Culturally Independent Gestures for In-Car Interactions

Sebastian Loehmann[1], Martin Knobel[2], Melanie Lamara[2], Andreas Butz[1]

[1]University of Munich (LMU), HCI Group, Amalienstr. 17, 80333 Munich, Germany
[2]BMW Group Research and Technology, Hanauer Str. 46, 80993 Munich, Germany

sebastian.loehmann@ifi.lmu.de, andreas.butz@ifi.lmu.de,
martin.knobel@bmw.de, melanie.lamara@bmw.de

Abstract. In this paper we report on our ongoing work to introduce freehand gestures in cars as an alternative input modality. Contactless gestures have hardly been successful in cars so far, but have received attention in other contexts recently. We propose a way to achieve a better acceptance by both drivers and car manufacturers. Using a four-step process, we developed a small set of culturally independent and therefore easy-to-learn gestures, which can be used universally across different devices. We built a first prototype using distance sensors to detect the stop gesture in front of several devices. We conducted a user study during actual driving situations, testing the pragmatic and hedonic quality of the approach as well as its attractiveness. The results show a high acceptance of our approach and confirm the potential of freehand gestures as an alternative input modality in the car.
Keywords: freehand gestures, automotive user interface, culturally independent

1 Introduction

Gestures have always been an important part of interpersonal communication. They are closely connected to speech, giving a deeper insight into speakers' thoughts [7]. When communication partners speak diverse languages, gestures help them to increase their mutual understanding. Furthermore, gestures are essential in situations in which other communication channels cannot be used. Examples are the gestures performed by divers or the sign language created for the deaf.

With the introduction of Microsoft Kinect[1] and the Leap Motion Controller[2], freehand gestures have moved into the focus of HCI research again. These sensors can be seen as an enabling technology bringing gestures into researchers' labs and people's living rooms by its easy handling and low cost. Being a novel input modality, gesture-controlled interfaces seem to stimulate people in a positive way and to evoke feelings like joy and pleasure. Another reason for the success of this type of interaction is its ease of use. Since gestures are monitored through the human's proprioception and can in many cases even be performed blindly, they need less visual attention compared to haptic controls such as buttons.

[1] http://www.microsoft.com/en-us/kinectforwindows/
[2] https://www.leapmotion.com/

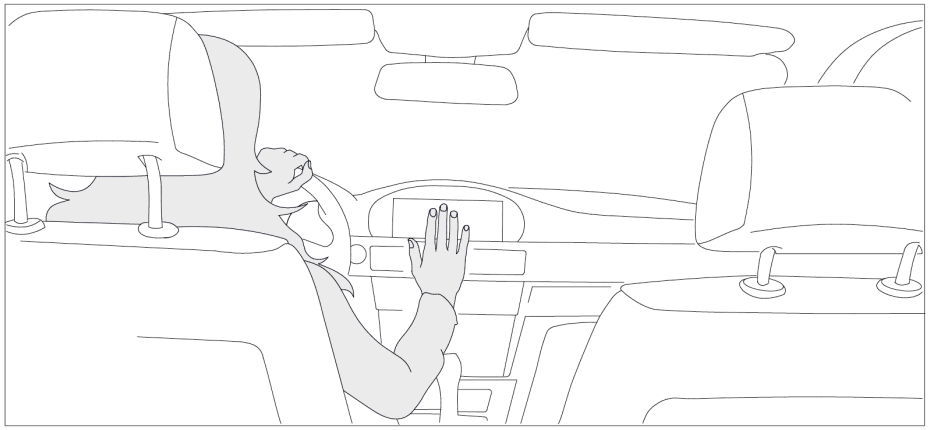

Fig. 1. Using freehand gestures as an input modality in cars

It has been shown that gestural input (see Fig. 1) significantly reduces drivers' visual distraction compared to haptic input [4]. This implies a potential increase in traffic safety by helping drivers to keep their eyes on the road [2], which makes gestures interesting again for car manufacturers. As a result of the increased interest, a first gestural interface has been introduced. The contactless tailgate opener [3] allows passengers to open the trunk by performing a foot gesture below the rear bumper. However, there is currently no interior gestural interface used in practice.

Taking into account lessons learnt from previous work on gestural interfaces and the limitations of the automotive context, we propose a novel approach for in-car gestures. Our goal is to increase their acceptance by manufacturers and drivers, which is often a challenge when moving interactive systems from the lab into practice.

In a pre-study, we identified a set of culturally independent gestures. The stop gesture was then successfully tested, providing early insights into drivers' acceptance of the approach. We built a prototype using infrared distance sensors and conducted a study in real driving situations, comparing the use of haptic controls and the freehand gesture in a quantitative and qualitative way as supposed in [11].

2 Related Work

In the past, several efforts (e.g. [1, 13]) have been made to bring gestures into the car. In spite of positive user feedback during studies, they were not commercially successful. Some drawbacks were caused by hand tracking based on infrared-cameras. Despite the complex software and expensive hardware, it is sensitive to varying lighting conditions [1], e.g. when driving out of a tunnel into direct sunlight. Furthermore, gestures are performed in a restricted interaction space above the gearshift [1], preventing direct interaction when controlling the dashboard's different devices. Another drawback seemed to be the rather large set of 22 gestures [13]. The need of remembering and recalling the correct movements might lead to a decreased acceptance.

More recent work introduced the Kinect to the automotive context. The combination of RGB- and depth-camera based tracking allows for a more robust gesture tracking and "is tolerant to the changing lighting conditions" [9]. In [10] the Kinect was mounted on the car's ceiling, successfully tracking gestures in the gearshift area. According to earlier work, direct interaction is not supported: all hand movements are tracked in the same spot but are used to control different applications. In [9] the sensor was placed on the dashboard and gestures were executed in front of the windshield. However, hand movements visible from the outside can cause misunderstandings and thus dangerous traffic situations. None of the systems described above were tested under realistic driving circumstances.

Regarding the creation of a gesture set, [12] stressed the importance of including end users into the design process. Nevertheless, [10] proposed to "balance user needs with technological requirements and limitations".

3 A New Approach

Taking lessons learnt from related work and the goal of using culturally independent gestures into account, our concept implemented the following principles:

Reduced Distraction. We utilize a general advantage of freehand gestures: As the movement needs to be less precise than finding and using a haptic control element, the interaction requires less hand-eye coordination and reduces visual distraction [4].

Direct Interaction. In prior automotive gestural systems, gestures were performed in a fixed interaction space [1, 9, 10, 13]. In contrast, we allow movements close to the intended device, enabling direct interaction. Additionally, visibility of the gestures to other traffic participants [9] can hereby be prevented, reducing the risk for possible misunderstandings.

Small Set of well-known Gestures. In [10] cultural differences were ignored to simplify the definition of the gesture set. We propose a process to identify few well-known and culturally independent gestures to support intuitiveness and achieve a low mental load [11]. According to [12], we included input by potential users but at the same time chose gestures to fit the specific automotive context as proposed by [10].

Universal Use of Gestures. We propose the application of gestures usable universally across different systems (e.g. radio or navigation). We believe that this will result in a good learnability, leading to instant usage and a higher acceptance by drivers.

4 Finding Universal Gestures

In order to find a small set of culturally independent gestures, we followed a four-step process. First, we collected 34 gestures in the corresponding literature (e.g. [7, 8]). They were chosen because of their explicit but also culturally independent meaning. All gestures are permanently used during interpersonal communication [7, 8]. Thus, they should be easy to understand and to remember.

Fig. 2. Reduced Gesture Set. Upper row from left: Slow Down, Clap and Fist. Lower row from left: Go Away, Stroke and Stop.

To reduce the initial set, we conducted a study involving possible end users (five participants, two female) in the second step. For each gesture, they received detailed descriptions, performed the corresponding movement and were interviewed about their subjective meaning. Using this information, we eliminated ambiguous gestures, resulting in a set of 18 gestures, each of them having a unique interpretation.

In the third step, four automotive experts (one female) were asked to rate the remaining 18 gestures. The goal was to identify redundancies concerning similar movement profiles. We produced 72 video clips showing each gesture from four different angles. During this expert review, twelve gestures were removed. An example is the similarity of the greeting gesture (whole hand, fingers close together) to the stop gesture (whole hand, fingers spread, see Fig. 3 on the right). The remaining set consists of six gestures, each unique in meaning and movement profile (see Fig. 2).

In the fourth and final step, the remaining gestures were rated by 91 participants (45 female) via an online survey. The results provided deeper insights into explicit meaning, implicated emotions as well as the compatibility of each gesture to the automotive context. To collect first evidence towards the acceptance of the identified gestures, we tested one of them in an early prototype. Based on the results of the online questionnaire, we chose to implement the stop-gesture (see Fig. 2, lower right).

5 User Study

We conducted a study to collect first evidence concerning the acceptance of the described approach. The experiment took place under real driving conditions. Our first goal was to determine if a device (e.g. radio) could be controlled by a gesture without accidentally triggering a spatially close second device (e.g. air conditioning). Second, we wanted to test whether participants preferred haptic or gestural input for the given tasks. Combining both quantitative and qualitative results, the goal was to verify whether a high acceptance for the use of the gesture was achieved by the approach.

5.1 Prototype

To detect the hand in front of interactive in-car devices, Sharp distance sensors emitting infrared light were installed on the dashboard of a car (see Fig. 3 left). The task was to interact with three different devices: The universal meaning of the stop-gesture was used to mute the radio, to turn off the ventilation of the air conditioning and to stop the navigation system. A stop-gesture was recognized, if the two sensors above and below a specific device were covered at the same time (see Fig. 3 right). All sensors were powered by an Arduino board, which was connected to a laptop. The system was able to recognize the stop-gesture for all three devices. Though, because of technical constraints, the corresponding car functions could not be triggered directly. Instead, the supervisor controlled the setup from the backseat. Whenever a gesture was recognized by the system, he triggered the corresponding functions via a laptop connected to the car's internal bus.

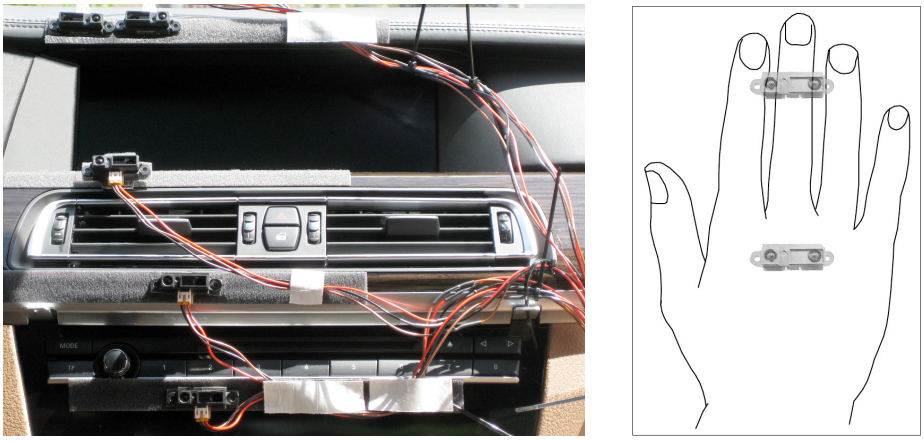

Fig. 3. Distance Sensors on the Dashboard of a car (left) detecting a Stop-Gesture (right)

5.2 Study Design

Participants were instructed to drive a predefined route in city traffic covering 9 kilometers, taking approximately 15 minutes. While driving, the three tasks (muting the radio, turning off the ventilation and stopping the navigation) were executed either by gestural or haptic input. The study followed a within-subjects design. Thus, each task was performed twice alternating the input modality. To avoid order effects, the sequence of the input modalities was counterbalanced. In a third round, both modalities were possible, giving participants the choice for every task. After each of the first two rounds, an AttrakDiff2 questionnaire [6] had to be completed. After round three, a questionnaire concerning the comparison of both modalities was provided.

5.3 Method

A German version of AttrakDiff2 [6] was used to evaluate the acceptance of both, gestural and haptic input. To evaluate the general appeal, we added the Attractiveness (ATT) scale from AttrakDiff1 [5]. In Hassenzahl's model [6] ATT consists of pragmatic quality (PQ), which is comparable to usability's "quality of use" [6] and hedonic quality (HQ), indicating the potential to fulfill pleasure in use and ownership. HQ is subdivided into stimulation (HQS, e.g. personal development) and identification (HQI, e.g. expressing oneself to others) [6]. The AttrakDiff2 including ATT comprises 28 seven-point semantic differential items, in each case seven for PQ (e.g. technical-human), HQI (e.g. isolating-connective.), HQS (e.g. conservative-innovative) and ATT (e.g. repelling-appealing).

Cronbach's alpha[3] for haptic interaction ranges between satisfying .66 (HQI) and .87 (ATT), both HQ values for gestures are not acceptable. After two HQI items (unprofessional-professional, tacky-stylish) were excluded based on an item analysis, Cronbach's alpha changes to .69 for both interaction types. Despite the inacceptable internal consistency of HQS for gestures (.45), we decided to include this scale exceptionally because composite reliability of .68 for both interaction types is acceptable.

Discriminant validity analysis show that ATT intercorrelates significantly with PQ and HQI between $r=.58$, $p<0.01$ and $r=.76$, $p<0.001$ for haptics and ranges between $r=.60$, $p<0.01$ and $r=.78$, $p<0.001$ for gestures, supporting Hassenzahl's model. For gesture interaction an additional significant correlation between PQ and HQI ($r=.66$, $p<0.01$) was found, whereas the averaged intercorrelation of HQI is lower. Since the composite correlation of PQ and HQI ($r=.44$) is higher than between HQI and HQS ($r=.25$) - both subscales of the higher order hedonic quality scale - we decided to analyze only the main factors ATT, PQ and HQ. As a side note: this study evaluates two interaction types for three functions instead of a general product. Thus, the limitations of the questionnaire might be a consequence of the narrowed interaction context.

5.4 Results

The 20 participants (two female) had an average age of 37 years, ranging from 26 to 57. All of them drive at least once a week, 85% of them more than 5000 kilometers a year. Results of AttrakDiff2 show that gestural interaction strengthens attractiveness and hedonic quality significantly compared to haptic interaction (see CIs in Table 1). Supporting previous results, there are no significant inter-correlations between both interaction forms. Findings suggest that well-known gestures like stop (see Fig. 3) are a desired interaction form in the automotive context. No significant Pragmatic differences show that both interaction types fulfill the intended functional requirements and are thus equally accepted. The comparison of the *technical-human* item between haptics (M=-1.35, SD=1.04) and gestures (M=1.15, SD=1.23) shows a significant difference (t=-7.43, df=19, $p<0.001$, d=1.66), indicating a gesture driven natural interaction.

[3] Cronbach's alpha is a psychometric criterion for test-quality and determines the internal consistency of a scale, respectively the correlation between the particular items.

The unpredictable-predictable item differs significantly (t=4.53, df=19, p<0.001, d=1.01), supporting haptics (M=2.20, SD=0.62) over gestures (M=0.75, SD=1.25). One reason is the participants' lack of experience in using gestures. All in all AttrakDiff2 results show that the gesture is accepted and favored in these use cases.

Table 1. Mean intensity (standard deviation, 95% confidence interval) Attractiveness (ATT) and AttrakDiff2 main factors Hedonic quality (HQ without unprofessional-professional and tacky-stylish) and Pragmatic quality (PQ).

N=20	Haptics (95% CI)			Gestures (95% CI)		
	M (*SD*)	Lower	Upper	M (*SD*)	Lower	Upper
ATT (7)	0.69 (*0.70*)	0.36	1.01	1.44 (*0.76*)	1.08	1.79
PQ (7)	1.02 (*0.71*)	0.69	1.36	1.30 (*0.83*)	0.91	1.69
HQ (12)	-0.48 (*0.57*)	-0.75	-0.21	1.27 (*0.38*)	1.09	1.45

To conclude the study, participants filled out a questionnaire. Being asked which input modality they liked better, 14 favored the gesture, 2 haptics and 4 both. This indicates the acceptance of the stop-gesture and reconfirms the AttrakDiff results. 16 stated that gestural input helped to focus on traffic (1 haptics, 3 both), indicating a potential reduction of visual distraction. Nobody said that gesturing influenced traffic safety negatively. All participants considered gestures to be a valuable alternative input modality and a helpful feature in the car. In round three, participants were able to choose between input modalities. Concerning the navigation task, gestural input was chosen in 30 out of 40 cases (ventilation 37 out of 40 and radio 24 out of 40).

Concerning the hardware, we discovered no influence of the lighting conditions throughout the study. Gesture detection was robust as no intention to perform a gesture was unsuccessful. Participants were continuously interacted with the correct device, meaning that the spatially close functions were easily discriminable.

6 Conclusions and Future Work

During real driving situations, we tested a first implementation of a new approach towards using freehand gestures as an alternative input modality in the automotive context. We tracked interactions of the driver's hand with devices in the dashboard. The stop-gesture was exemplarily used because of its universal use across devices and its cultural independency. Participants of the study performed the stop-gesture to turn off certain functions (e.g. mute the radio).

After exploring both haptic and gestural input, most drivers chose the gesture because of less presumed visual distraction and ease of use. Results of the AttrakDiff2 showed significantly higher attractiveness and hedonic qualities of gestural input. Furthermore, according to qualitative feedback, gesturing did not influence traffic safety negatively, but in a rather positive way. We can conclude that we found a high acceptance of our approach to gestural input in the case of a stop-gesture. This can be seen as an indication for the potential of choosing culturally independent freehand gestures to be an alternative input modality in automobiles in the near future. The results of this pilot-study motivate us to further develop our suggested approach.

As a next step, the improvement of the prototype will be most important. Due to the Wizard-of-Oz technique, there was a negatively rated delay of the functional feedback (e.g. the actual muting of the radio) after performing a gesture. This can be avoided by controlling the car's features directly via the internal bus so that no active supervisor is needed. Furthermore, we are considering the use of additional feedback. Some participants stated that the functional feedback (e.g. the sound of the radio) alone is not sufficient and for instance an audio signal could be helpful. We also noticed the request for more gestures. One driver spontaneously tried to turn the ventilation back on by waving the hand from the fan towards the body. Besides the fact that this is a positive observation concerning the acceptance of gestures, we will add functionality and further appropriate gestures to the prototype. For the implementation, we will need to improve tracking by using more flexible sensors like Kinect.

We are aware that the study setup with the use of one gesture is only a first step towards a comprehensive system. However, we are confident that the suggested approach can have a positive influence on the driver's visual distraction in the future.

7 Acknowledgments

We thank Simone Rodestock for her important contributions.

References

1. Akyol, S., Canzler, U., Bengler, K. and Hahn, W. Gesture control for use in automobiles. In Proc. IAPR 2000, pp. 28-30.
2. Alpern, M. and Minardo, K. Developing a car gesture interface for use as a secondary task. In Proc. CHI 2003 ext. abstracts, pp. 932-933.
3. BMW: Contactless Tailgate Opening, available at http://www.bmw.com/com/en/newvehicles/3series/sedan/2011/showroom/comfort/contactless_tailgate_opening.html, accessed May 15th 2013.
4. Geiger, M., Zobl, M., Bengler, K. and Lang, M. Intermodal Differences in Distraction Effects while Controlling Automotive User Interfaces. In Proc. HCII 2001, pp. 263-267.
5. Hassenzahl, M. The Effect of Perceived Hedonic Quality on Product Appealingness. International Journal of Human-Computer Interaction, 13 (4), pp. 481-499 (2001).
6. Hassenzahl, M. The interplay of beauty, goodness, and usability in interactive products. Human-Computer Interaction 19, 4, pp. 319-349 (2004).
7. McNeill, D. Hand and mind: What gestures reveal about thought. University of Chicago Press, Chicago (1995).
8. Morris, D. Gestures: Their Origins and Distribution. Jonathan Cape, London (1979).
9. Rahman, M., Saboune, J. and El Saddik, A. Motion-path Based in Car Gesture Control of the Multimedia Devices. In Proc. DIVANet 2011.
10. Riener, A. Hand and Finger Gestures in Vehicular Applications. Computer, 45 (4), (2012).
11. Wachs, J. P., Kölsch, M., Stern H., and Edan, Y. Vision-Bases Hand Gesture Applications. Communications of the ACM, 54 (2), pp. 60-71 (2011).
12. Wobbrock, J. O., Morris, M. R. and Wilson A. D. User-defined gestures for surface computing. In Proc. CHI 2009, pp. 1083-1092.
13. Zobl, M., Geiger, M., Schuller, B., Lang, M. and Rigoll, G. A real-time system for hand gesture controlled operation of in-car devices. In Proc. ICME 2003, pp. 541-544.