Universal Access in the Information Society (2020) 19:433–450
https://doi.org/10.1007/s10209-019-00647-0

LONG PAPER

The intuitive grasp interface: design and evaluation of micro-gestures on the steering wheel for driving scenario

Yiqi Xiao[1,2] · Renke He[2]

Published online: 5 April 2019

Abstract

Gestural inputs are nowadays widely applied to in-car interactive systems. The emerging sensing technologies allow for micro-gestures that can be achieved with less energy and are performed probably when the driver is grasping the steering wheel, thereby reducing the user attention to human–vehicle interaction. The movability of hands and fingers has to be considered before designing micro-gestures for driving scenarios, thus the newly defined gestures may be utterly different to the familiar multi-touch or body gestures. This paper presents a set of micro-gestures which are designed for intuitively commanding the in-car information system by taking both the gesture meanings and physical limitations into account. The study sets out the results of evaluating the feasibility of gesture sets with its performance in dual-task situation. The learnability and intuitiveness of micro-gestures as well as the effects they have on driving tasks are evaluated, and the effects of different grasping postures on the performance of novice drivers are also compared. It is concluded that the micro-gestures particularly designed for the grasp interface have advantages in multi-tasking, and they are appreciated by users who participated in the evaluation test.

Keywords Micro-gesture · Grasp · Gesture set · Evaluation · Dual-task

1 Introduction

A micro-gesture is a micro-motion that is defined as gestural input to minimize motor effort of manual actions. Comparing with the touch-based gestures for surface computing [34] and the three-dimensional gestures for TV control in a lean-back environment [38], the moving distance of a micro-gesture is evidently shorter. It does not require users to move the whole hand [37] so that they can take short amount of time [5] to quickly access the devices. With the development of pervasive computing, there are increasing concerns about such microinteractions. The computing power of wrist-worn personal devices [18, 22] and wearable rings offers an opportunity to regard the tiny but intentional manual actions which demand a little mental and motor resources as valid inputs.

Micro-gestures enable interacting with devices simultaneously while engaging in a more important work, since the less effort for gesturing will minimally interrupt that primary task. In our routine life, driving (including bicycle riding) is one of the typical scenarios where such multi-tasking is needed. To ensure safety, the driver has to keep his or her hands on steering wheel when, however, the fingers are able to move within certain limits. In consideration of this, researchers have made attempts to define usable micro-gestures to promote user experience of in-car environment, especially about the navigation and infotainment [1].

To hold the position, the design of grasping gestures can only use small variations in configuration and movement to achieve its goal, though these gestures should not be too similar to some subconscious body movement, such as tapping or wrist rotation [20], to prevent the natural motions from being erroneously recognized as meaningful inputs. Actually, the available hand motions when grasping an object are still rich; however, a distinctive gesture may increase the burden of learning for users who get used to touch-based

✉ Yiqi Xiao
vincentxiaoyiqi@hnu.edu.cn

Renke He
renke8@163.com

[1] University of Shanghai for Science and Technology, Jungong Road No. 516, Shanghai 200093, People's Republic of China

[2] Hunan University, South Lushan Road No. 2, Changsha 410082, People's Republic of China

gestures. On the other hand, not all multi-touch gestures can be directly borrowed to be micro-gestures [27] because the motions need to be more or less simplified. This study aims at developing a set of micro-gestures that fit different sizes of grasped objects and work for a large number of novice users. The gesture repertoire includes 12 single commands commonly used in infotainment applications.

The main purpose of this study is to evaluate the micro-gestures which are designed by considering the semantic and ergonomic issues in terms of learnability and intuitive use, to examine how much they can help drivers to lighten their workload and distractions. The paper is organized as follows. In Sect. 2 the previous work that relates to this research topic is reviewed so as to discuss the benefits from using grasp interactions and design strategies. Section 3 provides an overview of design principles and the design process. Section 4 details the evaluation experiment where a dual-task situation is simulated for users to practice the gestures in parallel to a Peripheral Detection Task (PDT). The results of a lab test are described in the next section, followed by discussion of the feasibility of gesture set and some design guidelines.

2 Related work

2.1 Grasping gestures in the driving scenarios

The driving-related activities consist of and are classified into primary tasks (maneuvering), secondary tasks (e.g., using turn signals, honking) and tertiary tasks (comfort and infotainment functions) according to the importance [10]. Buttons, gestures, voice commands and even eye movements have their own advantages in efficiently dealing with the tertiary tasks and adapting to various usage scenarios. Natural language interaction is increasingly popular, but verbally describing the in-vehicle concepts or the degree to which the window will be opened seems not as intuitive as button control is [23]. Bach et al. [6] argued that gesture-based inputs required less time for eye glances and resulted in fewer errors of lateral control, suggesting that gestures are beneficial complement for speech-based interaction. As noted by Reiner et al. [30], people tend to perform gestures in a triangle area formed by the steering wheel, rear mirror and gearshift [28] to keep their hand from completely not engaging in the driving. Therefore, it is necessary to combine input devices with the hardware used for primary or secondary tasks [8], and the gestures performed on the steering wheel have been investigated in this context.

The ergonomic characteristics of the steering wheel enable user-preferred postures of handling the steering wheel and the idea of considering the whole wheel as input field. Neßelrath et al. [25] proposed a solution for precise control using two-handed gestures. The index finger of left hand specified the context, i.e., turn signal, front window, and the right index-finger-specified functions, e.g., turning on/off. In a study of Lee et al. [17], a dashboard which allowed drivers to move a specific number of fingers to turn the switches on/off was discussed, and the multi-finger gestures they used were movements in the space. Since the thumb is relatively free when gripping on the steering wheel, a touchpad can be installed to the center of the wheel for the thumb to touch. Döring et al. [8] proved that touching this field only with the thumb leads to less driver distraction and visual demand in comparison with touch interaction with the middle console. Pfleging et al. [29] also carried out a study on the impact of performing touch gestures on such a device as soon as the name of object or function was verbally addressed. As the sensors (e.g., IR sensors [16]) can be embedded in the steering wheel in an array to detect small movements of fingers on its whole surface, users do not need to extend the finger towards a sensing area, thus the stroke gestures are easier to perform.

Most of the on-wheel gestures, in the previous studies, are micro-gestures because the freedom of manual actions is limited when users are unable to take their hands off the wheel. The available gestures in this situation include spatial movements of fingers, rubbing/tickling [36] and tiny modifications of the palm or wrist. Wolf [35] classified the grasping postures into three categories: holding a steering wheel or handle, picking a cash card, and drawing with a pen. In the ensuing research [37], she further explained that grasping gestures are superior in managing peripheral interactions. By defining some finger-motions as gestural commands, other researchers evaluated the practical effects of using certain grasping gestures. For example, Häuslschmid et al. [13] compared three pairs of finger gestures and freehand gestures. It was found that finger gestures frequently delayed steering; however, they contributed to more lane-change tasks to be successfully completed, and users reported that the perceived degree of autonomy of finger gestures was higher. The three finger motions presented in this work are performed when the steering wheel is under control of the palm and thumb. Based on capacitive sensors, Angelini et al. [4] developed a prototype named Wheelsense which can detect *tap* and *swipe* gestures. An evaluation test revealed that these micro-gestures required less time for task completion and had higher perceived usability as well as the learnability for users than touch and voice control.

The movability, flexibility and reaction ability of a part of the hand vary with different grasping postures. For designers it is important not only to find the micro-gestures that may be physically easier for users to make in any position, but also to consider how accurately they convey the meanings of the corresponding functions. According to Hoven and Mazalek [15], grasping gestures, or gestures with objects in hand, belong to the crossing research area of communicative gesture interaction and tangible gesture interaction that generally makes use of hand movements in a manipulative mode. In many cases, tangible gesture interactions are also communicative. Thus designers can define gestures by giving a manipulative movement a new meaning, and modifying the touch or three-dimensional gestures into the grasping gestures is an alternative approach. To put it simply, the grasping gestures should be ergonomically universal to generic objects and easy to understand.

2.2 Designing for the in-car micro-gestures

The design of micro-gesture vocabulary that is specifically used for the steering wheel has been explored by a number of researchers so far. Regarding the design method, user elicitation is frequently adopted as it can indicate user preferences. The main content of user elicitation is asking users to decide gesture inputs in accordance to the system outputs. To begin with, a range of functions are preset for users to assign gestures, then the experts collect gestures from the user domain. Finally some benchmarks, for example, semantic interpretation, generalization, intuitiveness and learning rate, should be used to select the appropriate gestures [26]. A typical case of using this method to design in-car gestures was conducted by Stecher et al. [31], however, all the user-elicited gestures were freehand motions, depicting 17 functions by means of sign language.

To find out the location of user-preferred grasping gestures, Angelini et al. [3] divided the surface of the steering wheel into four subareas and 44 zones before a user elicitation study. It was shown that most users were likely to put their right hand on the 1–2 o'clock position of the steering wheel. They also reported that the majority of user gestures copied the well-known multi-touch gestures to some extent. Tan et al. [32] focused on what micro-gestures users would prefer while grasping the handlebar. Well-experienced users were invited to think over the input mechanisms during cycling, and provide hand gestures for bike functions and peripheral control, consequently 10 natural micro-gestures (e.g., inter-finger contacts, bending, tilting) were illustrated. However, the researchers consider the performance mostly by comparing the gestures with mechanical inputs, instead of interpreting the gesture meanings for better intuitive use. Mahr et al. [21] tried to encourage users to co-create micro-gestures with a visible experimenter in the form of "playing roles in theater". This "theater approach" extends the Wizard-of-Oz method. As for the evaluation, researchers showed participants the graphical interface on which the input parameters and outputs displayed, so participants were informed of the effect of each gesture performance. With this setting, participants subjectively evaluated the gesture-function pairings and the perceived physical demand. Nevertheless, the study on natural use of micro-gestures in a simulated driving environment is not the focus of this work.

Comparing with user preference, the expert-level design of micro-gestures puts more emphasis on the systematic gesture set and technical feasibility. For example, the multi-finger chord vocabulary [33] devised by Wagner et al. categorized the identities that fingers touch the hand-held tablet as the way to long-term retention of chord-command mappings. In some cases, experts give priority to the effects of gesture recognition on interaction efficiency. González et al. [11] built a small touchpad called StampPad that allowed three thumb-based input techniques: clutching, dial method and displacement. It can be used by drivers to manipulate several text entry methods (i.e., list selection, linear keyboard, two-dimensional keyboard) for searching street names. Considering the high recognition rate of specific gestures with vision-based algorithms, Gupta et al. [12] dedicated to merging pairs of gestures with high miss-classification rates, to construct the robust gestures for an automotive interface. Moreover, researchers exploited the qualities of different sensors to develop gestural interaction systems. In the earlier version of Wheelsense [2], pressure sensors were embedded to detect tangible gestures including *dragging up/down* and *squeezing*. Researchers also implemented prototypes using the proximity sensor [9] and Gyro sensors [24] which can be attached to the human body so as to help disabled people to command the car, accordingly, the gestural inputs have to meet technical conditions of each recognizer.

To sum up, the micro-gestures on a steering wheel comprise of touch, 3D and manipulative gestures, aiming at reducing the motor and mental resources for the completion of concurrent tertiary tasks. The knowledge domain of both users and professional designers contributes to useful micro-gestures. However, to the best of our knowledge, there are few existing studies that focus specifically on assessing a whole set of micro-gestures in terms of intuitive use.

3 Designing micro-gestures for grasp interface

3.1 The tasks

The design process began with collecting the commonly used functions/tasks (i.e., the referents) to which gestures would be assigned. We referred to the task lists in a number of user elicitation studies, in particular the work of Chan et al. [7]. In this study we selected 12 tasks and then categorized them into three groups based on the 34 tasks and 6 categories in Chan et al. [7]. The tasks about *transforming, editing* and *selection* were subsumed into a new group called *Managing contents*, because they are used to process, move or transform the digital contents. We replaced *menu* and

Table 1 The 12 tasks used for the design stage

Group	Referents	Description	Number
(a) Managing contents	Enter option or accept	Enter a program; answer the phone	1
	Cancel or reject	Refuse to answer; shut down	2
	Copy and paste	Transfer file or information	3
	Send	Send or submit	4
	Save	Save files; favorites button	5
(b) Browsing pages	Go to the main menu	Go back to the main menu	6
	Zoom-in/out	Zoom-in or out of contents of a page	7
	Scroll	Scroll through a web page	8
	Previous/next	Go to the previous/next page	9
(c) Adjusting attributes	Fast-forward	Fast forward music or video to a good part	10
	Volume up/down	Increase or decrease volume	11
	Stop or mute	Stop playing	12

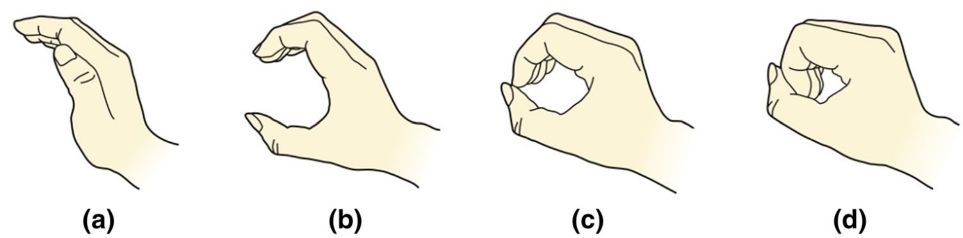

Fig. 1 Different postures of grasping an object with different levels of diameter

browsing with *Browsing pages* as it contains the tasks that aim to switch pages and receive the information that display on a screen. The *simulation* group [7] was renamed *Adjusting attributes* for directly describing the tasks of changing current state of outputs.

The *Managing contents* group contains five tasks, i.e., enter option/accept, cancel/reject, copy and paste, send and save. The four tasks in *Browsing pages* are (1) go to the main menu, (2) zoom-in/out, (3) scroll and (4) previous/next. Fast-forward, volume up/down and stop make up the *Adjusting attributes* group. We numbered these tasks in sequence, see Table 1. It is worth noting that five of the 12 tasks are dichotomous pairs, e.g., previous/next, and stop/replay is an example of state toggle.

3.2 Design process

Interaction designers played a leading role in the process of designing gestures for these tasks. As mentioned above, expert-defined and user-defined gestures are different in the standing point of design. Expert-defined gestures usually outperform the user-defined ones in respect of systematic design, though the prior knowledge or habits of users is what the user-elicited strategy is more concerned with. Gesture elicitation is a widely applied design method; however, we consider that the micro-gestures are not as familiar as screen-based gestures and buttons to normal users. Finally we opted for the designer-oriented strategy as it might not be easy for users to put forward a dozen of grasping gestures that worked well in a driving context.

Defining the size of grasped objects is the starting point of creating the universal gesture set. In the first place, we specified four grasping postures in the requirements document according to different levels of the diameter of grasped objects (Fig. 1). In addition, we formulated the design principles for professionals to define on-wheel micro-gestures based on the study of Wolf [37]. The following is a summary of these principles.

1. *Grasp stiffness*, as a result of exerting force on a grasped object for a long period of time, derives from the need of keeping a tight grip on that object. Drivers may not tend to put their hands on fixed positions of the wheel to avoid grasp stiffness, so the micro-gestures should adapt to the variations of grasping posture.
2. *Form-dependent limitation* refers to the fact that the movability of fingers depends on the thickness of grasped objects as shown in Fig. 1. Due to this limitation, the grasping gestures that bring together the thumb and the other four fingers are unable to perform in specific cases, i.e., the object is rather thick. Therefore such gestures are not recommended.
3. *Digit-dependent limitation* is the result of different flexibility of individual fingers. Anatomic studies suggest that fingers differ from each other in the separate mov-

ability. The middle and the ring finger are harder to flex independently than the thumb and the index finger are.
4. *Cognitive ergonomics* means gestures are easy to learn and remember with low complexity. Another important reason why a gesture is deemed easy-to-learn is that the meaning of the motion is self-evident. Furthermore, designers are recommended to carefully use the orientational metaphors to interpret grasping gestures, provided that the grasping posture limits fingers to horizontally or vertically move.
5. *Error-proofing* should also be seriously considered. For example, the unintentional single-finger tapping is inappropriate to be defined as the gesture of controlling the state of a device (e.g., open, close, music control, etc.).

In the design process, interaction designers were also provided with a checklist which incorporated 140 potential grasping gestures in total. We brainstormed this list by classifying the micro-gestures executed when grasping a physical object into three categories (i.e., movements in mid-air, movements on the surface of objects, and contact gestures) in advance. This document enabled designers to directly select gestures as the inputs, or they could develop different gestures as needed.

Five designers took part in our study as members of an expert panel. All the designers have experience in car driving (the driving experience they reported ranged from approximately 4500 to 30,000 km). The design process proceeded in two phases called elicitation and selection. In the elicitation phase, every designer was suggested to propose usable, learnable and intuitive gestures as many as possible for each of the tasks, on the basis of the aforementioned design principles. Designers have to fill in a form with details of the manual actions and the threshold of each micro-gesture they devised. Additionally, they also have to give a detailed account of the "prototypes" of a gesture, for example, routine activities, legacy-inspired gestures [32], cultural meanings [19] or any concepts that serve as the bases of an expert-defined gesture.

In many cases, a gesture may have more than one prototype that helps users to easily relate the motion to a rather familiar activity. The elicitation phase ended with the collection of 186 micro-gestures in total. After grouping the similar gestures for a task together and defining them as one proposal, 79 different gestures remained. For the dichotomous tasks designers preferred opposing gestures, and they believed that identical gestures were more suitable for representing toggles as have been claimed by [7].

In the first stage of the selection phase, designers ranked the frequency of gestures for a task, then found the most frequently suggested "prototypes" of these gestures. A ten-point Likert scale was adopted for designers to measure how much a gesture could be conceptually associated with a task. The gesture with the highest average score was selected among all the proposals for a task and considered as the best option for task-to-gesture mapping. Through this process, we finalized the 12 one-hand micro-gestures used on grasped objects (especially steering wheel in this paper) with the help of designers.

3.3 The gesture set

Figure 2 illustrates how these micro-gestures are performed. They are described as follows: (1) double tap: tapping the steering wheel for two times; (2) draw a "x" mark: the index finger draws two unparallel lines in mid-air; (3) tow-finger pinch: two fingertips must contact the steering wheel; (4) tap and flick: the index finger taps the steering wheel and then lift up; (5) pat on the wheel: the palm shortly disengage from the wheel, then pat it; (6) draw a circle sidewise: the movement trajectory of the thumb should be a circular arc over 270 degrees; (7) two-finger zoom-in/out: the degree of zooming in/out depends on how many times the thumb and the index finger stretch simultaneously; (8) scroll: each time the index finger vertically swipes on the steering wheel represents a page scrolls up/down; (9) thumb swipes: the thumb moves along a horizontal axis in mid-air; (10) "walking" gesture: alternate tapping on the steering wheel with index finger and middle finger; (11) swipe index finger with middle finger: the degree of volume increase/decrease depends on the distance of finger movements; (12) snip: the index finger and the middle finger both erect. These descriptions also detail the thresholds of some gestures.

4 Evaluation methods

4.1 Participants

Eighteen drivers (ten men and eight women) volunteered to participate in the evaluation study about the learnability and quick response of the micro-gestures (LQRMG). Their age range was in the 19–32 years old range ($M = 22.444$, $SD = 3.666$). None of them had been licensed to drive for over 1 year, and nor was experienced in driving the car which was equipped with a gesture control system. All the participants were right-handers. An examination confirmed that the visual acuity of any participant was not less than 0.8 (the Chinese standard measured by E chart, the USC equivalent of which is 20/25). Before the final participants were selected, the volunteers were involved in a pre-test where they were shown video clips of a racing video game in the first-person perspective. The volunteers who had played similar games were excluded in case they were easier to adapt to the simulated environment. These standards were made to ensure that the participants were physically qualified to

Fig. 2 The 12 gestures designed for the functions listed in Table 1, the numbers of gestures correspond to the task numbers as shown in Table 1

our test, and their prior knowledge would not have impacts on the test results.

4.2 Materials

The evaluation test was conducted in the lab. In order for the experiment room to imitate an interior space of the car, a laptop on a table and a chair were each placed at a stationary location to guarantee the distance between the screen and the chest of participants is about 60 ± 5 cm (Fig. 3a). A steering wheel was positioned 25 cm away from the screen. The computer used in the test was an original ASUS K550J with 1920×1080 resolution display. Other devices included a set of pedals, an Eyelink II eye tracker, and three cameras (two Sony 5N cameras and a Ordro HDV-V7). Two of the cameras mounted on tripods were used to capture the gesture performances from two different angles as illustrated in Fig. 3a, and the other one was placed under the table to monitor foot motion. The post-test analysis was mainly based on the videotaped data.

During the test, participants watched a game video as the substitute for a driving simulator. This video was edited from video recordings of playing the City Car Driving game, see Fig. 3b.

4.3 Procedure

The evaluation test generally continued for about 1 h and 10 min, including three phases: learning, training, and testing.

Before the learning session, first, the experimenter explained the purpose and steps of the test for the participants, stressing that the final results did not reflect their personal qualities in an attempt to lessen their feelings of stress and desire to perfect performance. For the repeated gestures, i.e., two-finger zoom-in/out, scroll and "walking" gesture, it was required for participants to perform such gestures only one time in each trial. A 2-min and 24 s teaching video was set for participants to learn the gestures. During the 12-s video time of showing each gesture, participants could see

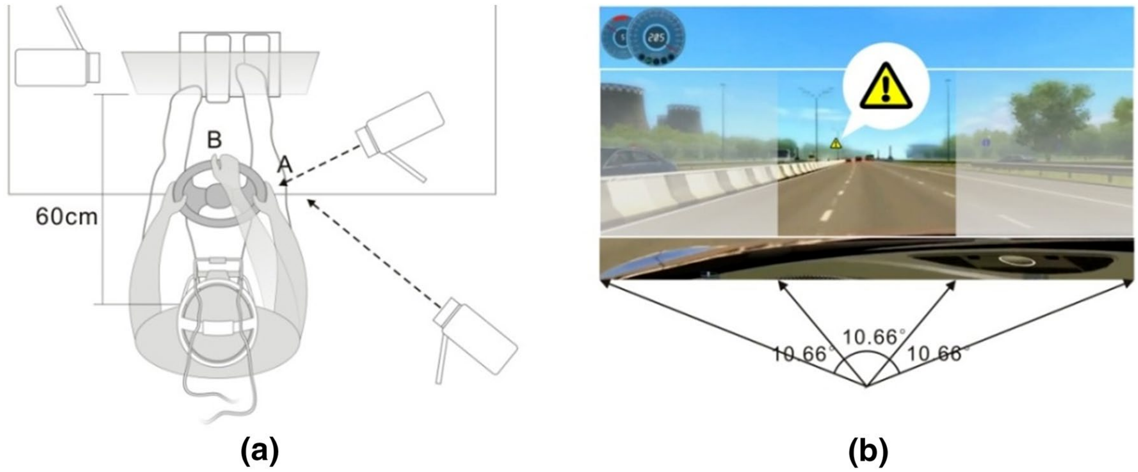

Fig. 3 Experimental setup: **a** overview of test environment and the grasping postures of user group A and B. **b** Screenshot of video game, the three areas of AOI are indicated on the screenshot by white squares

the standard performance, followed by a slow-motion video playback. The participants were allowed to watch the teaching video for one time, thereafter they began to learn the gestures and their corresponding functions by watching the video repeatedly and checking a typewritten command list. The participants informed the experimenter when they felt that they had remembered all the gestures.

In the training session, participants were asked to correctly answer the function a gesture represented, and then to accurately perform the gesture for a function the experimenter instructed them to achieve. We called the first part of training "tell me the function", and the second part "show me the gesture". The 12 commands were asked in random order for each participant. If a participant failed to give the right answer of a certain trial, he or she would review the teaching video until they were confident of proficiently using the 12 commands. Before testing participants had the chance to try out all gestures again. The main purpose of gesture training was to test the effect of learning; the two parts of training session would assist participants in fully mastering the gestures.

The primary task for the testing session is a modified version of Peripheral Detection Task [14] which is more attention-demanding than the original version is. In this study we did not adopt the Lane Change Task (LCT) as it was considered of higher learning effect for novices. Besides, the PDT is a proper method to evaluate the effects of driving situations on the workload of a primary task.

The test started with playing the game video. Participants had to identify the warning signs—each of which was an exclamation mark on a standard triangular sign with yellow background—that appeared in the field of vision, see Fig. 3b. The size of a warning sign was 3×3 mm, 17×17 pixels. In total 150 signs (i.e., signals, or stimuli) were displayed on the area of interest (AOI) in random time during the 12-min video. The inter-stimulus interval of two stimuli is no less than 1 s, and a stimulus duration is 0.5 s. Participants needed to immediately press the foot pedal as soon as they discerned a stimulus. The 150 stimuli were equally allocated to the left, right, and central area of AOI. The two lines between the sides of each area and the eye of participants formed a horizontal angle of approximately 10.66° (Fig. 3b). Ten volunteers tested this experimental setup prior to the formal testing to decide on the size and duration of a stimulus among four conditions (17×17 pixels, 500 ms; 25×25 pixels, 500 ms; 17×17 pixels, 750 ms; 25×25 pixels, 750 ms). The results of NASA-TLX questionnaire showed the workload of test in the condition of (17×17 pixels, 500 ms) was relatively moderate.

With such a setting, participants had to keep the gaze sweeping rapidly around the screen to identify as many signs as they could. They were reminded that not being too sensitive to wrongly react to the stimuli which looked like the warning sign. The experimenter calibrated the eye tracker, then asked participants to get familiar with the primary task by some try-outs. Participants watched the video twice, once for executing PDT without intervention of secondary tasks as the baseline data and once for the dual-task test. The number and location of stimuli in the videos for the two tests are the same, but the points of time at which they appeared are different. This design was set to prevent the results from being influenced by the priming effect, as we considered that participants were likely to be primed by remembering the fixed time point of a stimulus.

In the dual-task test, participants performed the gesture when they heard the name of a task, e.g., dubbed "Next Page". We used voice commands to be the gesture triggers to avoid visual overload [13]. The 12 tasks were performed

Table 2 Evaluation measures and indicators

Evaluation measures	Indicator
Learning time	The sum of elapsed time a learner takes to watch the video clip of each gesture
Effect of learning	Number of answers that wrongly match a gesture to a certain function
Fraction of missed signals and wrong reactions	% of signals to which a participant did not respond/total number of signals and number of steps on the brake pedal which were not the reactions to an observed or a true stimulus
Reaction time to stimuli	Time interval between the onset of the stimulus and the time point when the brake pedal was pressed down
Number of errors	Number of gesture performances do not visually meet the thresholds
Reaction time to voice prompts	Time interval between the time point when a voice prompt was emitted and the time point when participant starts to move finger(s)
Fixation count per second	Fixation count/time span between the time when a voice prompt was emitted and the time a gesture performance was ended

one by one. In the next round, tasks were repeated in a different order. For each participant 12 tasks × 4 repetitions = 48 trials were required. The order of tasks was counter-balanced between repetitions and participants; and the time interval between two trials is 4 s at least. Participants only had one chance to perform the right gesture in every trial. To further minimize learning effects, the driving speed that participants felt in video increased over time.

We hypothesized that the hand posture is a variable in intuitive use of grasping gestures. In testing session the 18 participants were equally divided into two groups, so that a comparison study could be conducted between two groups to know the effects of grasping postures on gesture performance. Participants in Group A held the right side of the steering wheel with their right hand, while the holding position of Group B members was the top side, see Fig. 3a. If the gesture set proved to enable different ways of grasping, the input field was not confined to a specific part of the steering wheel.

4.4 Measurements

The study procedure was organized to evaluate if the gesture set was intuitive enough to be quickly learned and then to be well performed in the extreme condition that users kept devoting much visual attention to a primary task. The results were subsequently analyzed by the following measures (Table 2) in three aspects.

4.4.1 Learnability

This was measured by the learning time (LT) and the effect, i.e., number of correct answers to the questions. As described in Table 2, the learning time for each gesture was the sum of every time a participant took to watch its teaching video. We used confusion matrix to show the effect that users learned the micro-gestures.

4.4.2 Intuitiveness

Number of errors and reaction time (RT) to voice prompts were used to measure how intuitively participants could use the gestures. The higher rate of task completion and less reaction time implied the gesture performance demanded very limited amount of mental resources. The reaction time is the time interval between two trials in case that the participant did not make a gesture as a response to the task.

4.4.3 Distraction

The measurement of distraction was conducted by retrieving the data about (1) missed signals (MS) and wrong reactions,[1] (2) reaction time to stimuli, and (3) fixation count per second. The first two measures aimed to quantify the effects that allocating users' attention to the gesturing for an instant would have on the completion level of primary task. As the participants had to move the visual focus constantly to identify the signals, we computed the recorded fixation count in AOI from the onset of a voice prompt to the end of a gesture enactment, to measure the degree that gesture performance might affect the efficiency of searching targets.

We used a post-experimental questionnaire for evaluating user satisfaction. The learnability, intuitive use and effortlessness of each micro-gesture were rated by participants according to their personal standards, based on a symmetric 7-point Likert scale. After that participants explained the grading and made some comments on the gesture design.

[1] If no reaction was detected in the time interval between two stimuli, the first stimulus was a missed signal (false negative). The wrong reactions (false positive) occurred on two situations. First, more than one reactions was detected in the time interval between two stimuli, in this case the reactions that were made later than the first one were regarded as mistakes. Second, if only one reaction was detected, and the fixation did not overlap or get close to the signal on the screen, experimenters judged it to be an unintentional response.

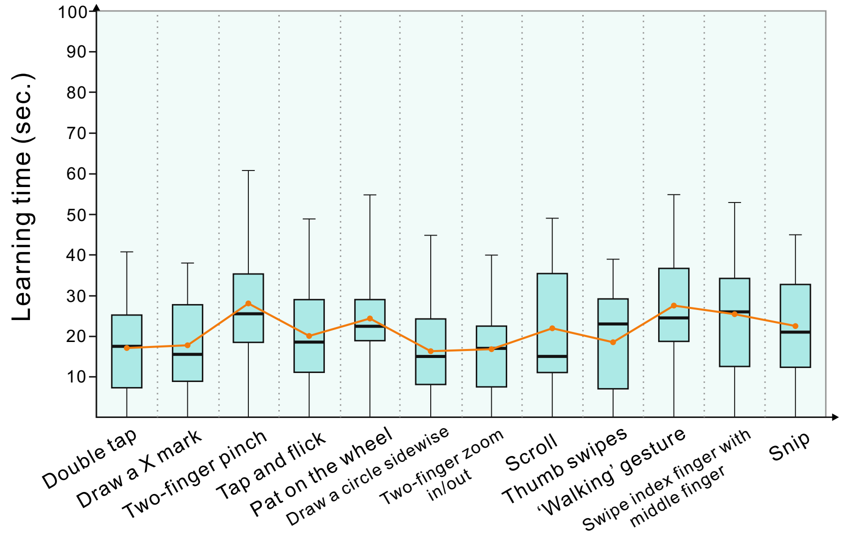

Fig. 4 Participants' learning time for the 12 micro-gestures, the orange points indicate mean values

Except for the fixation count, the summative data of the test were extracted from the video streams which were synchronized with the time points of stimuli and voice prompts, through frame-by-frame analysis. Without system feedback, participants did not know how many mistakes they had made, so it was difficult for them to correct wrong gestures or consolidate memory with repeated practices during the test. For identifying the correct performances, we conducted the postural analysis to compare the videotaped gestures with those in Fig. 2. All the quantitative and qualitative data were processed by SPSS 19.0.

5 Results

The results of the evaluation test were reported in this order: comparisons of gestures, completion of the PDT (primary task), comparisons between two groups and user satisfaction.

5.1 Gesture performance

In this section, we compare the micro-gestures in terms of the learning time, effect of learning, number of errors, reaction time to voice prompts and fixation count per second, thus the differences in the learnability and intuitive use of gestures were investigated.

Figure 4 illustrates the learning time participants spent in remembering each gesture. Based on the post-hoc analysis of a Friedman test, the gestures can be divided into four groups according to the mean rank differences of LT. In the eight gestures in Group 1, draw a circle sidewise, two-finger zoom-in/out, double tap and draw a "x" mark were not included in the other three groups which were significantly different from each other in LT ($p<0.05$). The difference of all the gestures in LT was statistically significant [$\chi^2_{(N=18)}=69.661$, $df=11$, $p<0.001$]. On average, "Draw a Circle Sidewise" took the shortest learning time while "Two-finger Pinch" required the longest time to learn. Some participants felt the "Two-finger Pinch" gesture was the most difficult to learn, although they could understand that designers defined this motion as the command of Copy and Paste with the intent of comparing the index finger and the middle finger to C key and V key respectively. Generally, participants thought "Pat on the Wheel" and "Walking" gesture were also less similar to the more familiar gestures besides "Two-finger Pinch".

Regarding the total learning time, the differences of participants were very evident ($M=261.778$, $\mathrm{SD}=145.528$). Surprisingly, the LT of one participant (P15) was zero because he said he had learned all gestures before the learning session (watching teaching video a second time). By contrast, the differences of participants in the training session was not so great. There were 11 participants correctly matched all the functions with the gestures, and eight out of 18 participants accurately performed all the gestures. Among 18 participants six responded correctly to all the $12\times2=24$ questions. The confusion matrix is shown in Fig. 5.

It is noteworthy that participants most frequently failed to answer what function the Scroll gesture represents. This result can be interpreted by the less obvious feature of Scroll. As the gesture suggests an "up and down" motion, it might easily remind participants of other functions such as turning a page or adjusting volume. A mistake the participants made when demonstrating to experimenter the "Thumb Swipes" gesture is representative, that is, they ended up swiping from right to left to complete the "Next" task. Five out of 18 participants reported that they were more used to the interactive mode of mobile phones which uses the "right-to-left" swipe gesture to denote that the current page is dragged to the left side. This result warns designers to specifically consider the

	1	2	3	4	5	6	7	8	9	10	11	12
1	16	0	0	0	0	0	0	0	0	0	0	0
2	0	18	0	0	0	0	0	0	0	0	0	0
3	0	0	16	0	0	0	0	0	0	0	0	0
4	0	0	0	18	1	0	0	0	0	0	0	0
5	0	0	1	0	16	0	0	0	0	0	0	0
6	0	0	0	0	0	18	0	0	0	0	0	0
7	0	0	0	0	0	0	18	0	0	0	0	0
8	0	0	1	0	0	0	0	15	0	0	0	0
9	0	0	0	0	0	0	0	1	17	0	0	0
10	0	0	0	0	0	0	0	1	0	17	0	0
11	0	0	0	0	0	0	0	1	0	0	18	0
12	1	0	0	0	0	0	0	0	0	0	0	18

(a)

	1	2	3	4	5	6	7	8	9	10	11	12
1	17	0	0	0	2	0	0	0	0	0	0	0
2	0	18	0	0	0	0	0	0	0	0	0	0
3	0	0	16	0	0	0	0	0	0	0	0	0
4	0	0	0	18	0	0	0	0	0	0	0	0
5	1	0	0	0	15	0	0	0	0	0	0	0
6	0	0	0	0	0	18	0	0	0	0	0	0
7	0	0	0	0	0	0	17	0	0	1	0	0
8	0	0	0	0	0	0	0	17	0	0	0	0
9	0	0	0	0	0	0	0	0	13	0	0	0
10	0	0	0	0	0	0	1	1	0	16	0	0
11	0	0	0	0	0	0	0	0	0	0	18	0
12	0	0	0	0	0	0	0	0	0	0	0	18

(b)

Fig. 5 Confusion matrixes: **a** "Tell me the function" part, row: gestures, column: functions. **b** "Show me the gesture" part, row: functions, column: gestures

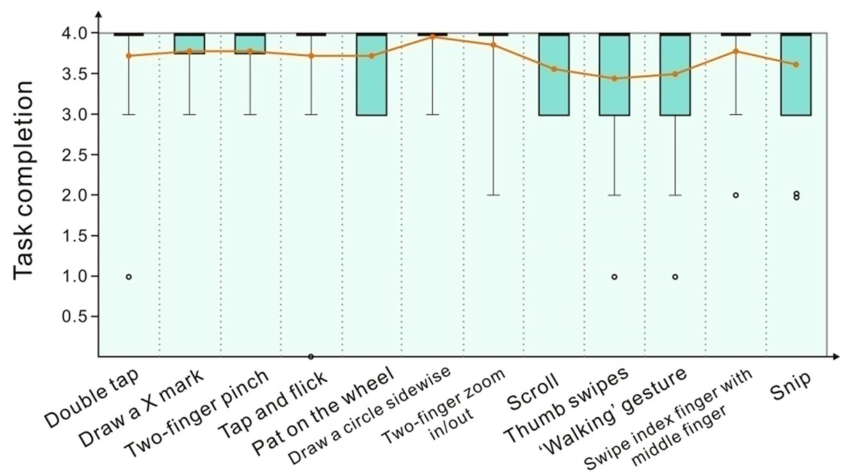

Fig. 6 Number of task completion in four repetitions of 12 trials

different mental models that will lead to opposite direction of a swipe gesture.

With respect to the number of errors, a Friedman test revealed that the effect of gestures was not statistically significant [$\chi^2_{(N=18)} = 15.434$, $df = 11$, $p = 0.163$]. As evident in Fig. 6, the error rate of "Thumb Swipes" and "Walking" gesture was relatively higher, whereas only one error was detected for performing the "Draw a Circle Sidewise" gesture. As is shown in Fig. 4, participants understood the meaning of this gesture at the cost of shortest learning time. By comparison, they forgot the "Scroll", "Thumb Swipe" and "Walking" gestures in the test several times. As a consequence, it was a bit difficult for participants to accurately perform these gestures to respond to the suddenly appeared voice prompts when their main attention was paid to reacting to the stimuli.

On comparing the reaction time, we performed a repeated measures ANOVA. The results indicated an interaction effect between gestures and repetitions on RT [Huynh-Feldt correction: $F_{(33,612)} = 2.060$, $p = 0.001$], and the significant differences of RT between gestures [$F_{(11,204)} = 2.026$, $p < 0.05$]. A pairwise comparison of RT was carried out using the Turkey HSD method. The result showed that the differences of "Walking" gesture (M_{round1}=1.869, M_{round2}=1.770, M_{round3}=1.471, M_{round4}=1.463) and "Draw a X Mark" (M_{round1}=1.219, M_{round2}=1.360, M_{round3}=1.115, M_{round4}=1.070, SE = 0.174, $p < 0.05$), "Tap and Flick" (M_{round1}=1.273, M_{round2}=1.134, M_{round3}=1.140, M_{round4}=1.186, SE = 0.168, $p < 0.05$), and as well as the "Draw a Circle Sidewise" gesture (M_{round1}=1.228, M_{round2} =1.228, M_{round3}=1.149, M_{round4}=1.072, SE = 0.152, $p < 0.05$) in RT were statistically significant (see Fig. 7). The gestures were not very different in the reaction time participants took to perform on the whole with the exception of the "Walking" gesture. Although a slight delayed response occurred more than

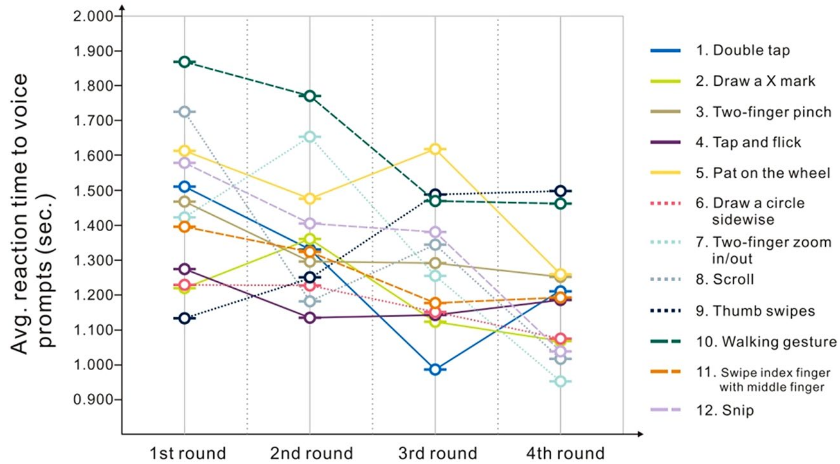

Fig. 7 Reaction time of micro-gestures to the voice prompts

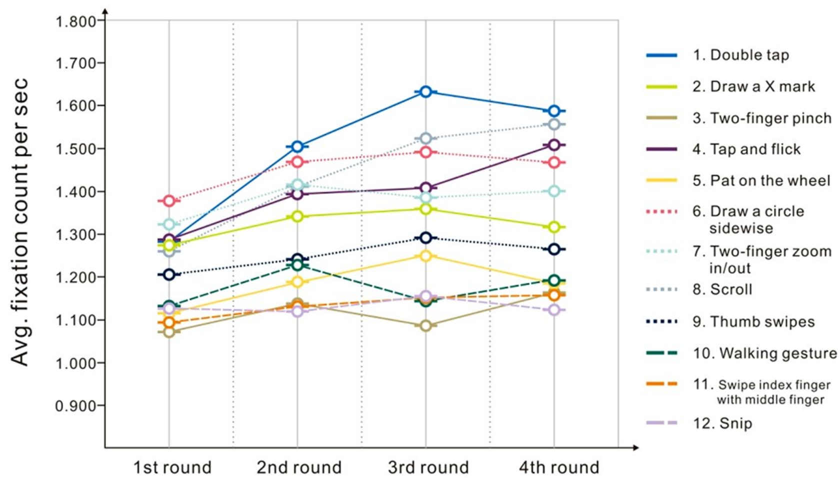

Fig. 8 Fixation count per second of participants in the duration of gestural interaction

once when performing each of the gestures, the reaction time of the "Walking" gesture was obviously much lower, which directly resulted in a significant data difference.

A repeated measures ANOVA revealed a statistically significant difference between gestures in terms of fixation count [$F_{(11,204)} = 9.344$, $p < 0.001$], however the results did not show an interaction effect between gestures and repetitions on fixation count (Mauchly's sphericity test: $\chi^2 = 6.708$, $p = 0.243$). In contrast to the RT which changed over time, fixation count per second was independent of different repetitions, perhaps due to the consistent strategy participants used to observe the stimuli and their inherent ability of visual search. As illustrated in Fig. 8, the mean values of fixation count per second were lower when performing the "Walking", "Swipe Index Finger with Middle Finger", "Snip" and "Two-finger Pinch" gestures. A likely explanation for this result is that all these gestures require the moving of the middle finger which is not very independently movable. The using of the middle finger might have diverted much of the users' attention, resulting in longer fixation duration in identifying the warning signals.

5.2 The primary task

In this section we report the effect of using micro-gestures on the PDT task performance. Regarding the percentage of missed signals, a Wilcoxon signed-rank test revealed statistically significant differences between the baseline data and dual-task performance [$Z(18) = -3.313$, $p = 0.001$]. Participants were more frequent to ignore the stimuli flashed on the screen in dual-task test, as seen Fig. 9a. Some of the participants reported that the shorter the time interval between a voice prompt and a stimulus was, the easier the stimulus was overlooked, especially when the voice came before stimulus. In this situation, they found that the attention they kept on the PDT was not enough to identify signals as accurately as usual. This result gives an overall indication of how much effect the gesture set has on the results of PDT.

The study delved further into the differences of missed signals in the three areas of AOI. A pairwise t test showed that the MS in central area was significantly less than both the left [$t(18) = -10.412$, $p < 0.001$] and right area [$t(18) = -11.020$, $p < 0.001$], and the differences between the left

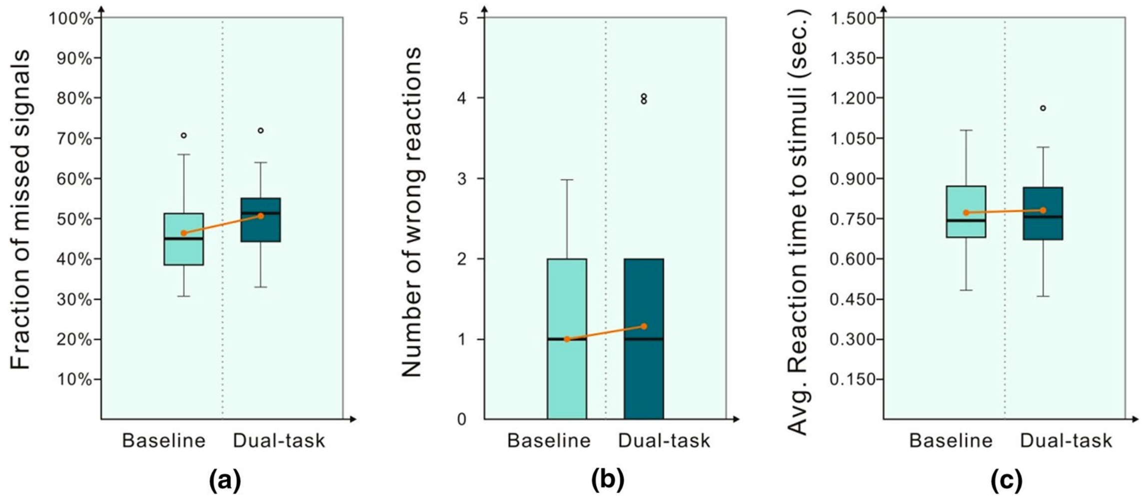

Fig. 9 Results of performing the PDT (primary task): **a** missed signals, shown as percentages. **b** Wrong reactions. **c** Reaction time to stimuli in two tests, shown as mean values

and right area were not statistically significant [$t(18) = 0.203$, $p = 0.842$]. This result demonstrated that the scope of users' visual search was narrowed when they were gesturing, and so they were more difficult to notice the stimuli appeared on either side area. This claim can be corroborated by comparing the dual-task results with baseline data. The differences between baseline data and dual-task in both the MS in the central area [$t(18) = -0.595$, $p = 0.560$] and the left area [$t(18) = -1.922$, $p = 0.072$] were not statistically significant, however, in the MS in the right area [$t(18) = -2.465$, $p < 0.05$], was statistically significant. In other words, as the secondary task, performing micro-gestures limited the movement of participants' visual focus in a certain range.

In regard to the wrong reactions, baseline data were slightly higher than the dual-task performance [$Z(18) = -0.632$, $p = 0.527$]. Overall, the frequency of wrong reactions in either of the two tests was low (Fig. 9b). In the video recordings we observed that participants occasionally reacted to non-existent stimuli, or tended to lightly step on the pedal as if they hesitated to confirm the stimuli at the first half part of two tests. When participants were debriefed, they said that they wrongly made a few body responses, and they thought the wrong reactions were mainly caused by the nervousness which was eased as the test proceeded. We can conclude that the dual-task did not cause more human errors in PDT, even if much attention was paid to performing the micro-gestures when needed.

A pairwise t-test showed that there were no statistically significant differences [$t(18) = -0.698$, $p = 0.495$] between the average reaction time to stimuli in baseline data and dual-task performance (Fig. 9c), suggesting that the gestures played a minor role in affecting the RT. However, the results of two tests in terms of the standard deviation of RT to stimuli were significantly different [$t(18) = -2.220$, $p < 0.05$]. Descriptive data showed the average RT to stimuli in dual-task test ranged from 1082 to 483 ms. This indicated that the participants' response ability varied greatly. In the dual-task test more quick responses were recorded in comparison with baseline data. It was probably because participants had been skilled in performing the PDT with repeated practices. Meanwhile, the larger number of delayed responses in the dual-task test could be related to the gestures which hindered participants from fully engaging in the PDT.

5.3 Comparative results

To test the usability of micro-gestures when users grasped the wheel in different postures, participants were split into two groups as mentioned in Sect. 4.3. A comparison of the two groups indicated that the grasping postures had insignificant effects on gesture performance, and it almost did not impact on the results of PDT.

Figure 10 showed a box-plot of the number of errors in dual-task test. A Mann–Whitney U test revealed that there were no statistically significant differences between the two groups [$U(18) = 26.5$, $p = 0.222$]. We then employed repeated measures ANOVA again to compare the RT to voice prompts and the fixation count per second of the two groups for performing every gestures. Regarding the RT to voice prompts, statistical analysis showed the differences between Group A and B were statistically significant when performing "Pat on the Wheel" [$F_{(1,16)} = 5.674$, $p < 0.05$] and "Scroll" [$F_{(1,16)} = 4.817$, $p < 0.05$], as compared with the other ten gestures.

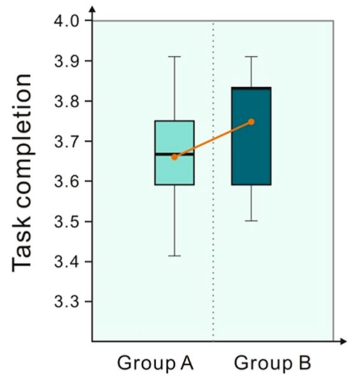

Fig. 10 Number of task completion of two user groups

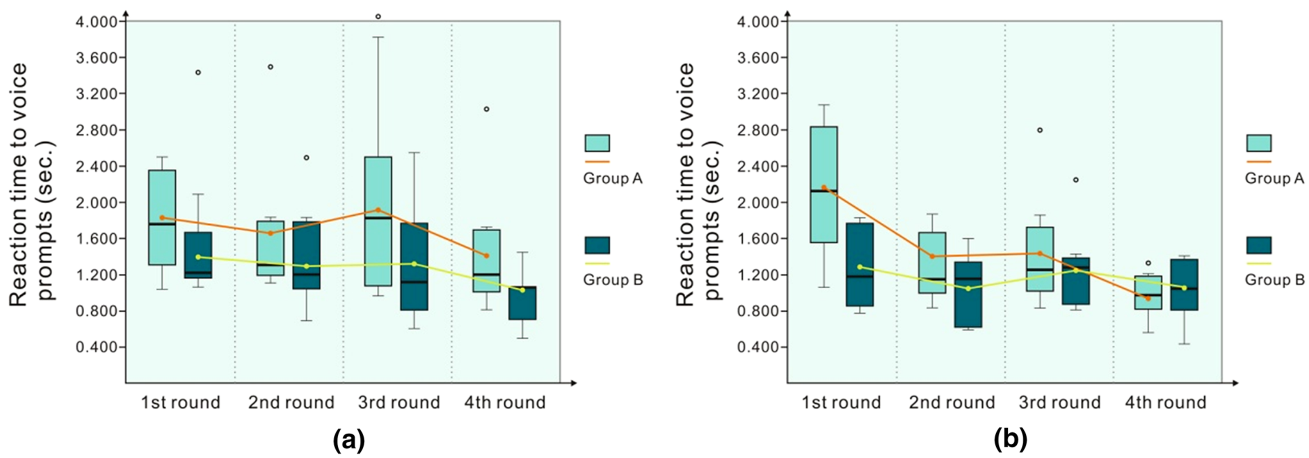

Fig. 11 Reaction time of two gestures. **a** Pat on the wheel. **b** Scroll

In Fig. 11b we can see the RT of Scroll gesture of Group A was far higher than Group B only in the first round. On the contrary, in the fourth round the average RT of Group B was slightly higher. This is likely to be the result of random grouping based on a small sample size. In fact, several delayed responses of Group A members occurred in the first round by chance.

However, in each round, Group A members needed longer RT to the voice prompt which instructed participants to perform the "Pat on the Wheel" gesture [Round 1: $U(18)=23$, $Z=-1.545$, $p=0.136$; Round 2: $U(18)=27$, $Z=-1.192$, $p=0.258$; Round 3: $U(18)=22$, $Z=-1.634$, $p=0.113$; Round 4: $U(18)=22$, $Z=-1.634$, $p=0.113$], see Fig. 11a. As a whole, the two groups were significantly different in this regard, even though none of the between-group differences in each round was statistically significant. For these results, participants in Group B felt that they did not need much preparation to perform the "Pat on the Wheel" gesture because they would not firmly grip on the wheel. Rather, they got used to casually put the hand on the top side of steering wheel just for a moment, in order that they could take a sharp turn next. On this account, the participants who gripped on the wheel with five fingers spent more time before starting that hand motion.

The data showed that in the duration of performing any of the gestures, the differences between two groups in terms of the fixation count were not statistically significant. Participants agreed that the grasping posture of Group B members was more likely to cause muscle stiffness and driving fatigue. Since the test duration is about 12 min, it seems the hand posture has little influence on the frequency of eye movements in such a short period of time.

For either of the MS [$U(18)=36$, $p=0.730$] and wrong reactions [$U(18)=26.5$, $p=0.222$] in PDT, a Mann–Whitney U test revealed that there were no statistically significant differences between the two user groups. The two groups were similar in the RT to stimuli [$Z(150)=-1.376$, $p=0.169$]. This means that the performance of PDT does not depend on the posture in which drivers hold the steering wheel.

5.4 User satisfaction

The user ratings (see Fig. 12) indicated an overall satisfaction on the learnability of the gestures set. However, the ratings of learnability for each gesture significantly varied [Friedman test: $\chi^2_{(N=18)}=47.227$, $df=11$, $p<0.001$]. According to the mean rank differences the 12 gestures can be divided into three groups, among which the least easy-to-learn group ($p=0.55$) consists of five gestures: "Thumb Swipes" ($M=5.833$, SD $=1.150$), "Pat on the Wheel" ($M=5.889$, SD $=1.079$), "Swipe Index Finger with Middle Finger" ($M=5.222$, SD $=1.114$), "Two-finger Pinch" ($M=4.944$, SD $=1.162$), and "Walking" ($M=4.778$, SD $=1.555$). The participants found it took longer time to remember these gestures and to map them to the corresponding functions.

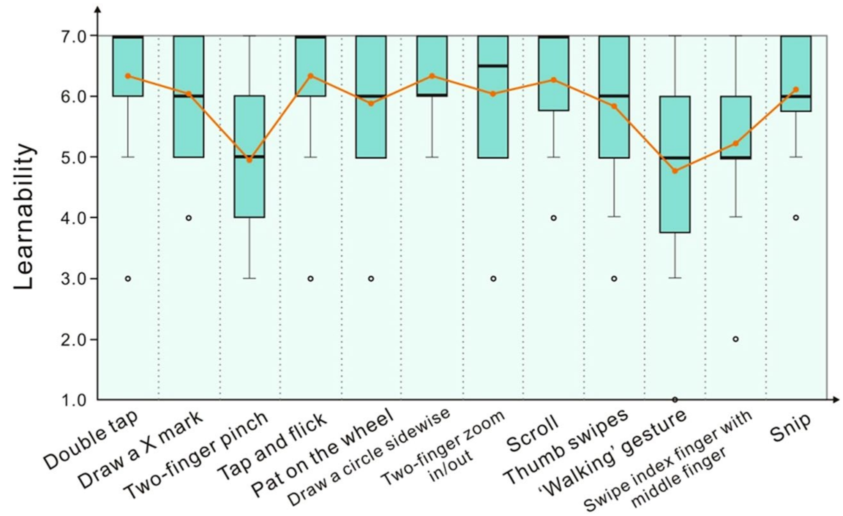

Fig. 12 Participants' evaluations of the learnability of micro-gestures

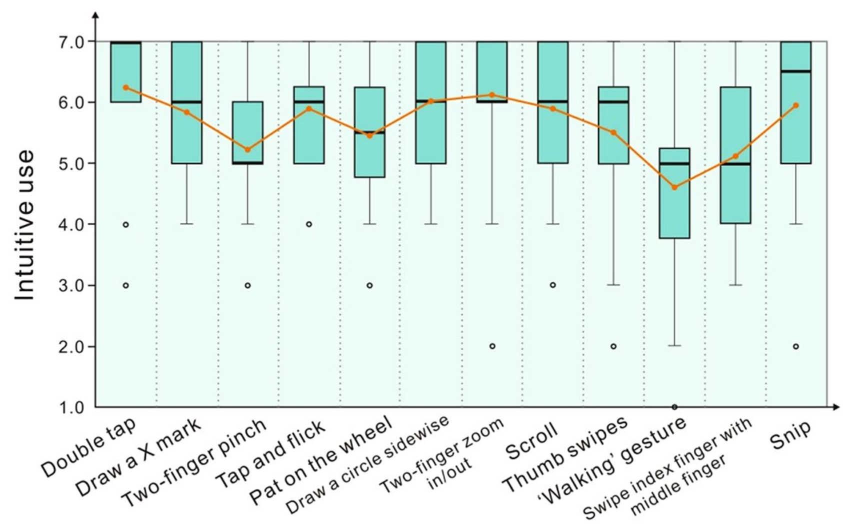

Fig. 13 Participants' evaluations of the intuitive use of micro-gestures

The learning time of gestures may correlate with the prior experiences and the memory strategies of users. After the end of the questionnaire survey, the three male participants (P8, P13, P15) who spent no more than 2 min in learning session were debriefed on the reason of quick learning. They said all gestures were easy to remember by relating them to the icons and the interactions with PC or mobile phone. Most of the participants considered that the metaphors "Two-finger Pinch" and "Walking" gesture used were intelligible, though not self-evident enough, consequently they needed much time to construct the mappings between gestures and meanings.

A Friedman test revealed that the differences of user ratings regarding the intuitive use between all gestures were statistically significant [$\chi^2_{(N=18)}=41.100$, $df=11$, $p<0.001$], see Fig. 13. The gesture with the highest score is "Double Tap" ($M=6.278$, $\mathrm{SD}=1.127$), followed by "Two-finger Zoom in/out" ($M=6.111$, $\mathrm{SD}=1.278$), "Draw a Circle Sidewise" ($M=6.000$, $\mathrm{SD}=1.029$) and "Snip" ($M=5.944$, $\mathrm{SD}=1.434$). The mean score of "Walking" gesture is the lowest ($M=4.611$), but it has a high standard deviation of user rating ($\mathrm{SD}=1.650$) as a few participants regarded it as a very intuitive input.

When considering how much the gestures can be intuitively used, participants appeared to be more likely to give an easy-to-learn gesture higher scores. They felt that the difficult-to-learn gestures can be retained by short-time memory in test duration indeed, but much task-switching cost is required to perform them in a dual task. For some participants, it was easier to evoke the memories of a gesture that makes use of iconic metaphors. For example, "Snip" uses the "two vertical bars" symbol to represent the "Stop" or "Mute" function. From the test results, however, we did not see rather short RTs to stimuli before performing this gesture (see Fig. 7).

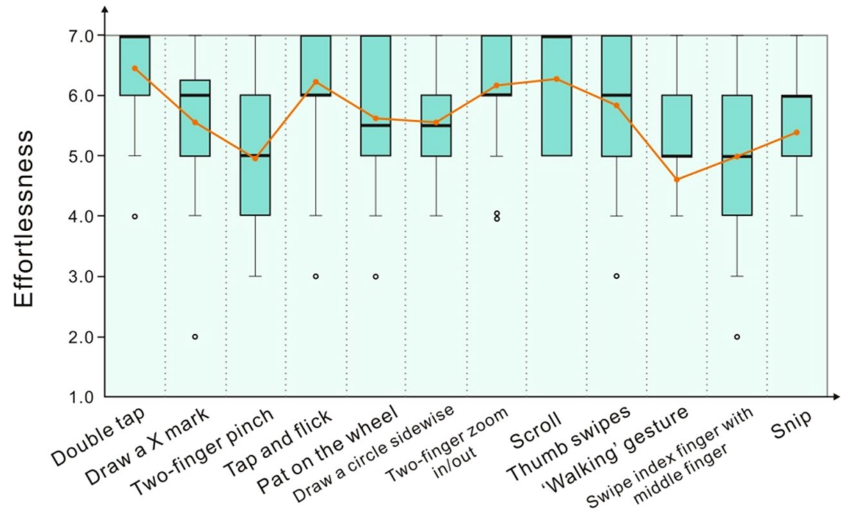

Fig. 14 Participants' evaluations of the effortlessness of micro-gestures

Regarding the effortlessness, a Friedman test revealed that there were statistically significant differences of user rating between the gestures [$\chi^2_{(N=18)}=46.812$, $df=11$, $p<0.001$], see Fig. 14. "Two-finger Pinch" ($M=4.944$, SD = 1.110), "Swipe Index Finger with Middle Finger" ($M=5.000$, SD = 1.328) and "Walking" ($M=5.389$, SD = 0.979) were rated as the gestures which require more motor effort to be performed. As for the mean value, Double Tap ranked first in all the three aspects, since it is a very simple gesture and has been intensively practiced in daily use. However, the performance (i.e., number of errors, RT) of this gesture in dual-task test was not always the best.

6 Discussion

The experiment is set for evaluating performance of the micro-gestures that are designed to interact with grasped objects in the periphery of vision and cognition. Using the PDT, we carried out a test to measure how much the visual attention is interrupted by the finger motions. By doing so the usability of micro-gestures in a driving scenario can be initially explored.

First of all, the learnability of micro-gestures was evaluated. It was found that participants took the shortest time to learn the "Draw a Circle Sidewise". The participants more frequently committed mistakes in training session when answering the four task-to-gesture mappings: "Previous/Next"-"Thumb Swipes" (6 mistakes), "Save"-"Pat on the Wheel" (5 mistakes), "Scroll"-"Scroll" and "Copy and Paste"-"Two-finger Pinch" (4 mistakes). Interestingly, five participants performed the "Swipe" gesture habitually in the opposite direction to the "Show me the gesture" part, instead they all correctly identified the function "Thumb Swipes" represents. This implies that swiping to the left means going to the next page is a deeply entrenched belief for some users.

The RT and fixation count vary significantly with different gestures. It is advisable to broaden the thresholds of gestures to further lower user attention. Actually, videotaped data showed the participants had a tendency to fine-tune the gestures and lessen their movement range in repetitions of trials for increasing the performance. This could indirectly result in a continuous decrease in RT (R1-R2: Mean difference = 0.084, SE = 0.47, $p=0.079$; R1-R3: Mean difference = 0.168, SE = 0.46, $p<0.001$; R1-R4: Mean difference = 0.269, SE = 0.42, $p<0.001$) and a slow increase in fixation count (R1-R2: Mean difference = −0.085, SE = 0.24, $p=0.001$; R1-R3: Mean difference = 0.110, SE = 0.24, $p<0.001$; R1-R4: Mean difference = −0.114, SE = 0.23, $p<0.001$).

According to the analysis in Sect. 5.2, drivers will take the risk of overlooking much visual information (e.g., road signs, pedestrians or tail-lights) on the periphery of vision field in the duration of using this gesture vocabulary. By comparison, the information in the center area of the field of vision is much easier to be noticed. The micro-gestures do not weaken the visual attention of users. The overall results of PDT indicated that micro-gestures mainly affected the frequency and range of eye movements, rather than the need of paying more attention to the primary task.

As indicated by the results, the postures of grasping the wheel have a negligible impact on performing both the gestures and the PDT. In the dual-task test, participants were able to quickly and accurately give gestural responses and identify as many as possible signals, wherever they put the dominant hand on the steering wheel. Since the designers have considered how to adapt the gestures to different sizes of devices to the maximum in design stage, the micro-gestures have potential for use in on-device interaction with other smart objects and for allocating less attention to the peripheral interaction.

The results of subjective evaluation suggest an evident user preference for some micro- gestures. Participants have strong adherence in legacy-based interactions, because they do not need to specially learn such kind of gestural inputs and they felt less frustration in using legacy-based gestures in testing session. The participants are satisfied with the user experience of this gesture set ($M = 5.667$, SD = 0.840), and also believe it can facilitate a better driving experience via promoting the efficiency of multimodal inputs. From the final results, we are convinced that it is necessary to reconsider the "Walking", "Thumb Swipes", "Two-finger Pinch" and "Pat on the Wheel" gestures, in pursuit of finding better ones for their corresponding tasks.

6.1 Design implications

In this section, we summarize evaluation data into the following design implications for micro-gestures on the grasped objects.

To design an easy-to-understand and learnable micro-gesture, it is useful to create a mapping between the gestural action and semantic meaning of function by image symbols and embodied metaphors: image symbols are direct metaphors that can clearly indicate the function. For example, "Draw a Circle Sidewise" means drawing a HOME button in mid-air, indicating the "Go to the Main Menu" task. In the "Swipe Index Finger with Middle Finger" gesture the index finger is compared to volume bar, thus the fingertip is naturally seen as the maximum volume. Embodied metaphors are constructed based on user's sensory-motor experience. For example, "Walking" gesture is a case of comparing the walking pace to the fast-forward function, though it seems not very intuitive as the relationship between these two concepts is not close enough for users.

A gesture motion should not be a very natural or a very intentional hand movement, to be more exact, there is a trade-off between the familiarity and intentionality for designing intuitive gestures: after a long period of time of grasping an object people may subconsciously move fingers to soothe the hand stiffness or express emotion. An error will occur when an accidental action is recognized as a meaningful gestural input of certain tasks. Thereupon, in the design stage gestures like single tap, finger contacts or bending a finger were not selected. In addition, the middle finger and the ring finger have no independent flexor and extensor, and they are physically connected by the Connexus intertendineus. For this reason, designers should carefully consider the flexibility of micro-gestures which use the middle finger or the ring finger, and avoid to define any multi-finger gestures or hand poses.

The micro-gestures should not damage the safety of human driver: in general, a finger motion will not negatively affect the manoeuvre of the steering wheel, unless the finger(s) or the hand exerts a force unintentionally that makes the steering wheel turn. If it happens that the steering wheel is suddenly turned when driving on a straight lane, e.g., the super highway, the vehicle is in danger of side and rear-end collision. In the checklist of potential grasping gestures, some of such gestures are listed, e.g., the palm drags along or around steering wheel.

Design gestures that are friendly to drivers of both left and right-hand drive vehicles: for this purpose, it is suggested that (1) a micro-gesture should not specify the direction of finger movements if not necessary, e.g., draw a circle sidewise, and (2) the moving trajectory of fingers or the hand is better to be shorter, so the gesture can be easier for a non-dominant hand to perform.

6.2 Future work

It should be noticed that the diverse usage scenarios, quality of drivers, and the experiment materials may become factors that bring about different evaluation results. Future research work could focus in the following.

First, based on the study regarding the steering wheel, we intended to propose a gesture set that could fit most of the graspable physical devices and thus to reduce users' learning burden. The physical devices differ widely in the usage scenarios and affordances. For example, a bicycle brake can limit specific hand movements. It is not known if this gesture set can be intuitively and freely used on the handles which have special shapes.

Second, the user study in this research is biased towards novice drivers, but does not concern more experienced drivers as they normally have better driving skill, more personal driving habits and different visual search strategy. A study of comparing novices and skillful drivers in terms of performing micro-gestures can be planned as future work.

Third, the research method used in the evaluation study was Wizard-of-Oz. Considering the necessity of analyzing the recognition technology for micro-gestures and the user performance in real situations, the in-situ deployment of high-fidelity prototypes should be implemented to refine the threshold values for gesture recognition.

7 Conclusion

This paper verified the learnability of a micro-gesture vocabulary, and tested the intuitive use of the vocabulary by asking users to complete a high-attentional primary task in the same time. The twelve micro-gestures, as a whole, turn out to be useful in general dual-task situations and promising for different input devices. However, they differ from each other in both learnability and performance. Despite grasping the steering wheel not being so necessary for driving as the self-piloting technology becomes increasingly mature, driving experience is still an important value people seek for. The development of tangible interaction interfaces, smart home devices and wearable computers will further the need for more tangible gestures. It is hoped that this study can provide empirical evidence on the design of micro-gestures for grasped objects.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study protocol conforms to the ethnical standard of the institutional committee and to the declaration of ethical principles adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964 and its later amendments. The Queen Mary Ethics of Research Committee approved the study protocol which includes all procedures performed in this study with the following reference number QMREC1749a.

Informed consent Informed consent was obtained from all individual participants involved in the study. This paper does not include any studies with animals conducted by any of the authors.

References

1. Alpern, M., Minardo, K.: (2003). Developing a car gesture interface for use as a secondary task. In: CHI EA'03, Florida, USA, pp 932–933
2. Angelini, L., Caon, M., Carrino, F., Carrino, S., Lalanne, D., Khaled, O.A., Mugellini, E.: WheelSense: Enabling Tangible Gestures on the Steering Wheel for In-Car Natural Interaction, 8005, pp. 531–540. Springer, Berlin Heidelberg (2013)
3. Angelini, L., Carrino, F., Carrino, S., et al.: (2014). Gesturing on the steering wheel: a user-elicited taxonomy. In: Proceedings of AutomotiveUI'14, pp. 1–8, Seattle, WA, USA
4. Angelini, L., Baumgartner, J., Carrino, F., et al.: (2016). Comparing gesture, speech and touch interaction modalities for in-vehicle infotainment system. In: IHM'16, Fribourg, Switzerland
5. Ashbrook, D.: (2007). Supporting mobile microinteractions. Georgia Institute of Technology. Dissertation
6. Bach, K.M., Jæger, M.G., Skov, M.S., Thomassen, N.G.: (2008). You can touch, but you can't look: Interacting with in-vehicle systems. In: Proceedings of CHI' 08, pp. 1139–1148, Florence, Italy
7. Chan, E., Seyed, T., Stuerzlinger, W., Yang, X.D., Maurer, F.: (2016). User elicitation on single-hand microgestures. In: Proceedings of CHI'16, pp. 3403–3414, San Jose, CA, USA
8. Döring, T., Kern, D., Marshall, P., et al.: (2011). Gestural interaction on the steering wheel: reducing the visual demand. In: Proc. CHI'11, pp. 483–492, Vancouver, Canada
9. Endres, C., Dimitrov., S.: (2010). Using a Theremin for micro-gesture recognition in an automotive environment. In: Proceedings of AutomotiveUI'10, Pittsburgh, Pennsylvania
10. Geiser, G.: (1985). Man machine interaction in vehicles. ATZ, 87, pp. 74–77
11. González, I.E., Wobbrock, J.O., Chau, D.H., Faulring, A., Myers, B.A.: (2007). Eyes on the road, hands on the wheel: thumb-based interaction techniques for input on steering wheels. In: Graphics interface conference 2007, pp. 95–102, Montréal, Canada
12. Gupta, S., Molchanov, P., Yang, X., Kim, K., Tyree, S., Kautz, J.: (2016). Towards selecting robust hand gestures for automotive interfaces. In: Intelligent vehicles symposium, IEEE, pp. 1350–1357
13. Häuslschmid, R., Menrad, B., Butz, A.: (2015). Freehand vs. micro gestures in the car: driving performance and user experience. In: 3D user interfaces, 2015 IEEE symposium, pp. 159–160, France
14. Horst, A.R.A., van der Martens, M.H.: The peripheral detection task (PDT): on-line measurement of driver cognitive workload and selective attention. SAE Int. **634**(4), 73–89 (2010)
15. Hoven, V.D.E., Mazalek, A.: Grasping gestures: gesturing with physical artifacts. Artif. Intell. Eng. Des. Anal. Manuf. **25**, 255–271 (2011)
16. Koyama, S., Sugiura, Y., Ogata, M., Withana, A., Uema, Y., Honda, M.: (2014). Multi-touch steering wheel for in-car tertiary applications using infrared sensors. In: Proceedings of AH'14, pp. 1–4, Kobe, Japan
17. Lee, S.H., Yoon, S.O., Shin, J.H.: (2015). On-wheel finger gesture control for in-vehicle systems on central consoles. In: Proceedings of Automotiveui'15, pp. 94–99, Nottingham, United Kingdom
18. Loclair, C., Gustafson, S., Baudisch, P.: (2010). PinchWatch: A wearable device for one-handed microinteractions. In: Proceedings of MobileHCI 2010, Lisbon, Portugal
19. Loehmann, S., Knobel, M., Lamara, M., Butz, A.: (2013). Culturally independent gestures for in-car interactions. In: Kotzé, P., et al. (eds.): INTERACT 2013, Part III, LNCS 8119, pp. 538–545
20. Ma, T., Wee, W., Han, C.Y., Zhou, X.: (2011). A method for single hand fist gesture input to enhance human computer interaction. In: Proceedings of HCI'13, pp. 291–300, Las Vegas, USA
21. Mahr, A., Endres, C., Schneeberger, T., Müller, C.: (2011). Determining human-centered parameters of ergonomic micro-gesture interaction for drivers using the theater approach. In: Proceedings of AutomotiveUI'11, pp. 151–158, Salzburg, Austria
22. Matthies, D.J.C., Lecolinet, E., Perrault, S.T., Zhao, S.: (2014). Peripheral microinteraction for wearable computing. In: Workshop on peripheral interaction: shaping the research and design space at CHI 2014
23. Müller, C., Weinburg, G.: Multimodal input in the car, today and tomorrow. IEEE Multimed. **18**(1), 98–103 (2011)
24. Murata, Y., Yoshida, K., Suzuki, K., Takahashi, D.: (2013). Proposal of an automobile driving interface using gesture operation for disabled people. ACHI, pp. 472–478
25. Neßelrath, R., Moniri, M.M., Feld, M.: (2016). Combining speech, gaze, and micro-gestures for the multimodal control of in-car functions. In: 12th International Conference on Intelligent Environment, pp. 190–193, London, UK
26. Nielson, M., Störring, M., Moeslund, T.B., Granum, E.: A procedure for developing intuitive and ergonomic gesture interfaces for HCI. Lect. Notes Comput. Sci. **17**(17), 1445–1453 (2003)
27. Oh, B.H., Hong, K.S.: Finger gesture-based three dimension mobile user interaction using a rear-facing camera. Int. J. Multimed. Ubiquitous Eng. **8**(5), 119–130 (2013)
28. Penn, F.: (2015). Micro-gesture interaction on the gear shift. Universität Passau, Dissertation
29. Pfleging, B., Schneegass, S., Schmidt, A.: (2012). Multimodal interaction in the car: Combining speech and gestures on the steering wheel. In: Proceedings of AutomotiveUI'12, Portsmouth, NH, USA
30. Riener, A., Ferscha, A., Bachmair, F., Hagmüller, P., Lemme, A., Muttenthaler, D.: (2013). Standardization of the in-car gesture interaction space. Proceedings of AutomotiveUI'13, pp. 155–162, Eindhoven, Netherlands

31. Stecher, M., Baseler, E., Draxler, L., Fricke, L., Michel, B., Zimmermann, A., Bengler, K.: (2015). Tracking down the intuitiveness of gesture interaction in the truck domain. In: 6th International conference on applied human factors and ergonomics (AHFE 2015) and the affiliated conferences, Vol. 3, pp. 3176–3183
32. Tan, Y., Yoon, S.H., Ramani, K.: (2017). BikeGesture: user elicitation and performance of micro hand gesture input for cycling. In: CHI EA'17, pp. 2147–2154, Denver, CO, USA
33. Wagner, J., Lecolinet, E., Selker, T.: (2014). Multi-finger chords for hand-held tablets: recognizable and memorable. In: Proceedings of CHI'14, pp. 2883–2892, Toronto, ON, Canada
34. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: (2009). User-defined gestures for surface computing. In: Proceedings of CHI'09, pp. 1083–1092
35. Wolf, K., Naumann, A., Rohs, M., Müller, J.: (2011). A taxonomy of microinteractions: defining microgestures based on ergonomic and scenario-dependent requirements. In: Campos, P., et al. (eds.), INTERACT 2011, Part I, LNCS 6946, pp. 559–575
36. Wolf, K., Schleicher, R., Kratz, S., Rohs, M.: (2013). Tickle: a surface-independent interaction technique for grasp interfaces. In: Proceedings of TEI'13, pp. 185–192, Barcelona, Spain
37. Wolf, K.: Microgestures—enabling gesture input with busy hands. In: Bakker, S., et al. (eds.) Peripheral Interaction, pp. 95–116. Springer, Switzerland (2016)
38. Zaiti, I.A., Pentiuc, S., G. and Vatavu, R.D.: On free-hand TV control: experimental results on user-elicited gestures with leap motion. Pers. Ubiquitous Comput. **19**, 821–838 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.