Implementation and Evaluation of Touch and Gesture Interaction Modalities for In-vehicle Infotainment Systems

Dan Zhao[1], Cong Wang[1,2], Yue Liu[1,2,3(✉)], and Tong Liu[1]

[1] Beijing Institute of Technology, Beijing 100081, China
liuyue@bit.edu.cn
[2] China Electronics Standardization Institute, Beijing 100007, China
[3] AICFVE of Beijing Film Academy, 4, Xitucheng Road, Haidian,
Beijing 100088, China

Abstract. The arrival of the Internet of Vehicles has promoted the transformation of in-vehicle human-computer interaction. As a new way of human-computer interaction, gesture interaction plays an important role in such interaction and technology update. This paper proposes the implementation of an in-vehicle driving simulator that allows infotainment system to be controlled by dynamic hand gestures and touch screen. Thirty participants were asked to interact with the infotainment system in two ways separately and randomly while performing the driving test. Moreover, three kinds of data related to the degradation of the driver's performance, including driving efficacy, visual attention and subjective workload were collected. In addition, the fuzzy comprehensive evaluation method is used to evaluate the usability of interactive gestures. The experimental results show that compared with the touch screen interaction, gesture control significantly reduces the driver's distraction and improves safety during driving.

Keywords: Driving simulation · Gestures · Touch · Driving distraction · Gesture usability

1 Introduction

With the development of intelligent automobile industry, vehicle products gradually developed from traditional means of transportation into a multi-functional, informational and intelligent human-machine interaction system. The introduction of information systems allows drivers to handle other events while driving the car, such as answering calls, turning on navigation devices, adjusting music and so on. However, the enrichment of functions increases the driving risk while providing convenience to the driver. To find an interactive modality that not only satisfies the driver's interaction demand but also minimize the driving risk is of great significance.

Commonly used in-vehicle interaction methods mainly include the button and knob control, which are the most widely used interaction method at present. Touch screen interaction is an interactive way developed with the promotion of a new generation of mobile smart devices. Due to the high penetration rate of mobile design, users are more

likely to accept such interaction [1]. However, the two interaction modes mentioned above may distract the driver and increase the driving risk. Speech is a more acceptable way of interaction with hand-free and eye-free, thus is one of the most popular modalities in use. However, the speech semantic recognition accuracy needs to be improved. At the same time, speech is susceptible to the environment, especially when the environment is noisy, the command understanding ability will be reduced [2]. In recent years, gesture acquisition devices such as Kinect and Leap Motion have been widely used in the field of human-computer interaction with the advantages of high precision and small size, which laid a technical foundation for the application of gesture control in the car [3, 4].

In this paper, we conducted an experimental evaluation of gesture interaction in a driving simulator and compared it to the direct touch interaction. The influence of the two interaction modalities on driving performance is summarized by analyzing the completion of driving tasks and eye tracking data. The driving behavior under different road conditions is analyzed by comparing to the existing works [5–7]. The driver's acceptance of gesture control in the vehicle is also studied.

2 Experiment Design

2.1 Driving Environment Simulation

In the driving environment, we usually refer to other human-vehicle interaction behavior except driving tasks as sub-tasks. When studying the completion of sub-tasks and the influence on the driver's attention under different interaction modalities, the simulated driving cockpit is usually used, so that the subjects can complete the driving task in the simulated environment.

Fig. 1. Simulated driving environment.

Our simulated driving platform (Fig. 1) allows the driver to manipulate infotainment functions via gestures and touch screen. The drivers complete the driving task through LG29 device. A Surface Pro mounted on the right side of the steering wheel serves as the center console and runs an application with typical infotainment scenarios like phone and music. Figure 2 shows the phone and music interfaces of the infotainment system. A Leap Motion controller is placed at the front of the Surface Pro to capture and recognize the user's gestures. In order to quantitatively evaluate the influence of touch and mid-air gestures on the performance of the driver, the SMI eye tracker placed in front of the display is used to measure the driver's gaze diversion data.

Fig. 2. Infotainment interfaces. (a) Menu interface. (b) Incoming phone. (c) Calling interface. (d) Music interface.

2.2 Experimental Task Design

The task for each participant was to complete the interactive task as efficiently and quickly as possible while driving the vehicle. Since the driving skills vary greatly among the subjects, a within-subject design was used, which compensated for fluctuations in performance between subjects. Inspired by the Lane Change Test (ISO 26022 standard) [7], the driving task is designed to include four separate tracks with two different road conditions (RD). For road condition 1, almost no road barriers are set in Track 1 and Track 2, but for road condition 2, a variety of different continuous road barriers as shown in Fig. 3 was set in Track 3 and Track 4. Each participant can familiarize with the road conditions in Track 1 and 3 and perform interactive tasks in Track 2 and 4 (For the convenience of description, we will simply refer to Track 1 and Track 3 without interactive commands as Track 1-N and Track 3-N). The design and implementation of the driving scene was developed with Unity3D software.

Four gestures (as shown in Fig. 4) were used in the experiment to complete the interactive task, i.e. answer/hang up the phone, turn up/down the volume and switch to the next/previous song. In order to reduce the user's memory demands, the gestures 'Swipe Right' and 'Swipe Left' were multiplexed, when the system has a telephone access, the user can complete the operation of answering/hanging up the call by Swipe Right/Left, and when the system is in the music playing mode, the Swipe Right/Left gesture can help the user to complete the operation of switching to the next/previous song. Since the system would pause the music being played while the phone was connected, the multiplexing of gestures does not cause any conflict. Each gesture was identified with the help of the Leap Motion SDK.
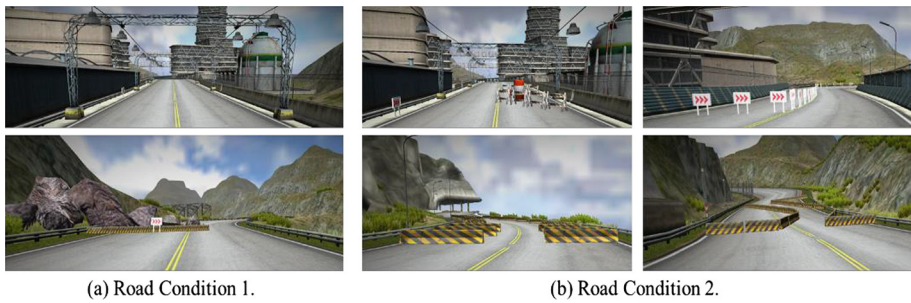

(a) Road Condition 1.   (b) Road Condition 2.

Fig. 3. Different road conditions.

(a) Swipe Right   (b) Swipe Left   (c) Clockwise   (d) Counter Clockwise
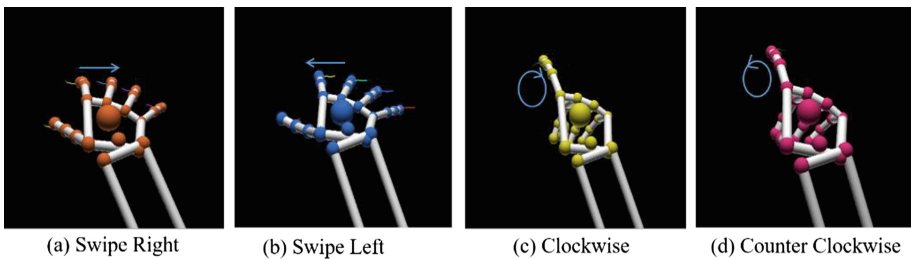
Fig. 4. Four hand gestures used in gesture interaction. (a) Swipe Right: answer the phone or switch to the next song. (b) Swipe Left: hang up the phone or switch to the previous song. (c) Clockwise: turn up the volume. (d) Counter Clockwise: turn down the volume.

2.3 Experiment Procedure

30 participants (17 males, 13 females) between 21 and 30 years (M = 24, SD = 1.94) were recruited for the experiment. Each subject would complete an experiment consisting of four parts with a total time span of one hour under the guidance of the experimental assistant. A pre-test exercise to ensure that the subject is familiar with the simulated driving environment and operation should be completed first. Then participants performed two test trials with touch and gesture interaction in random order. During the driving part, they were instructed to perform gestures through text-to-speech

output. The instructions were 'Answer/hang up the call' (occurred four times during each road condition), 'Switch to the next song' (three times each), 'Switch to the previous song' (two times each) and 'Adjust the volume' (three times each). During the gesture interaction, the gesture was executed once and the corresponding interactive task was completed once toon. In particular, the volume changes range from 0-1, the gesture was completed once and the volume changes by 0.1. The order of instructions was the same over all participants. After each trail, all participants were asked to evaluate the task load using NASA task load index (NASA-TLX). Finally, participants gave an overall rating for the usability of gesture interaction.

3   Results and Discussion

3.1   Driving Efficacy

Driving efficacy indexes are mainly reflected in experiment completion time, interactive task completion rate and autonomous interaction. Regarding all the tasks, the experiment completion time for the touch was larger than that for the gesture as shown in Fig. 5. However, the significance test results showed that both interaction modalities had no significant difference in the experimental completion time under various road conditions This is caused by the fact that the difference in driving habits between participants leads to the variation of driving speed and time.
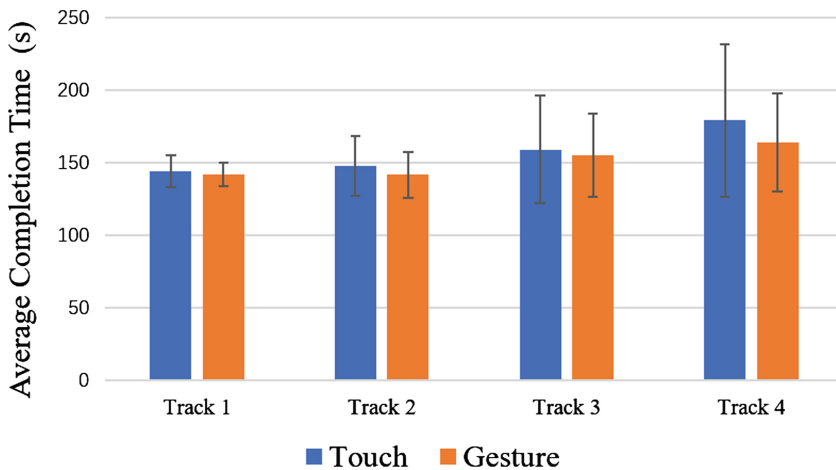

Fig. 5. Average experiment completion time of 30 participants for each track.

Another phenomenon that we found is that some participants chose not to execute interactive instructions when road conditions are complex to ensure driving safety. To account for this situation, the interactive task completion rate index was defined. As can be seen in Fig. 6, for road condition 1, the average task completion rate of the touch interaction was almost the same as that of gesture interaction (Track 2: Touch Mean,

M = 97.50%, Standard Deviation, SD = 6.97%; Gesture M = 99.17%, SD = 2.54%, F-value, F = 1.513, p-value, p = 0.224), however, for road condition 2, the average task completion rate of gesture interaction is significantly higher than that of the touch (Track 4: Touch M = 93.06%, SD = 8.21%; Gesture M = 98.61%, SD = 3.84%, F = 11.262, p < 0.01). The difference implies that the touch screen interaction has a high visual occupancy rate. For gesture interaction, an experienced driver can control the steering wheel with one hand while performing the driving task with the other hand.
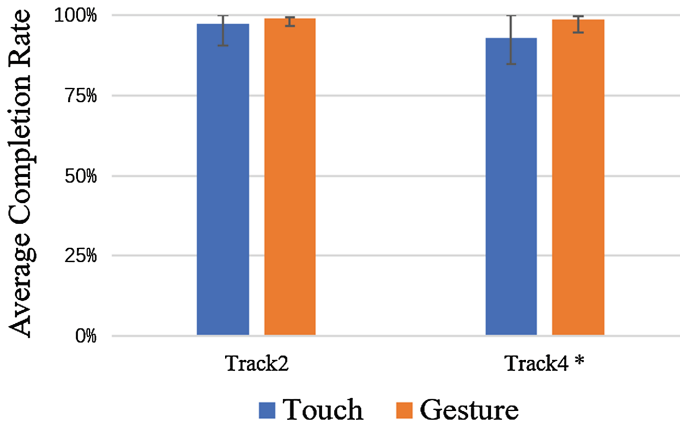

Fig. 6. Interactive task completion rate for different road conditions.

Other parameters, such as the number of accidents that occurred while performing interactive tasks and the frequency of autonomous interactions (Autonomous operation means the user's active interaction with the infotainment system without any interactive instructions.) are also counted as auxiliary indicators for experimental evaluation. The statistical results are shown in Table 1. It can be seen from the table that the accidents (over all 30 participants) caused by gesture interaction is lower than that of the touch under the simple road condition. For the complicated road conditions, although the number of accidents of the two interactions is similar, the task completion rate of touch is lower.

Table 1. Statistical results of auxiliary indexes (over all 30 participants).

Interactive modalities	Road conditions	Accidents	Autonomous operation
Touch	Track 2	12	21
	Track 4	12	3
Gesture	Track 2	6	22
	Track 4	14	8

3.2   Visual Attention

In order to analyze the driver's visual attention, SMI eye tracker is used to collect the subjects' gaze diversion data during the experiment. Through the eye tracker, we obtained the dwell time of primary visual attention lobe (PVAL), which refers to the road area that the driver pays attention to when there is no interactive task. Figure 7 shows the average dwell time for four tracks in two interaction modalities. The significant difference in dwell time occurs on track 2 for road condition 1 (Track2: Touch M = 54.64%, SD = 13.50%; Gesture M = 61.93%, SD = 13.39%; F = 4.407, $p < 0.05$), which means that under simple road condition gesture interactions reduce the distraction of the driver's attention relative to touch. For Track 4, the average dwell time for gesture interactions is higher, however there is no significant difference relative to touch. (Track4: Touch M = 57.13%, SD = 13.91%; Gesture M = 60.63%, SD = 14.29%; F = 0.920, p = 0.341) The reason for this phenomenon is that the task completion rate of the touch interaction on Track 4 is low. As mentioned in Sect. 3.1, the participant chooses not to execute interactive instructions to ensure driving safety.
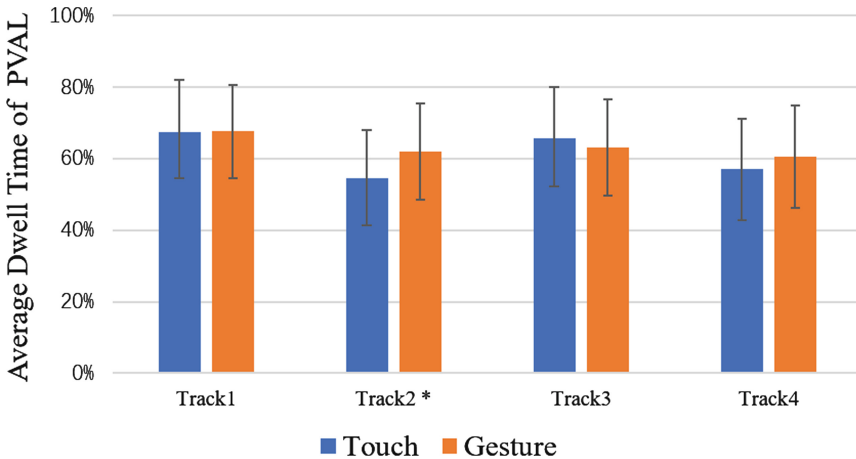

Fig. 7. The average dwell time of PVAL for four tracks in two interaction modalities.

The effects of sub-tasks on visual attention under different road conditions are illustrated in Fig. 8. It can be seen from Fig. 8 that for the touch screen interaction, the interactive task has a significant distraction of the driver's attention under both road conditions (Touch: Road Condition 1 F = 12.690, p < 0.01; Road Condition 2 F = 5.512, p < 0.05). However, for gesture interactions, results are completely opposite (Gesture: Road Condition 1 F = 2.810, p = 0.099; Road Condition 2 F = 0.529, p = 0.470). This result indicates that gesture interaction is superior in maintaining attention and is less affected by road conditions.
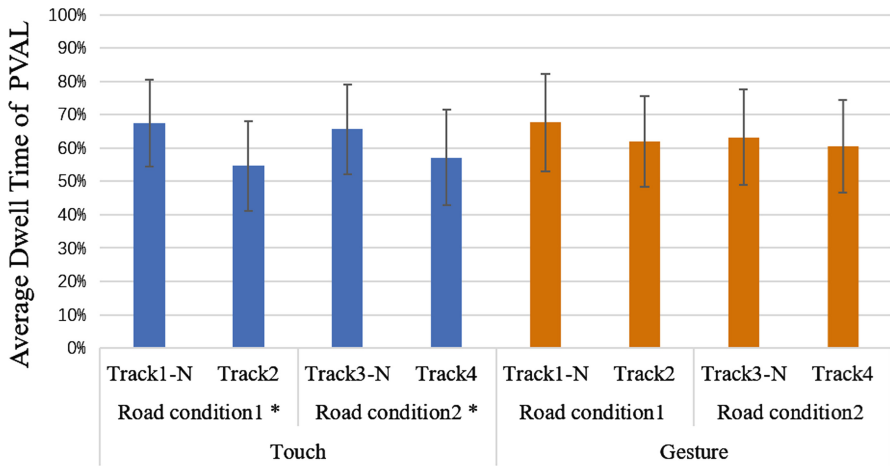
Fig. 8. Visual attention under different road conditions.

3.3 Subjective Task Load

In our experiments, the NASA-TLX is used to evaluate the interaction load. The NASA-TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Own Performance, Effort and Frustration [8]. The magnitude ratings on each subscale were set to 0 to 20 in our experiments. The overall task load score of touch interaction is higher than that of the gesture (Touch: M = 9.72, SD = 2.96, Max Load = 14.47, Min Load = 3.60; Gesture: M = 8.17, SD = 3.16, Max Load = 13.87, Min Load = 3.33). More details are shown in Fig. 9, which show that the two interaction methods only have significant differences in TD and Effort (TD: Touch M = 2.009; Gesture M = 1.236; $p < 0.05$. Effort: Touch M = 2.680; Gesture M = 1.856; $p < 0.05$). Regarding the TD, the participants indicated that they wanted to complete the instructions as soon as possible during the touch screen interaction, so that they could refocus their attentions on the driving task to prevent security problems, but they did not have such concerns for gesture interaction. For the effort difference, the participants explained that they need to pay more effort into the touch interaction.

3.4 Gesture Usability

In order to evaluate the usability of gesture control, we chose the usability principle proposed in [9], which evaluates each gesture from the following four dimensions: easy to learn and remember, effective, intuitive and comfortable and natural. We also added a 5th dimension indicator of the overall rating for the gesture interaction. The score for each dimension was set to 0 to 10. After counting all the participants' scores on gesture interaction, we found that the average overall rating was 7.73 (SD = 1.62), which means that gesture control can help the subject to complete the interaction
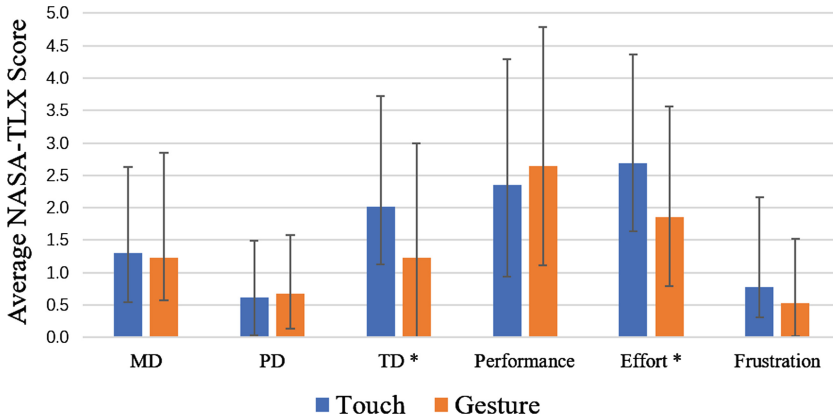
Fig. 9. The average task load of each subscale in two interaction modalities.

requirements. Figure 10 illustrates the average score of each gesture in each dimension. Based on the results, we can find that in terms of comfortable and natural, the scores of each gesture are relatively lower than other dimensions. We believe that this phenomenon is caused by the fact that participants use gesture interactions less frequently in their daily lives. In addition, for the Swipe Left gesture, in order to avoid misrecognition of the Leap Motion, the user needs to wrap the hand to the left side of the device and then perform the gesture, which increases the user's discomfort. For the volume adjustment gesture, the user needs to perform the gesture multiple times to get the expected volume level.
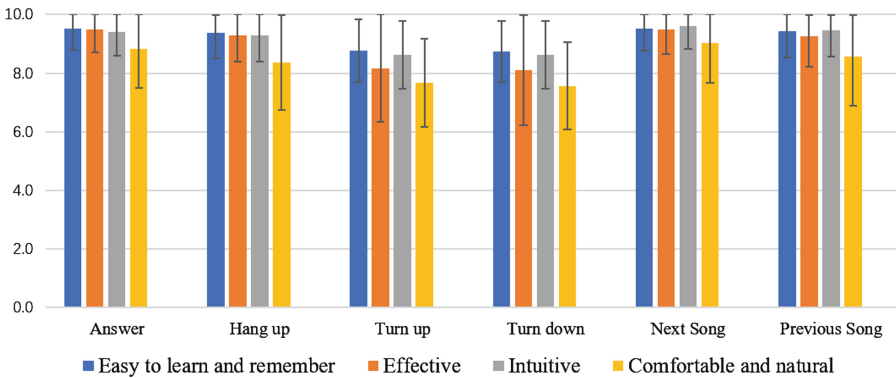

Fig. 10. The average score of each gesture in each dimension.

To make the evaluation results more significant, we use the fuzzy comprehensive evaluation method to quantify gesture preference degree. The evaluation results are shown in Table 2. Through these, we found that the quantified result of the gesture preference degree is consistent with the usability rating results, similar to the

phenomenon that the volume adjustment gesture has a relatively low score, and the user's preference for the gesture no longer tends to be excellent.

Table 2. Fuzzy comprehensive evaluation results.

	Excellent	Good	Medium	Bad	Low
Answer: Swipe right	85.27%	13.15%	1.58%	0.00%	0.00%
Hang up: Swipe left	77.39%	20.31%	0.77%	1.53%	0.00%
Turn up: Clockwise	47.20%	44.79%	6.41%	0.78%	0.82%
Turn down: Counter-clockwise	46.48%	44.73%	7.20%	0.77%	0.82%
Next song: Swipe right	86.00%	12.40%	1.60%	0.00%	0.00%
Previous song: Swipe left	78.83%	18.84%	0.78%	1.55%	0.00%

4   Conclusions

This paper presents a user study that compared the differences between gestures interaction with infotainment system and touch screen interaction in a simulated in-vehicle environment. The experimental results show that gesture interaction benefits from touch which has a better overall impression and better interaction efficiency. The efficacy index shows that although there is no significant difference in the completion time between the two interaction modalities, in terms of task completion rate, gesture interaction shows obvious advantages in complex road conditions, which means that gesture interaction can help drivers reduce distraction. Further proof of such conclusion is reflected in the result of visual attention analysis. By counting the average dwell time, we first conclude that under simple road conditions, gesture interactions can help drivers maintain their attentions on the road. Then we calculate the influence of the sub-task on the dwell time under different road conditions, and get the opposite result of the two interactions, that is, for touch interaction, the interactive task has a significant distraction of the driver's attention under both road conditions, but no significant difference with gesture. Evaluation of another indicator shows that the overall task load rating of touch interaction is higher than that of gesture, especially in the dimensions of time demand and effort. Finally, the usability of gestures is also considered, and different gesture solutions for each secondary task were comparatively analyzed based on the fuzzy comprehensive evaluation method.

In future work, simpler gestures should be explored to support gesture interactions for more control without increasing task load. At the same time, the gesture interaction system also needs to provide more gestures to users, for users to select a gesture that conforms to their interaction habits.

References

1. Kim, H., Song, H.: Evaluation of the safety and usability of touch gestures in operating in-vehicle information systems with visual occlusion. J. Appl. Ergon. **45**(3), 789–798 (2014)
2. Neßelrath, R., Moniri, M.M., Feld, M.: Combining speech, gaze, and micro-gestures for the multimodal control of in-car functions. In: International Conference on Intelligent Environments, London (2016)
3. Manawadu, U.E., Kamezaki, M., Ishikawa, M., et al.: A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle. In: IEEE International Conference on Systems, Budapest. IEEE (2017)
4. May, K.R., Gable, T.M., Walker, B.N.: A multimodal air gesture interface for in vehicle menu navigation. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications – AutomotiveUI (2014)
5. Angelini, L., Baumgartner, J., Carrino, F., et al.: Comparing gesture, speech and touch interaction modalities for in-vehicle infotainment systems. In: Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine on - IHM 2016 (2016)
6. Häuslschmid, R., Menrad, B., Butz, A.: Freehand vs. micro gestures in the car: driving performance and user experience. In: 3d User Interfaces, Arles (2015)
7. Kopinski, T., Eberwein, J., Geisler, S., et al.: Touch versus mid-air gesture interfaces in road scenarios - measuring driver performance degradation. In: IEEE International Conference on Intelligent Transportation Systems, Rio de Janeiro (2016)
8. Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors & Ergonomics Society Annual Meeting, vol. 50, no. 4661, pp. 904–908 (2006)
9. Ma, J., Du, Y.: Study on the evaluation method of in-vehicle gesture control. In: IEEE International Conference on Control Science & Systems Engineering, Beijing. IEEE (2017)