



# A Literature Review of the Research on Interaction Mode of Self-driving Cars

Zhongshi Hu<sup>1</sup>(✉), Xin Xin<sup>1</sup>, Wanting Xu<sup>1</sup>, Yuwei Sun<sup>1</sup>,  
Zhenyu Jiang<sup>1</sup>, Xiangyu Wang<sup>1</sup>, Yishan Liu<sup>1</sup>, Siyao Lu<sup>1</sup>,  
and Min Zhao<sup>2</sup>

<sup>1</sup> Beijing Normal University, 19 Xijiekouwai Street, Beijing, China  
HZS13546325555@outlook.com, xin.xin@bnu.edu.cn,  
15856911002@163.com, 1216198339@qq.com,  
710922793@qq.com, miya\_hi@163.com,  
liuyishan1313@126.com, Lynn6917@163.com  
<sup>2</sup> Baidu.com, Inc., No. 10 Xibeiwang East Road, Beijing, China  
zhaomin04@baidu.com

**Abstract.** With the developing of the technology on self-driving, more and more L3 driverless vehicles are launched in market, people get opportunities to experience the self-driving cars in their daily life. As a result, there is a growing demand for the autopilot experience. Natural and efficient human-computer interaction can not only improve the driving experience, but also accelerate the process of self-driving commercialization. This paper discusses eye-movement interaction, voice interaction and gesture interaction in self-driving car, and analyzes the technology, advantages and disadvantages of the existing interaction modes, and prospect the future development trend of self-driving human-computer interaction.

**Keywords:** Self-driving · Human-computer interaction · Voice interaction · Eye-movement interaction · Gesture interaction

## 1 Introduction

With the progress of artificial intelligence research and the rapid development of sensor technology, the research field of self-driving cars is also expanding. In recent years, the research and development of self-driving technology is mainly divided into two camps: Tesla, Audi mainly car companies and Google, Baidu led by the Internet enterprises. At present, the research in the field of self-driving is more related to the application of self-driving vehicle obstacle avoidance algorithm and self-driving vehicle sensor, but the research on the internal human-computer interaction of self-driving car is not very in-depth [1]. The human-computer interaction system of self-driving cars is the last threshold for the commercialization of self-driving cars. With the development of eye tracking technology, speech recognition technology and gesture interaction technology, based on these technologies, the human-computer interaction of autonomous cars must be the focus of the next research [1]. Starting from three kinds of human-computer interaction technologies, this paper discusses the application and development of the three technologies at the present stage, and analyzes the advantages and disadvantages

of each technology at the present stage, and prospect the development trend of human-computer interaction in autopilot situation in the future.

## 2 Eye-Tracking Based Human-Computer Interaction

### 2.1 Technical Background

In recent years, many universities and research institutes in Chinese and abroad began to carry out eye-movement based human-computer interaction related research. Professor Zhu Xichan of Tongji University once proposed that the human eye can be used as a sensor in the driving process, which is equivalent to a part of human-computer interaction media. The eye's line of sight changes in the driving process and the region of interest can reflect the driver's intention to some extent [2].

### 2.2 Technical Principle

The fixation point of the human eye is determined by two factors: the head position and the eye position. The orientation of the head determines the range of gaze, while the precise direction of gaze is determined by the orientation of the eye, but it is limited by the orientation of the head. Stiefelhagen et al. classified visual tracking technology, mainly divided into hardware-based and software-based [3].

The basic principle of hardware-based visual tracking technology is image processing technology, which mainly records the change of line of sight by locking special cameras of both eyes, so as to achieve the purpose of the line of sight tracking. Nowadays most eye trackers are implemented in this way. This method requires the user to wear a specific helmet or fixture, which has great influence and interference. Software-based visual tracking technology mainly captures faces through cameras and then locates faces and eyes respectively. Through the algorithm, the head and eye trajectories are analyzed, and the gaze position is estimated. To track the line of sight.

### 2.3 Application of Eye-Movement Technology

**The Application of Eye-Movement Technology in Psychology.** Eye-tracking technology has been widely used in psychology and other fields. For example, Yan Guoli and others have used eye-tracking technology to study driving behavior under different road conditions and the psychological and physiological factors affecting their behavior [4].

**The Application of Eye-Movement Technology in Interaction Design.** In the early days, eye-movement interaction was used to assist people with disabilities in operating and interacting with computer devices, such as patients with ALS (Amyotrophic Lateral Sclerosis) disease. And Stephen Hawking's "talking eyes" and so on [5].

The Human-Computer Interaction Technology and Intelligent Information Processing Laboratory of the Chinese Academy of Sciences have carried out a number of national-level "multi-channel human-computer interaction interface" research projects.

Professor Tan Hao's team has carried out research on eye-movement interaction in-vehicle music player, used eye-movement interaction to control the car music, designed a set of interactive mode, and achieved good results [6, 7].

## 2.4 Application of Eye-Movement Technology in Self-driving

Eye-movement technology is mainly used in automatic driving from two aspects. One is to detect drivers' visual perception and driving behavior, and the other is to improve existing vehicle interaction through eye-movement technology.

**Eye-Movement Technology Analysis of Driving Behavior.** Most of the research focuses on psychology. The visual tracking technology of eye-movement is used to analyze driver's response and eye-movement law in different scenes. The aim is to reduce a driver's cognitive load and distraction in order to improve a driver's driving safety. Yan Guoli et al. used eye-tracking technology to study driving behavior under different road conditions [4]. Zhou Yang et al. constructed a cognitive distraction recognition model by simulating driving and eye-tracking technology and using the random forest method to study driver's distraction behavior [8]. Yuan Wei of Chang'an University studied three typical urban traffic environment characteristics, namely, channel width, running speed and traffic sign height, through the simulation test in the test site, and analyzed the changing rule of driver's dynamic vision under each characteristic condition in order to reduce the probability of traffic accidents in urban environment [9]. Guo Yingshi used the high-speed eye-tracking system to record the driver's dynamic eye-movement data such as gaze time, gaze target, scan time and scan speed during driving and analyzed the influence of traffic environment and driving experience on driver's eye-movement and workload [10]. Yang Meng et al. used eye-tracking technology to study the effects of different background music rhythm and lyric language familiarity on driving behavior and eye-movement. The results show that music rhythm has a significant impact on driving speed, eye jump and vertical search breadth. Compared with slow rhythm, fast rhythm has faster driving speed, shorter average eye jump distance and shorter vertical search breadth. The familiarity of lyric language has a significant effect on driving speed, error number and average gaze time. Compared with unfamiliar language, familiar language has a slower driving speed, more errors, longer average gaze time for novices and no influence for veterans. The experiment suggests that drivers choose music with unfamiliar language lyrics when choosing music, and the rhythm of music depends on the situation [11].

**Eye-Movement Technology Interaction.** Nowadays eye-movement technology can realize text input and control by human-computer interaction with the device through the user's gaze, eye jump, and smooth tracking. Gao Jun of Southeast University improved the vehicle navigation system by using eye-movement interaction and on-board head-up display. Li Ting of Hunan University redesigned the vehicle music application through eye-movement interaction, which greatly reduced the distraction of drivers [2].

## 2.5 Defects of Eye-Movement Interaction in Self-driving

**Accuracy.** According to the foregoing, line-of-sight tracking technology is divided into hardware-based and software-based. When hardware-based, the accuracy will be relatively high. At present, most laboratories use hardware-based eye tracker, but in practical application scenarios, the comfort and convenience of hardware wearing are low, which affects the experience. While software-based, although there is no physical device, comfort and convenience are very high, but the existing technology has low recognition accuracy, it is very difficult to get accurate focus [12].

**Midas Touch.** Midas contact means that if the cursor always follows the user's implementation, it will make the user feel bored, because the user may want to see anything at will, not necessarily "mean" anything, let alone start a computer command every time he changes his sight. Therefore, how to avoid Midas contact is a challenge in the future [13].

## 2.6 Prospects of Eye-Movement Interaction in Self-driving

Self-driving has developed rapidly in recent years. According to the standard of SAE (SAE J3016), the role of the driver will no longer exist in the L5 stage, and the traditional way of human-vehicle interaction will no longer exist. In the vehicle control, the car entertainment music function, navigation function and so on. Eye-movement interaction has broad prospects for development.

# 3 Voice Interaction in Self-driving

## 3.1 Technical Principle

The process of voice interaction includes automatic speech recognition (ASR), intelligent dialogue system and speech synthesis. Automatic speech recognition converts vocabulary content in human speech into computer-readable input, such as buttons, binary codes, or sequences of characters [15]. The intelligent dialogue system first understands the information conveyed by human beings, regards it as an internal state, then adopts a series of corresponding behaviors according to the strategy of dialogue state, and finally transforms the action into the expression of natural language [16]. Speech synthesis is a technique for producing artificial speech by mechanical and electronic methods. TTS technology (also known as text-to-speech technology) is a part of speech synthesis. It is a technology that converts computer-generated or externally-entered text information into an audible and fluent Chinese spoken language output [17].

## 3.2 Application of Voice Interaction in Self-driving

In addition to the visual channel, hearing is also one of the important sensory channels, which can obtain outside information without taking up the main channel of the driving

operation - the visual channel [17]. Because of this feature, voice interaction has become the mainstream in the automatic driving interaction mode.

Dong Changqing divides the voice interaction in automatic driving into two types. The voice system is placed in the smartphone and the voice system is configured in the vehicle terminal [1]. The voice system is placed in the smartphone, which uses the mobile phone as the carrier for voice interaction. Ford, Hyundai, General Motors, BMW and other car companies use this as a solution for in-vehicle voice interaction. Tesla, Audi, GM, Geely, and Changan have adopted the solution of the car terminal voice system. Representatives of the existing automotive voice interaction systems include Nuance's Dragon Drive voice assistant, the Cloudrive 2.0 system jointly developed by Chery and HKUST.

At present, the applications of voice interaction in automatic driving mainly include car navigation system, fatigue driving analysis, in-vehicle infotainment system, and emotional voice interaction system.

**Car Navigation System.** Liu Wang divides the car navigation voice interaction system into several steps: dialogue mode, keyword recognition, voice control command, name recognition, and speech synthesis. Liu Wang's experimental results prove that the system can meet the requirements of car navigation human-machine voice interaction [18].

The practical application of the in-vehicle voice navigation system is mainly embodied in: (1) Operation command input: control various instructions of the interface of each layer of the car navigation system, as long as the car navigation device starts menu, navigation, game, music and enters the speech recognition library and the navigation device can be operated freely with a simple input command. (2) Destination input: The building name and main traffic road of a specific city can be entered into the speech recognition library if the system allows. (3) Auxiliary Facilities Query: Find out the auxiliary facilities in the service stations, hospitals, restaurants, etc.

**Fatigue Driving Analysis.** Li Xiang believes that the current mainstream human fatigue detection method is based on facial performance characteristics and eye, but it still has limitations that are inconvenient to measure [19]. The driver's voice signal covers a large amount of human physiological or psychological state information, and its collection means is simpler and more convenient than other indicators, and the voice processing system has real-time, non-contact, strong environmental adaptability, and the noise reduction technology is mature.

Li Xiang proposed a method to detect driving fatigue using speech psychoacoustic analysis [20]. By processing the psychoacoustic model of the speech signal, the method can highlight the frequency components which are susceptible to fatigue and give more critical band descriptions, so that more detailed and intuitive fatigue information expression can be obtained.

**In-Vehicle Infotainment.** The voice interaction system is also applied to an in-vehicle infotainment system. In-Vehicle Infotainment is an in-vehicle information processing system based on body bus system and Internet service, which can realize car information, navigation, entertainment and other services [21]. Peng Yuyuan believes that

voice interaction can help drivers put visual attention on road driving and help drive safety.

**Emotional Voice Interaction System.** Emotional speech interaction system refers to the ability to give emotional recognition and expression to interactive systems, making human-computer interaction more humane [22]. In recent years, this field has received great attention.

In the 2000 study, V. Kostov and S. Fukuda developed VIS (Emotional Information Voice Interaction System), the ability to perceive users by investing in acoustic similarities in eight states of emotion (neutrality, anger, sadness, happiness, disgust, surprise, nervousness, distress, and fear). VIS and its research methods have laid the foundation for contemporary emotional voice interaction systems.

The experimental results of Won-Joong Yoon in 2007 show that using k-NN and SVM with probability estimation to estimate the degree of emotion, the accuracy rate in the five basic emotions (neutral, happy, sad, angry and troubled) is 72.5% [23]. It is possible to use the voice interaction to estimate the emotional level in the automatic driving in the later stage. In 2018, Li Zhenzhen studied the emotional design of voice assistants from the instinct, behavioral and reflective layers, and combined the speech interaction design with emotional design to provide new ideas and methods for the promotion of voice assistant experience.

### 3.3 The Challenge of Voice Interaction in Self-driving

According to the US J.D. Power agency released 2016 new car quality survey shows that among the many interaction problems, the voice interaction failure rate is as high as 23%. It can be seen that the voice interactive application in autonomous driving still has certain challenges. Dong Changqing believes that the main challenges are: (1) The low naturalness of voice interaction: the current voice commands have low recognition of local accents, proper nouns and flexible collocations, which limits the recognition of voice interaction. (2) Insufficient flexibility: Even if the accuracy of system identification is improved, the driver's spoof will affect the correct interaction. (3) Poor noise resistance [1].

Wang Haifeng believes that the challenge may also include (1) The sound environment in driving is difficult to be compatible: different acoustic signals are easy to cover each other. (2) The sound signal is time-sensitive: unlike visual information, if the auditory information cannot be accepted by the listener when it is generated, it will lose its usefulness [23]. Yezi analyze the problems of voice interaction from three levels of product design (operational layer, functional layer, and emotional layer) [24]. He believes that at this stage, the voice interaction does not liberate the hands well in the operation layer, and the recognition accuracy is not high in the functional layer. The emotional layer does not satisfy the user's perceptual psychological needs well.

At the same time, the results of 2001 John D. Lee, David L. Straye in 2014 and Douglas Getty in 2018 can prove that some voice-based interactions in the car may increase the driver's cognitive load, which is not good for traffic safety [25–27]. How to avoid this negative impact is also a research issue in the future of voice interaction in autonomous driving.

### **3.4 Prospects for Voice Interaction in Self-driving**

Li Zhiyong believes that the future development of voice interaction is divided into three phases. The first phase should aim to improve the accuracy of speech recognition in order to accurately respond to the user's voice input in a typical environment. This is also the focus of existing voice interaction research. The second stage reflects the personalization of voice interaction, which realizes the initial anthropomorphism and enhances the naturalness of voice interaction [28]. The third stage reflects the borderless nature of back-end content expansion. In essence, the future development of voice interaction is the process of digitization and intelligence, and the process of materialization in the form of ideals.

## **4 Gesture Interaction in Self-driving**

Gesture interaction is an interactive way to complete human-computer interaction by recognizing human gestures through mathematical algorithms.

### **4.1 Technical Background**

In recent years, with the development of gesture detection and recognition technology, a new way of interaction has emerged in the field of human-computer interaction, called gesture interaction. At present, gesture interaction is mainly used in the control of navigation and music playing interface. In a specific area of the car, the content of the interface is manipulated by specific gestures. Gesture interaction is a new interactive mode of human and vehicle, which is different from the traditional interaction mode of button operation and touch screen operation.

### **4.2 Technical Principles**

The realization of gesture interaction technology is mainly determined by two parts, one is the capture and tracking of gesture action, the other is the recognition and conjecture of gesture action.

The capture and tracking of gesture action is mainly realized by hardware foundation. According to the different hardware implementation, there are about three kinds of gesture recognition technology used in the industry at present. The first is structure light technology, through laser refraction and algorithm to calculate the position and depth information of the object, and then restore the whole three-dimensional space. The second is time of flight technology. A luminous element is loaded in the hardware, and the flight time of photons is captured and calculated by CMOS sensor. According to the flight time of photons, the flight distance of photons is calculated, and the depth information of objects is obtained. The third is Multi-camera, which uses two or more cameras to capture images at the same time. By comparing the differences between the images obtained by these different cameras at the same time, an algorithm is used to calculate the depth information. It can restore the entire three-dimensional space. The recognition and conjecture of gesture action is mainly realized by software. According

to the development of gesture recognition technology, there are template matching technology, neural network technology of statistical sample features and neural network technology of deep learning. The template matching method is mainly used in two-dimensional gesture recognition, and the recognition method based on neural network technology is mainly used to recognize three-dimensional gestures.

### 4.3 Application of Gesture Interaction Technology

**Application of Gesture Interaction in Non-driving Field.** Interaction for virtual environments. It is mainly used in virtual manufacturing, virtual assembly and game fields. Virtual manufacturing and assembly is the integration of design and manufacturing processes under a unified model. It integrates all kinds of processes and technologies related to product manufacturing into the solid digital model of three-dimensional, dynamic simulation process. As a result, the production is simulated more effectively and economically. The backtracking changes brought by the pre-design to the later manufacturing are effectively reduced. The production cycle and cost of the product are minimized, and the design quality of the product is optimized. Virtual manufacturing and assembly directly carry out the assembly of parts in the virtual environment through the movement of the hand. On the other hand, gesture interaction is also used in the field of somatosensory games. Microsoft, for example, released Kinect, a somatosensory device for use with XBOX360, in 2010. Kinect can recognize human movements, including gestures [29].

For sign language recognition. Sign language is the language of deaf and mute people. it is a relatively stable expression system composed of hand movements supplemented by expression postures. The goal of sign language recognition is to make the machine understand the language of the deaf, so that the machine is a sign language translation machine, which is convenient for deaf people to communicate with the surrounding environment. The process of deaf people communicating with the machine is the process of sign language recognition. It is also the process of gesture interaction [29].

**Application of Gesture Interaction in Self-driving.** The somatosensory interaction in the self-driving interaction mainly focuses on gesture interaction. Different from other traditional interaction methods (such as physical interaction and touch-screen interaction), gesture interaction is considered to be a more natural and more in line with people's own cognitive style way of human-vehicle interaction [30]. Gesture interaction has become the focus of automotive human-computer interaction interface design, because gesture interaction can reduce the visual distraction and cognitive burden of drivers [31]. At present, gesture interaction is mainly used in the interaction of vehicle navigation system, video and audio system, and the interaction of light and air conditioning system. Fang Yinan of Donghua University has designed and developed a set of map and music programs for on-board tablets, which use Leap Motion to capture and recognize gestures and complete the driver's control of maps and music through gesture movements, with a success rate of 90.5%. Nie Xin of Beijing University of Technology obtained the degree of driving distraction and driving load of each gesture by testing the brain power, vision and pressure of gesture operation, observing the lane,



speed and distance, and designed a set of operation gestures [31]. Li Moyang of Hunan University has studied and designed a vehicle-borne system based on gesture interaction, which can answer phone calls, play music, change radio stations and operate navigation maps through hand movements [32]. Yang of Hunan University applied situation awareness to the study of human-computer interaction in self-driving cars, and designed and produced HMI system, which proved the feasibility of gesture interaction as a natural interaction for secondary driving tasks in the car [33].

#### **4.4 Advantages of Gesture Interaction in Self-driving**

In Alpen's study, they explored on-board systems based on gesture control by comparing gesture-operated entertainment tasks with traditional radio tuning tasks. It is found that gesture operation can make drivers make fewer mistakes in driving tasks than traditional radio adjustments. By comparing the control mode of the gesture operation interface with the touch screen operation interface to the music player, Pirhonen found that the gesture operation interface significantly reduces the user's workload and task operation time. Li Moyang of Hunan University believed that gestures provide a new space for human-computer interaction in the car. It can alleviate the contradiction between the growing functions and the limited space in the car. Gestures provide the possibility for the realization of these functions, especially the operation of the secondary driving task function.

#### **4.5 Challenges of Gesture Interaction in Self-driving**

Dong Changqing of the China Automobile Research Center summarizes the shortcomings of gesture interaction: (1) The current gesture recognition technology is not mature enough. It is difficult to accurately complete gesture control; (2) The similarity between different gestures and the fuzziness of gesture range are easy to lead to systematic misjudgment; (3) The content that gesture can convey is limited. Compared with speech interaction, gesture interaction is difficult to express the specific task content. In addition to solving the above difficulties, the design of gesture interaction in the future should also be in line with people's cognitive habits. Because gesture interaction requires users to invest more learning cost than speech interaction and eye movement interaction, gesture interaction in accordance with users' cognitive habits can greatly reduce its learning cost and make the interaction process more natural. It can also save users' cognitive resources. At the same time, with the gesture design on the basis of simple and easy understanding, designer need to consider the cultural customs of different regions and design the gesture which is close to the daily natural communication posture of local users [1].

#### **4.6 Development Prospect of Gesture Interaction in Self-driving**

Jing Chunhui of Sichuan University believes that with the rapid development of self-driving technology and the upgrading of in-vehicle electronic equipment, the traditional interaction mode will account for a smaller and smaller proportion in the future self-driving interaction mode. On the other hand, with the saturation of visual channel

interaction design, the design of other sensory channels will be paid attention to. In this trend, gesture interaction will become one of the main stream of interaction design [34]. At the same time, with the application of gesture interaction on other mobile devices, desktop devices and object-linked devices, users will naturally apply gestures learned on one device to another platform, which will lead to the emergence of gesture standards. Although the research on gesture interaction of self-driving cars is in the development stage, there is no unified standard in the industry. But in the near future, there will certainly be similar organizations to produce and develop the corresponding gesture standards.

In the era when self-driving cars liberate drivers' hands from the steering wheel, gesture interaction will play a great role in how to complete the operation of audio-visual system, navigation system and light air conditioning system more efficiently and comfortably.

## 5 Conclusion

In the future, the main task of human-computer interaction in the self-driving is to integrate eye movement, voice, gestures and other interaction technologies to achieve efficient human-car interaction. A better human-computer interaction will lay a foundation for the commercial use of self-driving car. The final form of the development of human-computer interaction in the self-driving is that there is no interaction. The existing human-computer interaction technology is still a long way to get 100% natural interaction. The interactions are issued by human and executed by cars. If user makes a mistake in the interactive process, the car would not correctly complete the tasks, which is a big hidden danger to the safety of the whole driving environment. Li Deyi, an academican of the Chinese Academy of Engineering, said in a speech in November 2016 that "interactive cognition is very important for smart cars to become interactive wheeled robots." If the unmanned car lacks of effective interaction, passengers will think it is a ghost and afraid to ride. After the further developing of human-computer interaction in the self-driving situation, the natural human-computer interaction is mainly based on situational recognition. Cars automatic recognize the driving environment to predict the needs of users. Such a system can further reduce the operation of users and more intelligently meet the needs of users.

In a word, the trend of the development of human-computer interaction of autopilot situation is to achieve intelligent and humanized user needs safely and stably. Industry needs to apply vehicle engineering, industrial design, artificial intelligence, psychology and other fields of knowledge and skills to achieve this goal.

## References

1. Dong, C., Din, T., Huang, X., et al.: A survey of research on unmanned human-computer interaction. *Times Car*, 11–12 (2017)
2. Gao, J.: Research on human-computer interaction of electric vehicle navigation system based on eye-movement information. (Doctoral dissertation) (2017)

3. Stiefelhagen, R., Yang, J.: Gaze tracking for multimodal human-computer interaction. In: IEEE International Conference on Acoustics (1997)
4. Yan, G., Tian, H., Zhang, X.: Eye-movement study of automobile driving behavior. *Psychol. Sci.* **28**(5), 1211–1212 (2005)
5. Vidal, M., Bulling, A., Gellersen, H.: Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In: ACM International Joint Conference on Pervasive & Ubiquitous Computing (2013)
6. Feng, F., He, R., Tan, H.: Research on eye-movement interaction in vehicle oriented to information entertainment system. *J. Automot. Eng.* **5**(2), 108–115 (2015)
7. Feng, F.: Design and research of eye-movement interaction in vehicle music application. (Doctoral dissertation, Hunan University) (2014)
8. Zhou, Y., Fu, R., Yuan, W., Wang, D., Zhang, R.: Study on random forest model for driver cognitive distraction recognition. *J. China Saf. Sci.* (1) (2018)
9. Yuan, W., Guo, Y., Fu, R., Chen, Y.: Effects of urban road types on drivers' workload. *J. Chang'an Univ. (Nat. Sci.)* **34**(5), 95–100 (2014)
10. Guo, Y.: Study on the effects of traffic environment and driving experience on drivers' eye movement and workload. (Doctoral dissertation, Chang'an University) (2009)
11. Yang, M., Wang, J., Xia, Y., Yang, F., Zhang, X.: The influence of background music rhythm and lyric language familiarity on driving behavior and eye-movement. *Psychol. Sci.* (5), 1056–1061 (2011)
12. Zhang, L., Li, H., Ge, L.: Application of Tobii eye tracker in human-computer interaction. *Chin. J. Ergon.* **15**(2), 67–69 (2009)
13. Feng, C., Shen, M.: Application of eye-tracking in human computer interaction. *J. Zhejiang Univ. (Sci. Ed.)* **29**(2), 225 (2002)
14. Li, T.: Eye-movement interaction interface design and case development. (Doctoral dissertation, Zhejiang University) (2012)
15. Harris, R.A.: *Voice Interaction Design: Crafting the New Conversational Speech Systems*. Elsevier, Amsterdam (2004)
16. Basir, O.A., Miners, W.B.: Multi-participant, mixed-initiative voice interaction system: U.S. Patent (2014)
17. Lee, K.J., Joo, Y.K., Nass, C.: Partially intelligent automobiles and driving experience at the moment of system transition, pp. 3631–3634. ACM (2014)
18. Liu, W., Yang, D., Lian, X.: Implementation of vehicle navigation human-computer voice interaction system. *Electron. Prod. World*, 127–130 (2007)
19. Li, X., Li, G., Shi, J., et al.: Driving fatigue detection based on speech psychoacoustic analysis. *J. Instrum.* **10** (2018)
20. Car Infotainment System: CN201115295[P] (2008)
21. Zhang, S., Li, L., Zhao, Z.: Research progress on speech emotion recognition in human-computer interaction. *J. Circuits Syst.* 440–451 (2013)
22. Yoon, W.J., Cho, Y.H., Park, K.S.: A study of speech emotion recognition and its application to mobile services. In: Indulska, J., Ma, J., Yang, L.T., Ungerer, T., Cao, J. (eds.) *UIC 2007. LNCS*, vol. 4611, pp. 758–766. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-73549-6\\_74](https://doi.org/10.1007/978-3-540-73549-6_74)
23. Wang, H.: Human-computer interaction interface design in the cab of driverless electric vehicle (2017)
24. Ye, Z.: Research on the problems existing in human-computer voice interaction. *Technol. Commun.* **9** (2017)
25. Lee, J.D., Caven, B., Haake, S., et al.: Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers' attention to the roadway. *Hum. Factors* **43**, 631–640 (2001)

26. Strayer, D.L., Turrill, J., Coleman, J.R., et al.: Measuring cognitive distraction in the automobile ii: assessing in-vehicle voice-based. *Accid. Anal. Prev.* **379** (2014)
27. Getty, D., Biondi, F., Morgan, S.D., et al.: The effects of voice system design components on driver workload. *Transp. Res. Rec.* (2018). <https://doi.org/10.1177/0361198118777382>
28. Li, Z.: Anatomy of the hierarchy behind the voice interaction. *China Acad. J. Electron. Publ. House*, 14–15 (2019)
29. Li, Q., Fang, Z., Shen, M., Chen, Y.: Gesture recognition technology and its application in human-computer interaction. *Chin. Ergon.* **01** (2002)
30. Tan, H.: Research on automotive human machine interface design based on complex interaction context. *Packag. Eng.* **18** (2012)
31. Nie, X.: Research on the gesture operation of BMW in the user experience mode. (Master dissertation, Beijing University of Technology) (2015)
32. Li, M.: Research and application of user-defined gesture design for automobile human machine interface. (Master dissertation, Hunan University) (2015)
33. Yang, W.: Research of intelligent automobile human-machine interaction design based on situational awareness. (Master dissertation, Hunan University) (2017)
34. Jing, C., Zhang, J., Deng, C.: Vehicle gesture interaction design. *Packag. Eng.* **8** (2018)