviour, 24*, 27–38, 2014/05/01.

Kohl, J., Gross, A., Henning, M., & Baumgarten, T. (2020). Driver glance behavior towards displayed images on in-vehicle information systems under real driving conditions. *Transportation Research Part F: Traffic Psychology and Behaviour, 70*, 163–174, 2020/04/01.

Kraft, A.-K., Maag, C., & Baumann, M. (2019). How to support cooperative driving by HMI design?". *Transportation Research Interdisciplinary Perspectives, 3*, Article 100064, 2019/12/01.

Kraft, A.-K., Maag, C., & Baumann, M. (2020). Comparing dynamic and static illustration of an HMI for cooperative driving. *Accident Analysis & Prevention, 144*, Article 105682, 2020/09/01.

Kramer, A. F., Cassavaugh, N., Horrey, W. J., Becic, E., & Mayhugh, J. L. (2007). Influence of Age and Proximity Warning Devices on Collision Avoidance in Simulated Driving. *Human Factors, 49*(5), 935–949, 2007/10/01.

Lal, S. K. L., Craig, A., 2002. Fatigue neurophysiology in professional and non-professional drivers. In: *Proceedings of the Australasian road safety research, policing and education conference,* vol. 6, no. 1, pp. 125-130.

Li, G., Chen, Y., Cao, D., Qu, X., Cheng, B., & Li, K. (2021). Extraction of descriptive driving patterns from driving data using unsupervised algorithms. *Mechanical Systems and Signal Processing, 156*, Article 107589, 2021/07/01.

Li, Y. C., Sze, N. N., Wong, S. C., Yan, W., Tsui, K. L., & So, F. L. (2016/10/01/, 2016.). A simulation study of the effects of alcohol on driving performance in a Chinese population. *Accident Analysis & Prevention, 95*, 334–342.

Li, X., Vaezipour, A., Rakotonirainy, A., Demmel, S., & Oviedo-Trespalacios, O. (2020). Exploring drivers' mental workload and visual demand while using an in-vehicle HMI for eco-safe driving. *Accident Analysis & Prevention, 146*, Article 105756.

Liu, Z., & Kircher, K. (2018). Comparison of a time-and a speed-based traffic light assistance system. *Cognition, Technology & Work, 20*(1), 93–103.

Lobjois, R., Faure, V., Désiré, L., & Benguigui, N. (2021). Behavioral and workload measures in real and simulated driving: Do they tell us the same thing about the validity of driving simulation? *Safety Science, 134*, Article 105046, 2021/02/01.
Long, K., Liu, Y., & Han, L. D. (2013). Impact of countdown timer on driving maneuvers after the yellow onset at signalized intersections: An empirical study in Changsha, China. *Safety Science, 54*, 8–16, 2013/04/01.
M. o. H. a. U.-R. D. o. t. P. s. R. o. China, "Specification for layout of urban road traffic signs and markings," China Planning Press, 2015.
M. o. T. o. t. P. s. R. o. China, "Specification for Layout of Highway Traffic Signs and Markings," People's Communications Press, 2009.
Ma, W., Liu, Y., & Yang, X. (2010). Investigating the Impacts of Green signal countdown devices: Empirical approach and case study in China. *Journal of Transportation Engineering, 136*(11), 1049–1055, 2010/11/01.
Maag, C., Kraft, A.-K., Neukum, A., & Baumann, M. (2022). Supporting cooperative driving behaviour by technology – HMI solution, acceptance by drivers and effects on workload and driving behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 84*, 139–154, 2022/01/01.
Mahajan, K., & Velaga, N. R. (2022). Effects of partial sleep deprivation: A comparative assessment of young non-professional and professional taxi drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 85*, 209–220, 2022/02/01.
Monsaingeon, N., Caroux, L., Mouginé, A., Langlois, S., & Lemercier, C. (2021). Impact of interface design on drivers' behavior in partially automated cars: An on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 508–521, 2021/08/01.
Montgomery, J., Kusano, K. D., & Gabler, H. C. (2014). Age and gender differences in time to collision at braking from the 100-car naturalistic driving study. *Traffic Injury Prevention, 15*(sup1), S15–S20, 2014/09/26.
Newnam, S., Koppel, S., Molnar, L. J., Zakrajsek, J. S., Eby, D. W., & Blower, D. (2020). Older truck drivers: How can we keep them in the workforce for as long as safely possible? *Safety Science, 121*, 589–593, 2020/01/01.
Ni, R., Kang, J. J., & Andersen, G. J. (2010.). Age-related declines in car following performance under simulated fog conditions. *Accident Analysis & Prevention, 42*(3), 818–826, 2010/05/01.
O. f. N. Statistics. "Which jobs do men and women do? occupational breakdown by gender," March 10, 2024; https://careersmart.org.uk/occupations/equality/which-jobs-do-men-and-women-do-occupational-breakdown-gender.
Öz, B., Özkan, T., & Lajunen, T. (2010). Professional and non-professional drivers' stress reactions and risky driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*(1), 32–40, 2010/01/01.
Paul, M., Ghosh, I., & Mazharul Haque, M. (2022). The effects of green signal countdown timer and retiming of signal intervals on dilemma zone related crash risk at signalized intersections under heterogeneous traffic conditions. *Safety Science, 154*, Article 105862, 2022/10/01.
Prabhudesai, K. S., Mainsah, B. O., Collins, L. M., Throckmorton, C. S.,"Augmented latent Dirichlet allocation (LDA) topic model with Gaussian mixture topics." pp. 2451-2455.
Romoser, M. R. E., & Fisher, D. L. (2009). The effect of active versus passive training strategies on improving older drivers' scanning in intersections. *Human Factors, 51*(5), 652–668, 2009/10/01.
Romoser, M. R. E., Pollatsek, A., Fisher, D. L., & Williams, C. C. (2013). Comparing the glance patterns of older versus younger experienced drivers: Scanning for hazards while approaching and entering the intersection. *Transportation Research Part F: Traffic Psychology and Behaviour, 16*, 104–116, 2013/01/01.
Rosenbloom, T., & Shahar, A. (2007). Differences between taxi and nonprofessional male drivers in attitudes towards traffic-violation penalties. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(5), 428–435, 2007/09/01.
Rouzikhah, H., King, M., & Rakotonirainy, A. (2013). Examining the effects of an eco-driving message on driver distraction. *Accident Analysis & Prevention, 50*, 975–983, 2013/01/01.
Saifuzzaman, M., Zheng, Z., Mazharul Haque, M., & Washington, S. (2015). Revisiting the Task-Capability Interface model for incorporating human factors into car-following models. *Transportation Research Part B: Methodological, 82*, 1–19, 2015/12/01.
Sato, R., Kamezaki, M., Sugano, S., Iwata, H. "Gaze pattern analysis in multi-display systems for teleoperated disaster response robots." pp. 003534-003539.
Savage, S. W., Zhang, L., Swan, G., & Bowers, A. R. (2020). The effects of age on the contributions of head and eye movements to scanning behavior at intersections. *Transportation Research Part F: Traffic Psychology and Behaviour, 73*, 128–142, 2020/08/01.
Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., & Haque, M. M. (2019). Estimating and comparing response times in traditional and connected environments. *Transportation Research Record, 2673*(4), 674–684.
Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D. J., & Biondi, F. (2019). Assessing the visual and cognitive demands of in-vehicle information systems. *Cognitive Research: Principles and Implications, 4*(1), 18, 2019/06/21.
Sullivan, J. M., Tsimhoni, O., & Bogard, S. (2008). Warning reliability and driver performance in naturalistic driving. *Human Factors, 50*(5), 845–852, 2008/10/01.
Tao, D., Zhang, R., & Qu, X. (2017). The role of personality traits and driving experience in self-reported risky driving behaviors and accident risk among Chinese drivers. *Accident Analysis & Prevention, 99*, 228–235, 2017/02/01.
Tielert, T., Killat, M., Hartenstein, H., Luz, R., Hausberger, S., Benz, T. "The impact of traffic-light-to-vehicle communication on fuel consumption and emissions." pp. 1-8.
Vaezipour, A., Rakotonirainy, A., Haworth, N., & Delhomme, P. (2017). Enhancing eco-safe driving behaviour through the use of in-vehicle human-machine interface: A qualitative study. *Transportation Research Part A: Policy and Practice, 100*, 247–263, 2017/06/01.
Wynne, R. A., Beanland, V., & Salmon, P. M. (2019). Systematic review of driving simulator validation studies. *Safety Science, 117*, 138–151, 2019/08/01.
Yan, W., Wong, S. C., Loo, B. P. Y., Wu, C. Y. H., Huang, H., Pei, X., & Meng, F. (2022). An assessment of the effect of green signal countdown timers on drivers' behavior and on road safety at intersections, based on driving simulator experiments and naturalistic observation studies. *Journal of Safety Research, 82*, 1–12, 2022/09/01.
Yan, X., Zhang, Y., & Ma, L. (2015). The influence of in-vehicle speech warning timing on drivers' collision avoidance performance at signalized intersections. *Transportation Research Part C-Emerging Technologies, 51*, 231–242.
Yang, H., Almutairi, F., & Rakha, H. (2021). Eco-driving at signalized intersections: a multiple signal optimization approach. *IEEE Transactions on Intelligent Transportation Systems, 22*(5), 2943–2955.
Yang, B., Zheng, R., Shimono, K., Kaizuka, T., & Nakano, K. (2017). Evaluation of the effects of in-vehicle traffic lights on driving performances for unsignalised intersections. *IET Intelligent Transport Systems, 11*(2), 76–83.
Yared, T., & Patterson, P. (2020). The impact of navigation system display size and environmental illumination on young driver mental workload. *Transportation Research Part F: Traffic Psychology and Behaviour, 74*, 330–344, 2020/10/01.
Yu, B., Bao, S., Feng, F., & Sayer, J. (2019). Examination and prediction of drivers' reaction when provided with V2I communication-based intersection maneuver strategies. *Transportation Research Part C: Emerging Technologies, 106*, 17–28.
Zhang, W., & Wang, W. (2019). Learning V2V interactive driving patterns at signalized intersections. *Transportation Research Part C: Emerging Technologies, 108*, 151–166.
Zhang, Y., Wu, C., Qiao, C., & Hou, Y. (2019). The effects of warning characteristics on driver behavior in connected vehicles systems with missed warnings. *Accident Analysis & Prevention, 124*, 138–145, 2019/03/01/.
Zheng, Z., Ahn, S., Chen, D., & Laval, J. (2011). Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations. *Transportation Research Part B: Methodological, 45*(2), 372–384, 2011/02/01.
Zheng, Z., & Washington, S. (2012). On selecting an optimal wavelet for detecting singularities in traffic and vehicular data. *Transportation Research Part C: Emerging Technologies, 25*, 18–33, 2012/12/01.
Zhou, M., Yu, Y., & Qu, X. (2019). Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: a reinforcement learning approach. *IEEE Transactions on Intelligent Transportation Systems*.