Journal of Computing and Electronic Information Management

ISSN: 2413-1660

Impact of Intelligent Connected Vehicles Voice Interaction System on Driving Safety

Wenxia Xu [1, a], Jianjie Tan [1, b], Jun Ma[1, c] and Zaiyan Gong[1, *]

[1]School of Automotive Studies, Tongji University, Shanghai, China

*Corresponding author e-mail: gongzaiyan@foxmail.com

[a]wenxiaxu@hotmail.com, [b]tan_jianjie@foxmail.com, [c] majun.tongji@foxmail.com

Abstract: The development of intelligent connected vehicle(ICV) makes voice interaction an important interaction method in the car. The ICV's voice interaction system supports natural speech recognition and can complete a large number of secondary tasks. This paper is based on the Chinese market and 5 ICVs are screened out. Application scenarios for voice interaction systems are constructed through pre-survey. Using simulated driving platform which can be connected with mass-produced vehicle, data on driving performance can be obtained. From the lateral and longitude movement aspects, the corresponding indicators are selected to analyze the impact of ICV's voice interaction system on driving safety.

Keywords: voice interaction, driving Simulation, driving safety

1. Introduction

The progress of science and technology and the change of life style have brought about the change of people's habits, and vehicles have also become an important living space for people's lives. In this space, the In-Vehicle Information System(IVIS) plays an important role. When driving, drivers often need to listen to the audio, answer the phone, find the point of interest (POI) and etc., which will make the driver's hand away from the steering wheel, eyes away from the road and can't focus on the driving. According to the data statistics, more than 90% of road traffic accidents in the United States and Japan are caused by driver factors [1], and more than 80% of road traffic accidents in China[2]. Among the road traffic accidents caused by drivers in 2016, 54.18% were caused by perception errors, 35.88% by judgment and decision-making errors, and 9.15% by operation and decision-making errors[3].

Voice interaction is being applied to cockpit more and more because of its high degree of naturalness and rich realizable functions. It has become an essential function of Intelligent Connected Vehicle(ICV). According to iFLYTEK's official data, in 2018, it had more than 4 million mobile intelligent voice kits. With more and more users, we need to study the impact of different intelligent voice interaction systems on driving safety from the perspective of users.

2. Methods

Considering the complexity of the in-vehicle voice interaction system and factors such as drivers, vehicle condition, road environment, and pedestrians, the method of driving simulation test was finally selected. This method does not require real driving, which can ensure the safety. At the same time, it can highly restore the driving scenario, provide true and accurate data, and support for subsequent research.

According to the hardware configuration, the experimental platform can be divided into three modules: simulation driving platform, driving simulation environment and eye tracker. It can be connected with any segment and form of passenger vehicles, and use mass-produced vehicles to drive on the simulated driving road.

When the simulated driving platform is running, the steering wheel rotates synchronously when the front wheels rotate. The left steering wheel is connected to an angle sensor, and is input into the virtual driving scene as a steering wheel steering signal. The other two sensors are installed on the vehicle's brake pedal and accelerator pedal to obtain pedal data. The driving behavior is transformed by the steering wheel and pedals into system signals and transmitted to the virtual driving scene to realize vehicle control.

The simulation environment is designed based on the Unity 3D platform. The definition and design of all roads are based on real Chinese road scenes. In this paper, the experiment chose a 13.8km route, including urban roads, elevated roads and suburban roads. The road design was consistent with the real environment.

3. Pre-survey and experiment

Currently, there have been many mass-produced ICVs equipped with voice interaction systems, which have natural voice interaction and rich vehicle control functions. This paper is based on the Chinese market and selects mainstream ICV voice system manufacturers. Michael Braun[4] studied that the different visual schemes that affect the safety performance. Finally, 5 ICVs with typical voice interaction system were selected, which have network and natural voice interaction ability, as shown in the Fig.1 below.

Fig.1 Interior of 5 tested cars: the first row from left to right: RR Model, GB Model, LK model, WE Model and DA Model

Before the experiment, a pre-survey was conducted for the owners of 5 models, including quantitative questionnaires through the owner forum, targeted connection and etc., and qualitative interviews with owners of 5 models. In the pre-survey questionnaire, the common driving secondary tasks are listed and the owner will choice the secondary tasks finished by the voice interaction system usually. At the same time, the owner will be asked in the interview if there are common tasks to add. During the interview, it was found that when different owners conduct tasks among the 25 tasks previously selected, the language expression is very close. The difference is mainly the individual words, such as "end navigation" or "exit navigation". There were no instances of task errors or failures due to differential expression for the 5 models. In terms of language expression, for commonly used and traditional function modules such as air-conditioning, radio / music, etc., the user's language expression is often simple, direct, and has little difference. For some innovative features or emotional voice interactions, the user's voice expression would be significantly different.

Finally, through the pre-survey, the following test tasks are determined:

Table.1 selected tasks

Function	Secondary task
Air conditioner	Turn air conditioning on / off
	Turn temperature up / down
	Increase / decrease air volume
Radio/music	Previous / Next
	Switch music mode (radio, music, Bluetooth music, etc.)
	Search music
Navigation	Navigate to "Tongji Technology Plaza"
	Navigate to the POI "Nearest Gas Station"
Telephone	Call "Mr. Wu"
	Call "10086"

When designing the testing scenario, the user's usage habits, the operability of the experiment, and the scientificity of the result were considered. In the final voice interaction scenario in this paper, the steering wheel buttons were used as the wakeup method during the task execution. And the drivers can complete the relevant task by their own expression habits.

After the tested driver arrives in the laboratory, the experimenter first introduces the principle and basic process. Then the experimenter records the basic information of the tested driver, including age, driving age, driving license check, etc.. Then a simple Mandarin test and visual test would be done. After completing the above process and completing the test, the tested driver enters the formal experiment. The formal experiment is divided into 5 stages: adaptive practice, interactive practice, non-task driving, tasked driving, and subjective evaluation. Throughout the experiment, the experimenter is located in the co-pilot position, and the Gopro camera assumed to be in the sunroof is used to record the entire course and recording of the experiment. A total of 12 people, 60 people participated in the experimental process.

4. Results and analysis

The data of the simulated driving platform is collected by sensors and matched with the task-record, including the real-time speed of the vehicle, the offset distance from the lane centerline, and the steering wheel angle. In this paper, the time segment, from the time when the driver receives the task instruction to the time when the task is completed, is used as the cut-off point to divide the data segment, and the time period data is used for analysis.

This paper adopts the pauta criterion method for data cleaning, which is suitable for the case where the sample data is relatively large[5]. This is consistent with the actual situation of this topic, so this method is selected for cleaning abnormal data.

The calculation formula used by pauta criterion for inspection is:

$$|\text{value} - \text{mean}| > 3\sigma$$

Where value refers to the actual data, mean refers to the mean of the sample data, and σ refers to the standard deviation of the sample. According to the normal distribution law of random variables, after repeated tests, the test data are 99.73% in the range of (mean-3 * σ, mean + 3 * σ), so the data not in this range is considered as small probability event, which is determined as abnormal data.

Driving performance and visual demand are two key factors for evaluating safety during when finishing secondary tasks. This paper will evaluate the driving safety from the aspect of driving performance when driers finish secondary tasks using voice interaction.

4.1 Lateral directions

Lateral movement of the vehicle comes from the driver's control of the vehicle's steering wheel. Therefore, the steering wheel angle is an important data index to represent the lateral motion of vehicles. When using voice interaction, the attention shifts to the operation of the secondary task, so that the vehicle's driving track will be off from the center. When the driver finds that the vehicle's driving track deviates from the center of the lane, he will make rapid adjustment while carrying out the secondary task. Compared with the focus driving, the smoothness and connection of the steering angle will be affected when the voice interaction is carried out. Fig. 2 below is a comparison of 30s steering wheel angle data during voice interaction and 30s data during normal driving. It can be seen that the fluctuation frequency and range of steering wheel angle during voice interaction is significantly higher than that normal driving. In the process of voice interaction, the steering wheel angle fluctuates between the positive and negative values for many times, which indicates that the operation stability of the driver is affected.

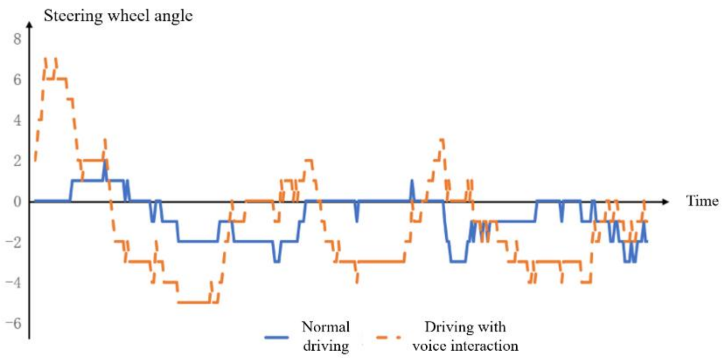

Fig. 2 Steering wheel angle comparison

In order to evaluate the lateral motion of the vehicle, this paper uses the Steering Wheel Angle Standard Deviation（SWASD）and lane departure standard deviation (LDSD) to evaluate. Where, the definition of steering wheel angle standard deviation is shown in the following formula.

The Steering Wheel Angle Standard Deviation(SWASD）is defined as follows

$$SWASD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (deg_i - \overline{deg})^2}$$

Among them，N refers to the number of steering wheel samples obtained, deg_i refers to the steering wheel angle value, and $\overline{deg}$ refers to the average value of the steering wheel angle.

LDSD is defined as the standard deviation of the lateral distance between the car center and the lane center. Because it is not possible to ensure that the vehicle center coincides with the lane center, so the standard deviation is more significant. According to Ma Jun’s research[6] , the larger the standard deviation, the greater the lateral position fluctuations, which reflects the driver's difficulty in driving, and the risk of collision with an adjacent lane or off the road.

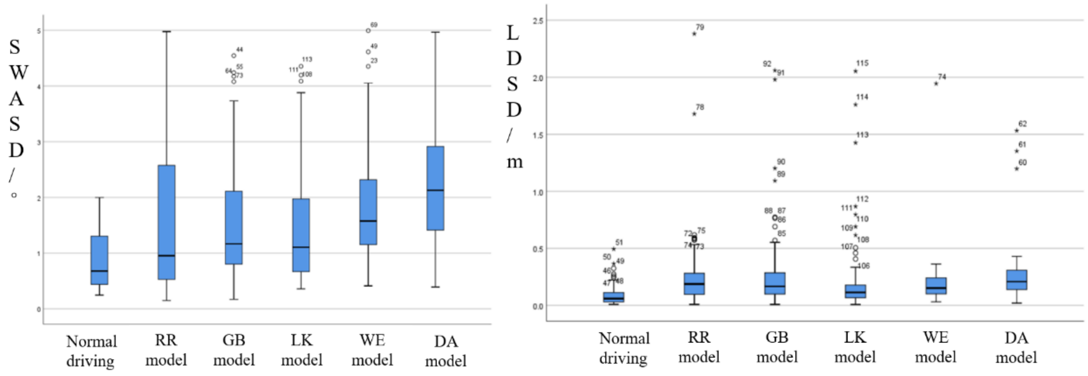

Fig.3 Box-plot of comparison results of SWASD and LDSD

The following Fig.3 shows the comparison results of SWASD and LDSD of different models when performing voice interaction secondary tasks with the normal driving. The results show that the SWASD and LDSD are higher than that of normal driving when the drivers finish secondary tasks using voice interaction systems. Significance tests are performed on the SWASD and LDSD during normal driving and when finishing secondary tasks using the voice interaction system for different vehicle models. The results are shown in Table 2 below:

Table.2 Significance tests of SWASD and LDSD

	RR model	GB model	LK model	WE model	DA model
SWASD	0.001	0.000	0.004	0.000	0.000
LDSD	0.000	0.000	0.004	0.002	0.000

According to the test results, for SWASD and LDSD, it can be found that when driving 5 models to complete secondary task using voice interaction system, the difference is significant($p < 0.05$), which is significantly different from the normal driving state.

4.2 Longitude directions

The driver controls the vehicle's longitudinal movement through the brake and accelerator pedal, so that the vehicle completes acceleration, uniform speed and braking. The driver will control the vehicle speed according to the external environment. The driver controls the vehicle's longitudinal movement through the brake and accelerator pedal, so that the vehicle completes acceleration, uniform speed and braking. The driver will control the vehicle speed according to the external environment.

The following Fig.4 is a comparison of the 30s segment of no secondary tasks and finishing secondary tasks through the voice interaction system when the driver is driving at the target speed of 45km / h. It can be seen that when driving with concentration, the speed is basically stable at the target speed, with occasional small changes. But during finishing voice interaction, the speed of the vehicle fluctuates significantly and rapidly.

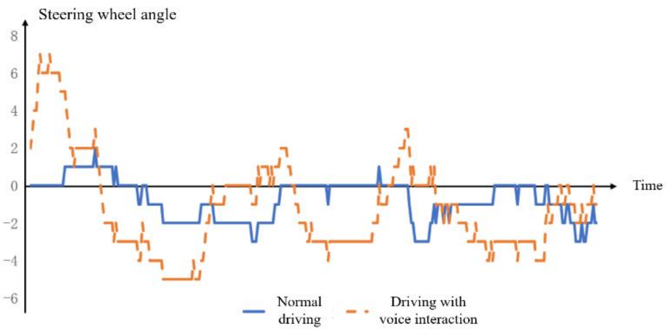

Fig. 4 Speed comparison

In the simulation environment, in order to approach the real road environment, urban roads, suburban roads and elevated roads are set up. Since the target speed is different in different road environments, this paper takes the difference between the actual speed and the target vehicle speed The absolute value of the value is used to characterize the longitudinal velocity. The Average Value of Speed Difference(AVSD) is calculated as follows:

$$AVSD = \frac{1}{n}\sum_{i=1}^{n}\left|v_i - v_{target}\right|$$

Among them, v_i refers to the longitudinal speed, and v_{target} refers to the target speed of the current road segment.

In order to investigate the longitudinal speed, the driving speeds of urban road sections are compared separately, as shown in Table 3. The comparison shows that the average speed of the driver will decrease during the secondary task. Researchers believe that the driver 's attention is distracted, which is a compensation behavior that should be dangerous because of distraction[7]. This is because when subjective cognition feels insecurity, it will subconsciously reduce the speed to reduce the stress of interaction.

Table.3 Average speed of different models in urban road

	Normal driving	RR model	GB model	LK model	WE model	DA model
v_{target}(km/h)	45.45	43.08	43.90	43.61	44.55	44.62

The standard deviation of speed (v_{sd}) is one of the important indicators for evaluating the change of the longitudinal speed of the vehicle. It represents the driver's ability to drive at a fixed speed, specifically the standard deviation of the sampling speed of the vehicle during the secondary task. The specific calculation method is as follows:

$$v_{sd} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(v_i - \bar{v})^2}$$

The following Fig.5 shows the comparison results of AVSD and v_{sd} of different models when performing voice interaction secondary tasks with the normal driving.

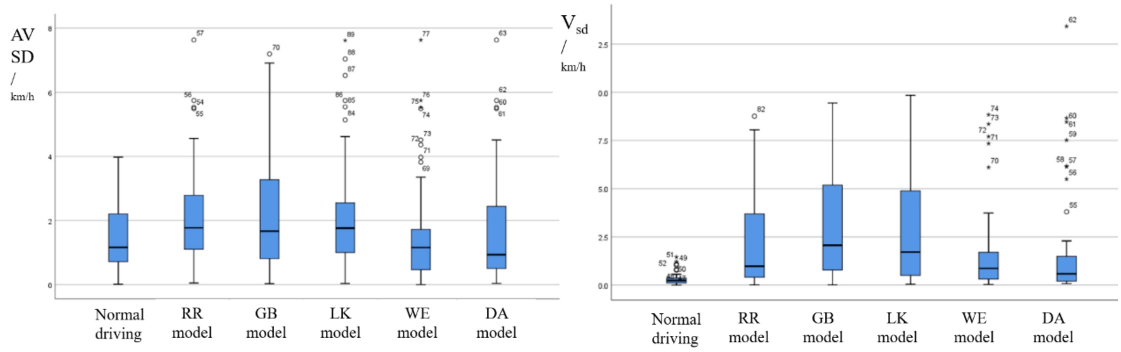

Fig.5 Box-plot of comparison results of AVSD and v_{sd}

From the data, it can be found that although the average difference in longitudinal speed difference is not large, the standard deviation of speed is obviously different from normal driving state, which shows that under the influence of secondary tasks of voice interaction, the driver will frequently modify the speed of the vehicle.

Significance tests are performed on the AVSD and v_{sd} during normal driving and when finishing secondary tasks using the voice interaction system for different vehicle models. The results are shown in Table 4 below:

Table.4 Significance tests of AVSD and v_{sd}

	RR model	GB model	LK model	WE model	DA model
AVSD	0.012	0.003	0.004	0.759	0.388
v_{sd}	0.000	0.000	0.000	0.000	0.000

From the significance difference results, it can be seen that the RR model, GB model, and LK model all have significant p-values less than 0.05, which is significantly different from normal driving, but the p-values of WE model and DA model are greater than 0.05, which is not significantly different from normal driving conditions. Therefore, AVSD cannot be used for the longitudinal performance when finishing secondary tasks using voice interaction. But for v_{sd}, the difference is significant($p < 0.05$), which is significantly different from the normal driving state.

5. Conclusion

This paper studies the voice interaction system of ICV in China market, and analyzes its impact on driving safety. Through the data, it is found that the use of voice interaction system to complete secondary tasks will have an impact on driving performance, and then affect driving safety. For lateral directions, fluctuation frequency and range of steering wheel angle during voice interaction is significantly higher than that normal driving. And SWASD and LDSD are significantly different from the normal driving state. For longitude directions, the speed of the vehicle fluctuates significantly and rapidly. And v_{sd} is significantly different from the normal driving state.

References

[1]Wang Wuhong, Sun Fengchun, Caoqi. Theory and Method of Driving Behavior in Road Traffic System [M]. Beijing: Science Press, 2001
[2]Yu Jiaying. Study on the characteristics of direct driving behavior of drivers at intersections [D]. Beijing：Tsinghua University. 2012.
[3]Wang Wuhong. Vehicle human-machine interaction safety and assisted driving [M]. Beijing People's Communications Press, 2012
[4]Braun M , Broy N , Pfleging B , et al. Visualizing natural language interaction for conversational in-vehicle information systems to minimize driver distraction[J]. Journal on Multimodal User Interfaces, 2019,13:71-88.
[5]Zhang Wentong. SPSS 11 statistical analysis tutorial. Advanced [M].Beijing: Hope Electronic Press, 2002.
[6]Ma Jun, Xu Maofei, Du Yuchun. A Usability Study on In-Vehicle Gesture Control. SAE Technical Paper 2016-01-1870, 2016, doi: 10.4271/2016-01-1870.
[7] Nelson E, Atchley P, Little T D. The effects of perception of risk and importance of answering and initiating a cellular phone call while driving[J]. Accident; analysis and prevention, 2009, 41(3):438-444.