Letter

Exploring the Effectiveness of Gesture Interaction in Driver Assistance Systems via Virtual Reality

Tong Liu, Mingwei Hu, Shining Ma, Yi Xiao, Yue Liu, and Weitao Song

Dear editor,

This letter presents a user study to explore the effectiveness of gesture interaction in driver assistance system (DAS). Distracted driving is a specific form of driver inattention and distraction occurs when the drivers' attention is diverted from the driving to other activities. According to the report of the National Highway Traffic Safety Administration [1], distracted driving is responsible for 15% of all motor vehicle accidents in USA, resulting in 3142 traffic fatalities and an estimated additional 424 000 people injured in motor-vehicle crashes in 2019 alone.

In-vehicle secondary tasks, such as answering the cell phone and adjusting the audio volume, are one of the main causes of the drivers' visual distractions and are responsible for 25% (no passengers) to 40% (with passengers) visual distractions [2]. Therefore, more considerations and carefulness are necessary when designing driver-operated human machine interfaces (HMIs). "Eyes on the road, hands on the wheel" is a crucial principle to be considered when designing the interface for DAS. Currently, the touch screen is the main interface of the auxiliary control in the car as a result of the popularity of smart mobile devices and the wide acceptance of touch-screen-based interactions [3]. However, such interfaces are inherently visually distracting because the driver needs to keep an eye on where they touch.

To this end, a promising solution that can provide less visual distraction to in-vehicle interface is that of micro gesture interaction. The micro gesture involves a micro-motion with minimized hand actions, allowing users to make any gestures that they want without paying extra attention to their hands. Compared with touch-based gestures and air gestures, micro gestures enable the drivers to keep attention on more important tasks due to the much shorter distance of the hand movement. So, micro gesture interaction is a very promising approach consistent with the design principle.

In this letter we set up an immersive virtual reality driving environment to simulate real driving conditions with unity and HTC VIVE PRO headsets. A driving test was conducted where 20 participants were asked to interact with the in-vehicle infotainment system using air gestures, micro gestures and a touch screen, respectively. Such performance-related metrics as driving efficiency and subjective load from the drivers were collected and analyzed in this study. During the experiments, the eye-tracking technology was also adopted to collect users' gaze points which can indicate the driving attention. Our results show that micro gesture interaction has great potential in the design of the interface of DAS.

Related work: The voice interaction interface is a more natural method with handless and eyeless interactions. As voice interaction is in line with human interaction habits, and the feedback is very clear, it is one of the most popular in-vehicle interfaces. However, the audio signals can be easily affected by the environment noise, especially in a noisy environment [4], indicating that the voice interaction has some limitations for the in-vehicle scenes. Recently, gesture capture devices such as Kinect [5] and Leap Motion Controller [6] have been widely used for gesture interaction because of their advantages such as high accuracy and small size. Using gestures as the input has also been explored in the automotive industry [7]. Many studies on gesture interaction have found that the evaluation score of the gesture interface is higher than that of traditional control methods, indicating that drivers prefer to use gesture interaction [8]. However, the main disadvantage of the in-vehicle gesture interface is the interference of large-scale hand movements with the ongoing driving behavior. Lack of adequate feedback also creates more visual distraction for the driver, which lowers the satisfaction, safety and usability while interacting with gestures.

It has been recognized that micro gestures can support direct and subtle interactions [9]. With the increasing applications of hand-tracking and electronic devices, micro gesture interaction provides a new approach to the interaction design for portable ubiquitous computing. Since drivers need to keep their hands on the steering wheel as much as possible to ensure driving safety [10], many researchers have studied micro gesture interaction while grasping. Considering the great mobility of hands and fingers, Sharma *et al*. [11] systematically investigated the effects of grasps and object sizes on user-elicitation gesture design. By taking both the gesture semantics and action trajectories into account, Xiao and He [12] explored how to use micro gestures to interact with the in-vehicle infotainment system efficiently and intuitively. Meanwhile, since gesture interaction requires some extra time to learn complex gestures [13], the designed gestures need to be "simple" and "natural".

Driving environment simulation: Our simulated interface includes the cell phone and audio interfaces, which allows drivers to manipulate infotainment functions via gestures and a touch screen. HTC VIVE PRO headsets were used to provide an immersive virtual reality driving environment and the Logitech G29 device was employed for drivers to complete driving tasks. A white panel was placed next to the steering wheel serving as the real benchmark for the virtual interface, which can provide proper haptic feedback when the user is interacting with the virtual touch screen. The eye-tracking module in HTC VIVE PRO was used to collect drivers' gaze diversion data for quantifying the driving distraction caused by the touch and gesture interaction. The overall workflow is shown in Fig. 1.

We set up a gesture dataset according to Chan *et al*. [14], in which all gestures were designed based on human factors and ergonomics. We conduct a user elicitation study where the participants were asked to choose the preferred gestures from the dataset based on the interaction function. We also calculate the agreement rates among participants and filter out the final selected gesture set, which includes 6 air gestures (gestures performed in the air) and 6 micro gestures (gestures performed while grasping the steering wheel). Fig. 2 shows the selected gesture set and the corresponding functions.

Experiment procedure: Lane-Change-Test [15] is a standard experiment for testing driving-related metrics, through which we can measure how the performance of a primary task (PT) declines while a secondary task (ST) is conducted. The PT was required to follow a predefined route on three lanes at a fixed speed of 60 km/h and change the lane according to the road signs. During the ST, participants were instructed to perform the interaction tasks through audio instructions.

Twenty volunteers participated in the study (12 male, 8 female) and they are all college students. The age of participants ranges from 22 to 30 years (Mean = 25, SD = 1.76). All participants had driving licenses and their driving experience range from 1 to 10 years (Mean = 4, SD = 1.92).

Each participant was required to complete the experiment under the guidance of the laboratory assistant. All participants need to

Corresponding author: Yue Liu and Weitao Song.

Citation: T. Liu, M. W. Hu, S. N. Ma, Y. Xiao, Y. Liu, and W. T. Song, "Exploring the effectiveness of gesture interaction in driver assistance systems via virtual reality," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1520–1523, Aug. 2022.

The authors are with Beijing Engineering Research Center of Mixed Reality and Advanced Display, the School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China (e-mail: liutong967@163.com; 718199906@qq.com; shining.ma@bit.edu.cn; yixiao0202@gmail.com; liuyue@bit.edu.cn; swt@bit.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2022.105764

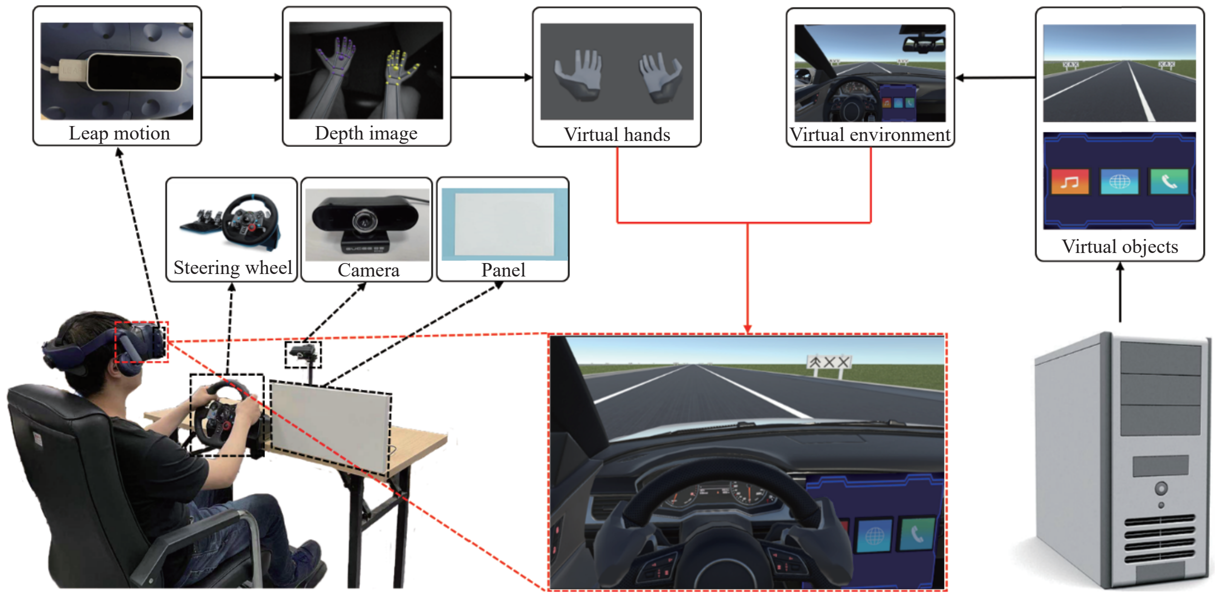

Fig. 1. System workflow. When the user is wearing the helmet mounted display (HMD), the Leap Motion camera can capture the depth image of the user's hands. At the same time, virtual objects, such as the virtual steering wheel and in-vehicle virtual interface, are created by a computer. We calibrate the position of the virtual interface in the simulated environment with the position of the white panel in reality. Finally, the hand image and the virtual objects can be fused into a virtual reality scene.

Air gesture code	Example	Micro gesture code	Example	Corresponding functions
Whole hand swipe right		Index swipe right		Answer the phone/ Switch to the next song
Whole hand swipe left		Index swipe left		Hang up the phone/ Switch to the previous song
Index circle clockwise		Index circle clockwise		Volume up
Index circle anticlockwise		Index circle anticlockwise		Volume down
Thumb point		Thumb tap		Turn into music menu/ Turn back to main menu
Thumb & Index tap		Index tap		Play/Pause music

Fig. 2. Gesture set and corresponding functions.

accomplish the pre-study questionnaire for his/her anonymous information, and sign the consent form voluntarily. Then the laboratory assistant helped the participants to learn how to perform the air gestures and micro gestures according to the video presentation. Before the formal experiment, a pre-test practice was completed to make participants get used to the simulated driving environment and the interaction operations. Then three groups of formal tests with air gestures, micro gestures and touchscreen were performed in random order. Each gesture was recorded by a camera as a video for the further analysis of the completion effect of the gestures. In order to make the experiment run smoothly, the Wizard of Oz [16] approach was adopted to ensure that users can get simultaneous feedback after the interactive task was completed. The experiment procedure is shown in Fig. 3.

To compare the performance of the three interaction methods, both objective and subjective measurements were recorded during the experiments. Both the driving efficacy and visual attention showed the influence of different interaction methods on driving distraction. All the data were automatically recorded in a log file. NASA task load index (NASA-TLX) was used to evaluate the task load of participants after each trial. Nonparametric tests were adopted in the analysis since the data did not pass the variance homogeneity test. The overall p-values were estimated using the Friedman test, and if significant, the Wilcoxon signed-rank test will be used to perform pair-wise comparisons.

Driving efficacy: We adopt 4 indicators to measure the driving efficacy of participants while using different interaction methods, whose definitions are as follow:

1) Experiment completion time: The total time between the participant starting the experiment and crossing the finish line.

2) Task completion time: The completion time of each interaction task, from the beginning of the interactive instruction to the end of the interactive operation by each participant.

3) Task completion rate: The average completion rate of each task.

4) Driving errors: Vehicle collisions and lane-changing errors during the experiment.

Fig. 4 shows that the experiment completion time of touch interaction (M = 537.83, SD = 51.14) was longer than that of air gesture interaction (M = 490.86, SD = 45.73) and micro gesture interaction (M = 472.32, SD = 49.84) in all tasks. Fig. 4(c) shows the difference in the completion time among the three interactive tasks, and it can be seen from Fig. 4 (c) that for each interaction task, the completion time of micro gesture interaction (M = 3.78, SD = 0.17) is shorter than that of air gesture interaction (M = 4.47, SD = 0.31) and touch interaction (M = 5.63, SD = 0.39), which means that micro gesture interaction has smaller distraction with the driver. In addition, compared with other interaction tasks, the "adjust the volume" tasks (including " Volume up" and " Volume down") take more time to complete. Some participants reported that it was caused by the operational inertia, while others reported that it was caused by the fact that they hoped to get more obvious auditory feedback.

Another phenomenon observed from this experiment was that sometimes participants did not perform interaction or perform it later when they were changing lanes or turning, whose reason was explained to prevent driving accidents. To measure the impact of this situation, the task completion rate was introduced to reflect the completion effect of interactive tasks. As shown in Fig. 4 (b), for touch interaction, the task completion rate is significantly lower than that of gesture interaction (Touch M = 93.06%, SD = 6.97%; Air gesture M = 99.17%, SD = 2.54%; Micro gesture M = 99.23%, SD = 2.12%; $p < 0.01$), while the air gesture interaction has almost the same task completion rate as micro gesture interaction ($p = 0.214$).

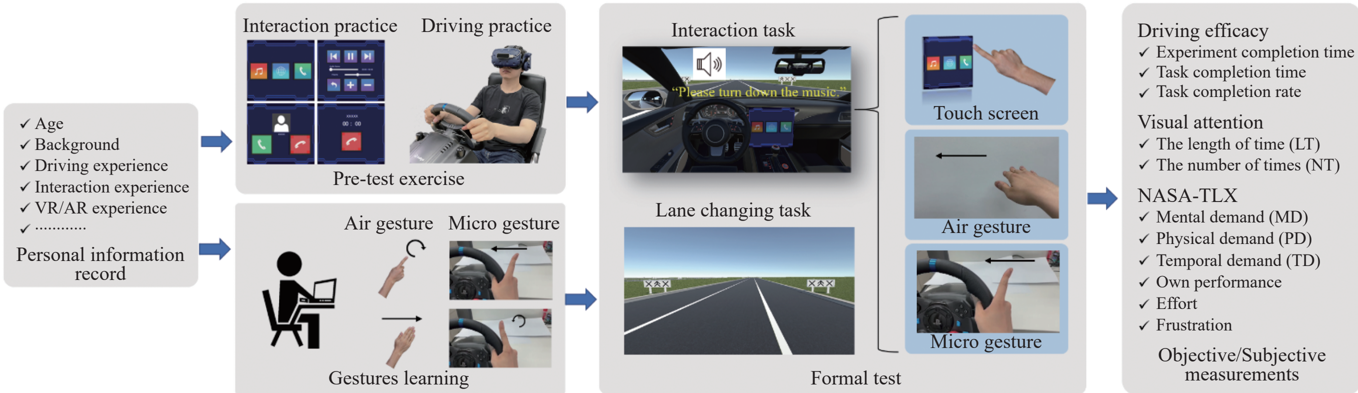

Fig. 3. Experiment procedure. The participants need to fill the pre-study questionnaire for his/her anonymous information, and sign the consent form voluntarily. Before the formal experiment, a pre-test practice was completed to ensure that the participants make participants get used to the simulated driving environment and the interaction operations. During the formal test, the participants were asked to finish both interaction task and lane changing task in three groups of tests. The order of interaction modes in three groups of tests is random.

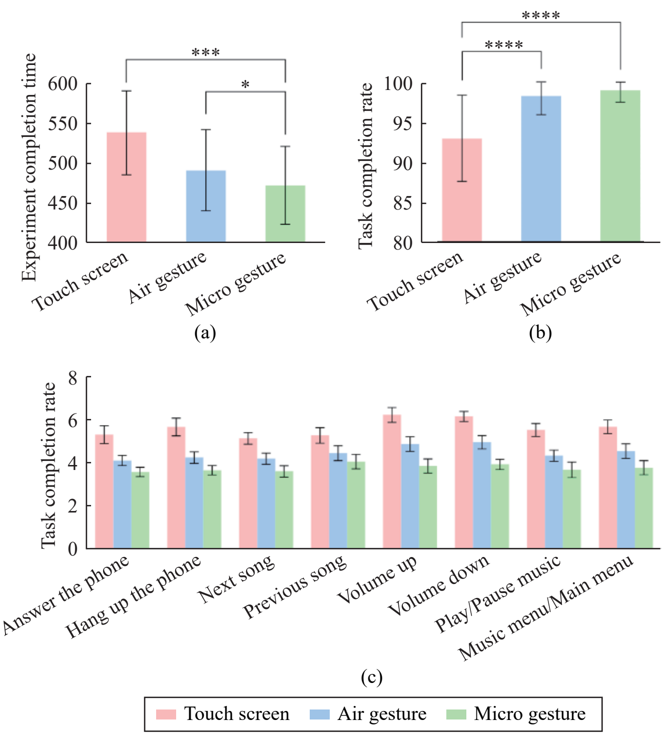

Fig. 4. Driving efficacy. (a) The total duration of the experiment in three interactive modes; (b) The average completion rate of each task in three interactive modes; (c) The completion time of each interactive task in three interactive modes. "*****" means that the significant difference between the two sets of data is extremely high. "***" means that the significant difference between the two sets of data is high. "*" means that the significant difference between the two sets of data is low.

In addition, the number of vehicle collisions and lane-changing errors occurred when performing interactive was also introduced as auxiliary indicators for the performance evaluation. Vehicle collision refers to a vehicle collision event caused by driving errors or interaction. The lane-changing error refers to the failure to change the lane as directed. The statistical results are shown in Table 1. There is no substantial difference in the number of vehicle collisions and lane-changing errors without performing any interaction task for the three interactive methods. When participants performed interactive operations, the number of vehicle collisions and lane-changing errors increased, indicating that the interactive operations distracted the participants from driving. When participants performed air gesture interaction and touch interaction, more errors occurred than that of the micro gesture interaction, indicating that these two interaction methods have a greater impact on participants' driving distraction. Some participants reported that sometimes they would ignore the ongoing driving operation during the interactive operation, but most participants would choose to complete the interactive task after completing the driving task.

Table 1. Statistical Results of Vehicle Collisions and Lane-Changing Errors (Over all 20 Participants)

Interactive method	Vehicle collisions while interacting	Vehicle collisions without interacting	Lane-changing errors while interacting	Lane-changing errors without interacting
Air gesture	17	10	13	6
Micro gesture	12	10	7	3
Touch	29	9	24	5

Visual attention: The eye-tracking technology of HTC VIVE PRO was conducted to collect the gaze diversion data of the participants while driving. We set up a safe gaze area (SGA) in front of the participants, including the front windshield of the car as well as the left and right rearview mirrors. The distribution of the participants' gaze position was collected and the length of time (LT) and the number of times (NT) when the gaze positions were not in SGA was counted. Fig. 5 shows the LT and NT in the three test trials. In terms of LT, touch interaction is significantly higher than air gesture interaction and micro gesture interaction ($p < 0.05$). Similarly, compared to air gesture interaction and micro gesture interaction, touch interaction is extremely significant in terms of NT ($p < 0.05$). This result shows that gesture interaction is more effective than touch interaction in reducing driving distraction. It can be seen from Figs. 5(a) and 5(c) that there is no significant effect between two gesture interaction methods on NT (Air M = 61, SD = 6.85; Micro M = 55, SD = 6.36; p = 0.24), but micro gesture interaction shows significance on LT (Air M = 72.86, SD = 18.45; Micro M = 61.87, SD = 16.32; $p < 0.05$), which means that micro gesture interaction can significantly reduce driving distraction time. Some participants reported that when they performed gestures, they would use peripheral vision to observe the gestures. We speculate that the operating position of micro gestures is closer to SGA, so the observation time for operating micro gestures will be shorter, thereby reducing the driving distraction.

For each interactive task, there is no significant difference in LT and NT between the air gesture interaction and micro gesture interaction. When completing the interactive task of "adjust the volume", the LT and NT are higher compared with other interactive tasks. The reason is that "adjust the volume" is a quantitative interactive task. Without a certain volume range, the auditory feedback of "adjust the volume" is not obviously enough, so

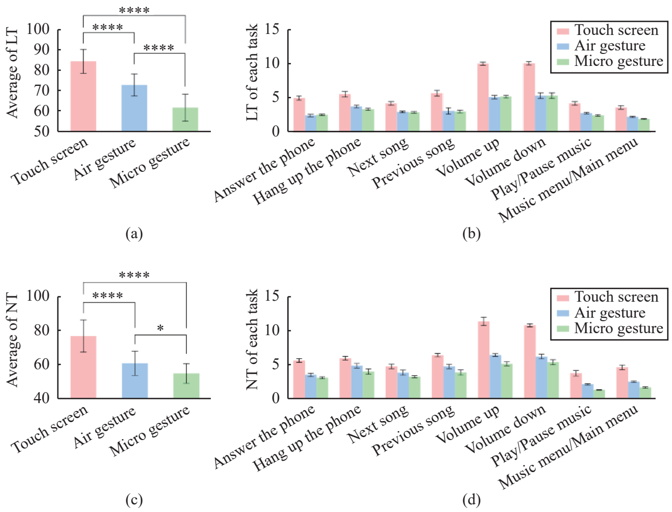

Fig. 5. Visual attention of three interaction methods. (a) Average LT of each interaction method; (b) Average LT of each task; (c) Average NT of each interaction method; (d) Average NT of each task.

participants have to visually determine whether they have completed the interactive task.

Subjective task load: In our experiments, the NASA-TLX was used to evaluate the subjective load of different interaction methods. The NASA-TLX provides an overall workload score based on a weighted average rating on six subscales: Mental demand (MD), physical demand (PD), temporal demand (TD), own performance, effort, and frustration [17]. In our experiments, the magnitude ratings on each subscale were set between 0 and 20. The overall task load score of touch interaction is higher than that of gesture interaction (Air gesture: M = 8.27, SD = 3.24, max load = 13.96, min load = 3.61; Micro gesture: M = 7.96, SD = 3.04, max load = 11.63, min load = 3.74; Touch: M = 9.64, SD = 2.78, max load = 15.63, min load = 3.50). More details are shown in Fig. 6.

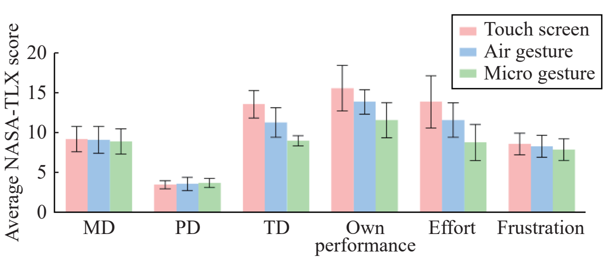

Fig. 6. The average task load of each subscale in three interaction methods.

The results show that the three interaction methods have significant differences in TD, own performance, and effort (TD: Air gesture M = 1.236; Micro gesture M= 1.174; Touch M = 2.009; $p < 0.05$. Own performance: Air gesture M = 1.856; Micro gesture M = 1.912; Touch M = 2.680; $p < 0.05$. Effort: Air gesture M=1.673; Micro gesture M = 1.129; Touch M = 2.352; $p < 0.05$). Regarding the TD, the participants stated that more visual attention was required in touch interaction when looking for the buttons to be operated with. Regarding the own performance, some participants reported that there were more problems with driving when using the touch interaction, and sometimes they were not satisfied with their performance. As for the difference in terms of effort, some participants explained that they had to pay more attention to the touch screen as more visual attention was distracted.

Conclusions: We conducted a user study to explore the effectiveness of gesture interaction in driver assistance systems via virtual reality to effectively reduce driving distraction. An immersive virtual reality driving environment based on users' driving habits was built up which fulfilled task requirements. Twenty participants were required to use air gestures, micro gestures, and a touch screen to interact with the in-vehicle infotainment system. The results show that gesture interaction can significantly reduce driving distraction. Micro gesture interaction can not only significantly reduce the time of driving distraction, but also improve the driving efficiency and the interaction efficiency of drivers. In addition, the overall task load rate of micro gesture interaction is also lower than that of air gesture and touch interaction. Our results indicate that gesture interaction methods, especially micro gesture interaction, have a promising prospect in the DAS.

Acknowledgments: This work was supported by the Key-Area Research and Development Program of Guangdong Province (2019B010149001), the National Natural Science Foundation of China (61960206007, 62002018, 62007001) and the 111 Project (B18005).

References

[1] National Highway Traffic Safety Administration (NHTSA), "Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices," 2019. [Online], Available: https://www.nhtsa.gov/risky-driving/distracted-driving.

[2] B. Metz, A. Landau, and M. Just, " Frequency of secondary tasks in driving – Results from naturalistic driving data," *Safety Science*, vol. 68, pp. 195–203, 2014.

[3] S. Rümelin and A. Butz, " How to make large touch screens usable while driving," in *Proc. 5th Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, 2013, pp. 48–55.

[4] Y. Saito, M. Itoh, and T. Inagaki, "Driver assistance system with a dual control scheme: Effectiveness of identifying driver drowsiness and preventing lane departure accidents," *IEEE Trans. Human-Machine Systems*, vol. 46, no. 5, pp. 660–671, Oct. 2016.

[5] Microsoft, "Kinect," [Online], Available: https://developer.microsoft.com/zh-cn/windows/kinect/. Accessed on: Jun. 2022.

[6] Leap, " Motion Controller," [Online], Available: https://www.ultraleap.com/product/leap-motion-controller/. Accessed on: Jun. 2022.

[7] K. May, T. Gable, and B. Walker, "A multimodal air gesture interface for in vehicle menu navigation," in *Proc. 6th Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, 2014, pp. 1–6.

[8] V. Laack and A. Walter, " Measurement of sensory and cultural influences on haptic quality perception of vehicle interiors," *BoD–Books on Demand*, 2014.

[9] G. Young, H. Milne, D. Griffiths, E. Padfield, R. Blenkinsopp, and O. Georgiou, " Designing mid-air haptic gesture controlled user interfaces for cars," in *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. 81, pp. 1–23, Jun. 2020.

[10] L. Angelini, F. Carrino, S. Carrino, M. Caon, O. Khaled, J. Baumgartner, A. Sonderegger, D. Lalanne, and E. Mugellini, " Gesturing on the steering wheel: A user-elicited taxonomy," in *Proc. 6th Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, 2014, pp. 1–8.

[11] A. Sharma, J. Roo, and J. Steimle, " Grasping microgestures: Eliciting single-hand microgestures for handheld objects," in *Proc. CHI Conf. Human Factors Computing Syst.*, 2019, pp. 1–13.

[12] Y. Xiao and R. He, " The intuitive grasp interface: Design and evaluation of micro-gestures on the steering wheel for driving scenario," *Univ Access Inf Soc*, vol. 19, p. 433, Apr. 2020. DOI: 10.1007/s10209-019-00647-0.

[13] B. Barbara, W. Jia, C. Deepa, M. Linda, C. Michael, and V. Federico, "Brain-based limitations in attention and secondary task engagement during high-fidelity driving simulation among young adults," *Neuroreport*, vol. 31, no. 8, pp. 619–623, 2020.

[14] E. Chan, T. Seyed, W. Stuerzlinger, X. Yang, and F. Maurer, "User elicitation on single-hand microgestures." in *Proc. CHI Conf. Human Factors Computing Syst.*, 2016, pp. 3403–3414.

[15] S. Mattes, " The Lane-Change-Task as a tool for driver distraction evaluation," in *Proc. Quality of Work and Products in Enterprises of the Future*, pp. 57–60, 2003.

[16] N. Dahlbck, A. Jnsson, and L. Ahrenberg, " Wizard of Oz studies," in *Proc. 1st int. Conf. Intelligent User Interfaces*, 1993, pp. 193–200.

[17] S. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proc. Human Factors & Ergonomics Society Meeting*, Oct. 2006, vol. 50, no. 9, pp. 904–908.