I See Your Point: Integrating Gaze to Enhance Pointing Gesture Accuracy While Driving

Florian Roider[1,2], Tom Gross[2]
[1]BMW Group Research, New Technologies, Innovations, Munich, Germany
[2]Human-Computer Interaction Group, University of Bamberg, Bamberg, Germany
florian.roider@bmw.de, tom.gross@uni-bamberg.de

ABSTRACT
Mid-air pointing gestures enable drivers to interact with a wide range of vehicle functions, without requiring drivers to learn a specific set of gestures. A sufficient pointing accuracy is needed, so that targeted elements can be correctly identified. However, people make relatively large pointing errors, especially in demanding situations such as driving a car. Eye-gaze provides additional information about the drivers' focus of attention that can be used to compensate imprecise pointing. We present a practical implementation of an algorithm that integrates gaze data, in order to increase the accuracy of pointing gestures. A user experiment with 91 participants showed that our approach led to an overall increase of pointing accuracy. However, the benefits depended on the participants' initial gesture performance and on the position of the target elements. The results indicate a great potential to support gesture accuracy, but also the need for a more sophisticated fusion algorithm.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI); Pointing; Gestural input;** *Empirical studies in HCI;*

Author Keywords
Attentive interfaces; automotive interface; gaze-added interface; mid-air gestures; multimodal fusion; pointing gestures

INTRODUCTION
Mid-air gestures have been repeatedly shown as a promising method for interaction with secondary functions in the vehicle while driving. In this context, it is very important to distinguish, which form of gesture interaction is used, because they vary in regards of usability and demands on the driver. Exisiting classifications differentiate between deictic gestures (i.e. pointing) and other forms of gestures, such as iconic, or

AutomotiveUI '18, September 23–25, 2018, Toronto, ON, Canada

DOI: https://doi.org/10.1145/3239060.3239084

Figure 1. Pointing gestures allow drivers to create references to all kinds of vehicle functions just by pointing at them.

metaphoric gestures [11]. For human-vehicle interaction, latter ones can also be classified as symbolic gestures. Symbolic gestures are pre-learned gesture shapes that users have to know beforehand to operate certain functions in the vehicle. An advantage of symbolic gestures is that they can be performed blindly, altough it has been shown that control glances occur to check the correct posture and positon of the hand. One downside is the effort for memorizing gesture commands. It becomes more difficult to remember the entire gesture set as the number of gesture-controlled functions increases [15]. An increased learning effort might be acceptable for expert users, but not for the majority of drivers. This limits the amount of in-vehicle functions that can be efficiently supported with symbolic gestures.

Pointing gestures, on the other hand, do not need to be learned by the user. Pointing creates a simple deictic reference to all kinds of real and on-screen objects (as shown in Figure 1). Users are enabled to interact with a wide range of vehicle functions without having to learn new gestures, which is particularly helpful for novice users [1]. During the execution of a pointing gesture, users have to localize a pointing target and make a coordinated pointing movement with their hands. Compared to symbolic gestures, this requires a greater amount of the users' visual attention. However, with regard to the advances in autonomous driving and the increasing number of driver assisting functions in modern vehicles, increased visual attention is acceptable, when user experience, effectiveness and ease-of-use for operating secondary functions are increased in return.

Recent experiments have shown that drivers make relatively large pointing errors while driving [6, 17]. This is in line with findings from earlier experiments, which found that pointing errors especially occurred, if users cannot move their head towards the target [5]. The authors conclude that the data of eye-gaze fixations provides information about the pointing target, before the arm movement has event started. In fact, it has been found that the user's gaze is actually anchored to the pointing target during a pointing movement [12]. Just recently, this has also been shown in an experiment in the automotive domain. The drivers' eye-gaze is fixed on pointing targets during freehand pointing while driving [1]. This knowledge suggests that the close relationship between eye-gaze behavior and pointing gestures might be used to improve pointing performance while driving, without creating additional (visual and cognitive) load on the user.

In this paper, we describe a practical implementation of a simple algorithm that integrates gaze information with the aim to increase the accuracy of pointing gestures. Moreover, we report the results of a user experiment that reveals benefits and downfalls of the approach.

RELATED WORK

The combination of gaze information with gestures was examined in a number of HCI experiements that explicitly use gaze information to select objects on a screen and gestures to manipulate selected objects. Chatterjee et al. showed how the combination of gaze and gesture input can overcome gaze-only or gesture-only systems [7]. Zhang et al. presented a similar approach that enhanced the interaction efficiency compared to a gesture-only interface, but they also encountered problems regarding the participants eye-hand coordination [20]. Similar approaches have been presented in the automotive domain. Nesselrath et al. used gaze information to select real objects of the vehicle, such as side mirrors or windows. Gestures on the steering wheel or speech commands could then be used to control these objects [13]. Kern et al. showed the application of gaze information to select objects on a screen in combination with a haptic button on the steering wheel to confirm the selection [9].

All these prototypes have in common that they share the same approach for the integration of gaze information. It is used as an conscious, active selection tool, combined with a second modality for modification. Gaze input replaces the function of another input mode (e.g. touch or mouse input for selection). Salvucci et al. point out two problems that emerge from such gaze-based interfaces: the noise and limited availability of eye tracking information, and the dissociation between the user's glance behavior and the actual visual attention [18]. Especially the latter one, is very relevant in the automotive context. Driving is typically a dual-task situation with driving the car as the primary task, and operating non-driving related functions as the secondary task. Since steering a vehicle is visually very demanding, the driver's glances may be directed towards the street, although the mental attention is on the completion of a secondary task. Therefore, the authors propose the usage of gaze-added interfaces, which provide the same basic functionality as non-gaze interfaces, but add the ability to incorporate gaze information, if available [18].

Zhai et al. presented such a gaze-added prototype that combines passive gaze information with active mouse input. It passively tracked users' eye movements to predict the pointing target of the mouse and uses this information to enhance the movement of the mouse cursor [19]. Oviatt et al. describes this form of multimodal integration as *blended* multimodal interaction. The passive input mode is used to improve the multimodal system's prediction and interpretation of the active mode [14].

A number of studies over the last years presented other promising approaches on how to increase the accuracy of mid-air pointing gestures. Mayer et al. demonstrated how systematic displacement of different ray-casting approaches can be compensated using two-dimensional polynomials [10]. Plaumann et al. showed the influence of ocular dominance and handness on pointing gestures. They present a selection algorithm, which uses this information, to increase the users' pointing accuracy [16]. For pointing in the automotive domain the visual demand of the driving task, noisy sensor data, or unintended movement due to driving and road conditions further limit the effectiveness of pointing gestures. Ahmad et al. presented a Bayesian framework that takes additional sensory data from the vehicle (e.g. such as suspension travel data) into account in order to preditct freehand pointing targets. Though not evaluated, they also propose that eye-gaze data could offer valuable information on areas of interest on the display [4].

Summary

Pointing gestures enable drivers to interact with a wide range of vehicle functions, without the increased learning effort of symbolic gestures. While technical challenges to detect user's pointing direction exist, a more fundamental problem is that users pointing movements often lack sufficient accuracy to identify user intentions. Existing approaches make use of mathematical functions, users' ocular dominance or vehicle sensor data to increase pointing performance. Eye-gaze data has been shown to provide meaningful information about the users' attention, especially while performing finger pointing movements. In the automotive domain, gaze input has been used as an active selection tool, but it has not been used as a passive input modality to improve pointing gestures. We propose a first practical approach, how gesture data could be enhanced with passive eye-gaze data, in order to compensate for the lack of pointing accuracy.

PROTOTYPE

In the following section, we describe the physical setup of the prototype and the selection algorithm for the fusion of gesture and gaze data.

Apparatus

The setup of the prototype for the experiment is shown in Figure 2. In front of the mock-up, there was a 65 inch screen displaying a driving simulation. A Thrustmaster force-feedback

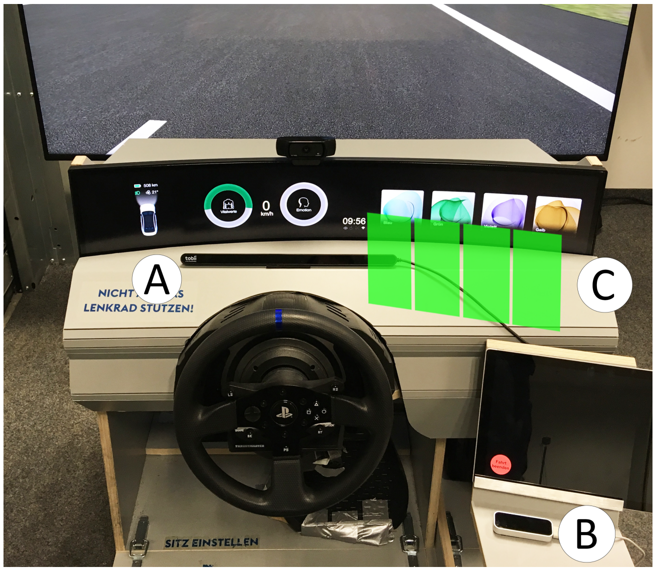

Figure 2. The vehicle mockup included a gesture camera and an eye-tracker. Selectable elements were displayed on the right side of the curved screen.

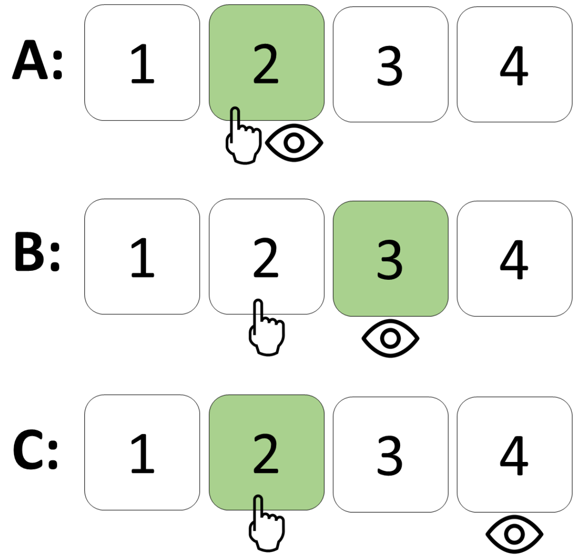

Figure 3. The algorithm distinguished three possible cases how to integrate gaze information for the selection.

steering wheel was used in combination with foot pedals for gas and brake to control the vehicle. There was a 34 inch LG curved screen integrated in the mock-up. It was placed approximately 34 cm behind the steering wheel. The display was sunk in the mock-up, so that only 18 centimeters of the screen were visible. The left part of the screen displayed information about the vehicle, such as speed. The right part of the screen displayed four selectable elements. Elements were squared with a side length of 6.75 cm x 6.75 cm. A Tobii 4C eye-tracker[1] (A) was placed centered behind the steering wheel to track users' gaze. The users' fingers were tracked with a Leap Motion[2] (B). It was placed to the right of the steering wheel on the middle console, so that it covers the mid-air gestures area. An important part of this area is the mid-air gesture interaction pane (C) in front of the four elements, which is used for selection. The tablet on the middle console was only used for instructions and collecting demographic data beforehand and was not part of the interaction during trials.

Selection Algorithm

In order to select one of the four elements on the screen, participants made a pointing movement with their right index finger by moving the outstretched finger towards the screen (see Figure 1). The mid-air gesture interaction pane was 25 cm wide and placed approximately 26 cm in front of the four selectable elements on the screen (C in Figure 2) and was tilted towards the driver. A selection was triggered when the fingertip of the index finger entered the interaction pane. It was horizontally split into four equally spaced portions (6.25 cm each). Based on these portions, the horizontal position of the fingertip in the moment of selection determined the gesture target element ($elem_{\text{Gesture}}$). Eye-gaze information was incorporated in the moment of the gesture selection. The user's gaze point on the screen determined which of the four elements the user is gazing at in the moment of selection ($elem_{\text{Gaze}}$).

This approach aims to integrate only passive gaze data, which means that users do not know that their gaze is taken into account. Thus, their gaze is not necessarily focusing the target element in the moment of selection (e.g. for control glances back to the street). In this case, integrating gaze data would falsify the selection. To avoid this, we assume that the gaze does not refer to the pointing selection, if $elem_{\text{Gesture}}$ and $elem_{\text{Gaze}}$ differ by more than one element. Based in these considerations, we apply an algorithm that uses three simple fusion rules to determine the selected element. It is illustrated in Figure 3.

$$diff = |elem_{\text{Gesture}} - elem_{\text{Gaze}}|$$

$A : diff == 0$
Both input modes indicate the same element. No correction is needed. $\Rightarrow$ Select $elem_{\text{Gesture}}$ (2).

$B : diff == 1$
Gesture and gaze indicated different, but neighboring elements. We assume that the gaze and gesture refer to the same target, but $\text{elem}_{\text{Gesture}}$ is wrong due to inaccurate pointing. $\Rightarrow$ Select $elem_{\text{Gaze}}$ (3).

$C : diff > 1$
Both elements differ by more than one position. We assume that the gaze does not refer to the target element. $\Rightarrow$ Select $elem_{\text{Gesture}}$ (2).

EXPERIMENT

We conducted a user experiment that examined the benefits of a gaze-added selection algorithm for mid-air pointing gestures while driving. In a number of pre-tests in a desk setup, we

[1] https://tobiigaming.com
[2] https://www.leapmotion.com/

observed that the benefit of our approach might depend on the participants' individual experience with gesture interaction and their resulting pointing performance. We decided to investigate the algorithm in a study with a a variety of different users, in order to get a general overview of the potential of the approach.

Participants

A total of 91 participants (28 females) between 17 and 66 years ($M = 35.52, SD = 11.79$) took part in our study. 66 of the participants reported that had never used any form of gesture interaction with computers before, 25 reported to already have some experience with the usage of mid-air gestures. 39 participants were driving daily, 48 participants were driving several times a month and four participants stated that they don't drive a car.

Study Design

Each participant completed one 5 minute trial in a driving simulator. They did a quick gaze calibration process, but were told that this was used to measure visual distraction. Therefore, participants did not know about the integration of gaze information for the selection. We wanted to avoid an artificial gaze behavior and instead focus on the benefit of naturally occurring eye-gaze information while pointing. During the ride, participants were repeatedly instructed to select one out of the four elements on the screen in the vehicle mock-up. The selection was made using gesture input by pointing in the direction of the selected element. Additionally, eye-tracking information was used to modify the users' gesture selection as described in the previous sections. Participants received visual feedback, which element was selected. The output element of the fusion algorithm was highlighted for one second and an earcon was played. The feedback did not explicitly differentiate between right or wrong selections. For each participant, we logged the percentage of correct selections with the gaze-added selection algorithm (fused accuracy). Additionally, we recorded the percentage of correct indications based on gesture information only (gesture accuracy). By comparing gesture accuracy and fused accuracy, we can determine the increase of pointing accuracy with the gaze-added system instead of an unimodal approach.

Procedure

At first, participants adapted the position of the seat (height and distance to the steering wheel) according to their size and length of arms, so that they could comfortably reach the interaction pane. Then they started the experiment by tapping on the tablet to the right (see Figure 2). Textual instructions on the tablet led the participants through the whole procedure. For the majority of the participants (85%) there was also an additional examiner who supervised the procedure. At first, demographic data was collected. In the next step, the eye-tracking system was calibrated. Participants had to focus two pulsing points on the screen to determine the left and right boundaries of the gaze interaction area. This was followed by a textual instruction on how to select elements on the screen, namely by pointing towards the requested element and physically crossing the virtual pane with the outstretched index finger. An example element appeared to practice this selection mechanism. If participants tapped too far left or right, an according hint was displayed ("e.g. Tap further left"). During this practice phase there was no integration of gaze information. After this selection was successfully completed, the driving simulation started automatically. The driving task was to follow a leading vehicle on a highway road with three lanes at 120 km/h. The road was slightly winding and there was only little traffic so that the leading vehicle stayed on the rightmost lane most of the time. During the first 40 seconds, no selections had to be made, so that participants could get used to the driving simulation. After that, participants completed nine selections. Each selection started with an acoustic notification and a textual instruction which element to select (e.g. "Select Blue"). Then the four elements appeared and participants tried to select the instructed element. The elements disappeared after a selection was made. The next selection started with a delay of 10 seconds.

Case	*All*	*Elem 1*	*Elem 2*	*Elem 3*	*Elem 4*
Case A	48%	60%	51%	50%	30%
Case B	17%	11%	17%	18%	22%
Case C	1%	1%	1%	3%	0%
No gaze	34%	29%	31%	29%	47%

Table 1. Occurence of the three cases in percent, summarized for all selections and depending on the target element. The last row shows the percentage of selections for which participants did not gaze at the UI.

RESULTS

We excluded about 17% of the total data set due to two reasons. First, there was a number of selections for which the eye-tracker did not output any gaze data. This might be caused by several reasons, such as failure of the tracking hardware, or by greater changes in posture of the participants after the eye-gaze calibration. We only analyzed selections with available gaze data, because the tracking performance was not focus of this experiment. Second, we observed some wrong selections, because participants could not remember the instructed element, when a certain amount of time passed between instruction and selection (because participants concentrated on driving). For this reason, we excluded selections from the analysis, if the selection time (from the moment of instruction until selection) was longer than the third quartile of all selections plus 1.5 * interquartile range and thereby classified as an outlier according to Tukey.

Table 1 summarizes the occurrence of the three cases described in Figure 3. In 48% of selections, gesture and gaze identified the same element (Case A). They differed by one element in 17% of selections (Case B). Deviations of more than one position occurred only in 1% (Case C). In the remaining 34%, the participants' gaze did not focus any UI element in the moment of selection. Although $elem_{\text{Gaze}}$ was not available and the algorithm could not be applied in these cases, they are still part of the reported results.

The following sections focus on the accuracy (percent of correct selections) that was achieved with the presented gaze-

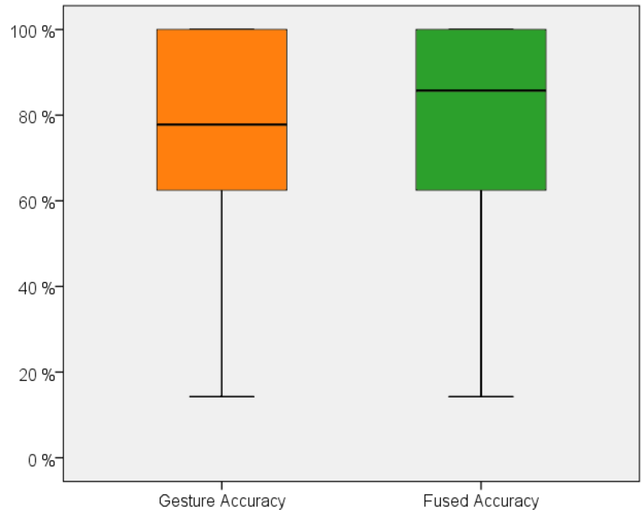

Figure 4. Pointing gesture accuracy (left plot) could be increased by incorporating gaze data, resulting in a slightly improved fused accuracy (right plot).

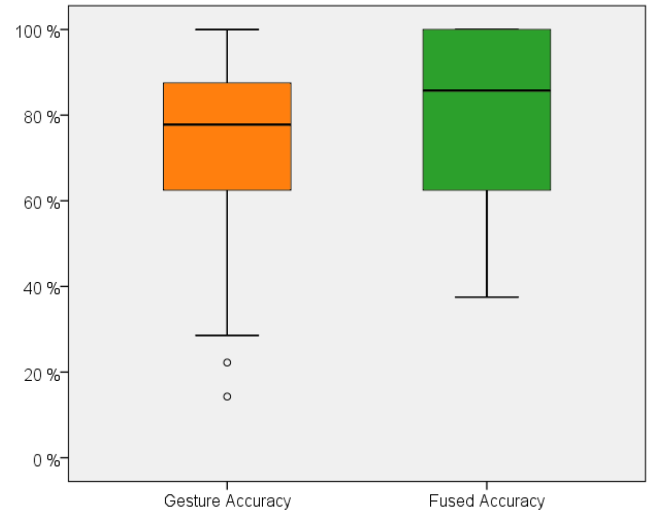

Figure 5. The increase of fused accuracy (right plot) compared to gesture accuracy (left plot) was greater for novice participants, who did not have any experience with gesture interaction.

added selection algorithm in comparison to the accuracy only based on gesture information. To measure the benefit of our approach we calculated the difference between the gesture accuracy and the fused accuracy for each participants (Δa).

$$\Delta a = \textit{fused accuracy - gesture accuracy}$$

Gesture Accuracy and Fused Accuracy

Gesture accuracy and the fused accuracy are illustrated in 4. The left boxplot shows that there was a big spread in gesture accuracy between participants. The mean gesture pointing accuracy was 75.52% ($Median = 77.78\%$) over all participants. Although the algorithm influenced only 17% of all selections (Case B in Table 1), it led to an improved fused accuracy of 80.56% ($Median = 85.71\%$). A Wilcoxon signed-rank test showed that the increase was significant ($Z = -2.12, p < .05, r = 0.22$). The data was not normally distributed.

The overall Δa is 5.04%, which is composed of 10.05% improved selections, but also 5.33% selections where the gesture indicated the correct element, but the gaze element was selected. Furthermore, the increase of accuracy was not equally distributed among participants. For 42.86%, there was neither an increase nor a decrease of accuracy. 37.36% of the participants were overall supported by the algorithm, but for 19.78% the algorithm led to a decrease of selection accuracy.

Influence of Experience

One strength of pointing gestures is its simplicity, which makes it particularly suitable for novice user [1]. We hypothesized that our algorithm will be more valuable for those people who are less experienced with gesture interaction. Therefore, the participants reported their level of previous experiences with gestures interaction at the beginning of the experiment.

Based on this data, we compare the pointing gesture performance and the benefit of the algorithm for two different groups of participants: Novice users, who reported to have no experience with gesture interaction (66 out of 91 participants), and experienced users, who reported to already have some experience with gestures interaction (25 out of 91). The group of novice users is shown in Figure 5. Gesture accuracy was 74.57% ($Median = 77.78\%$). Additional gaze information increased the mean accuracy by 7.14% up to 81.45% ($Median = 85.71\%$). The data was not normally distributed. A Wilcoxon signed-rank test showed that the improvement was significant ($Z = -2.38, p < .05, r = 0.25$).

For the experienced group of participants, the fused accuracy of 78.03% ($Median = 87.50\%$) showed no improvement compared to a gesture accuracy of 78.20% ($Median = 75.00\%$). Accordingly, a Wilcoxon test showed no significant differences between both variables in this case ($Z = -0.35, ns.$)

Influence of Element Position

Based on the results of previous studies, we assumed that the horizontal position of the instructed elements would have an influence on the users' pointing performance [17]. For this reason, we analyzed the benefit of our algorithm for each of the four elements.

Figure 6 shows that the accuracy of the pointing gesture was highest for the leftmost element, 87.57%. It decreased to 63.98% for the element in the rightmost position. With the gaze-added algorithm, the accuracy for the leftmost element is also at 85.57%. There is no improvement for this element. There is also a decrease in accuracy for elements further right, but it is less critical than for the gesture accuracy. The rightmost element was still selected with an accuracy of 71.43%. The biggest improvement was made for the third element (8.98%). However, Cochran's Q test determined that the influence of the position of the four elements were statistically significant in both cases, gesture accuracy ($\chi^2(3) = 28.27, p < .01$), and fused accuracy ($\chi^2(3) = 16.41, p < .01$).

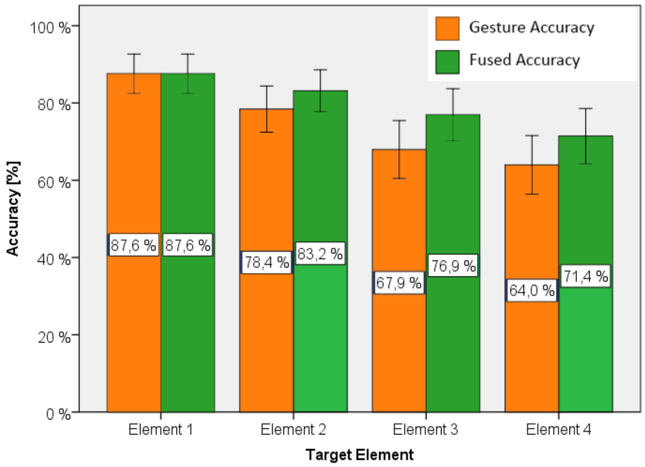

Figure 6. The position of the elements had a significant effect on gesture accurracy and fused accuracy. Error bars display the 95% confidence intervall of the mean.

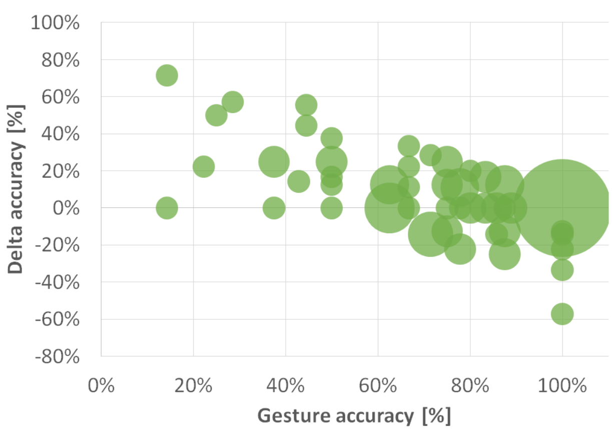

Figure 7. There is a negative correlation between the gesture accuracy and the improvement made by the algorithm (Δa).

Dependency on Gesture Accuracy

The results of the previous two sections showed that the benefit of our approach is greater for novice users, and for element positions that are further away from the users line of sight. In other words, we observed that the improvement of the gaze-added approach is greater when the accuracy of the pointing gestures decreases. We found a highly significant negative correlation between the gesture accuracy and Δa ($r_s = -.53, p < .01$). Figure 7 shows this relationship. For participants with lower gesture accuracy (below 70%), Δa is throughout positive and led to major improvements. However, when the gesture accuracy is higher (above 70%), Δa is lower and also drops below zero. For those people, the additional gaze information had a negative effect on the fused accuracy.

DISCUSSION

In summary, our results show that our gaze-added algorithm led to an overall improvement of selection accuracy for pointing gestures while driving. Users did not know that their gaze behavior influences the selection algorithm. Therefore, we claim that the users' gaze behavior does not differ compared to a normal pointing gesture selection. The increased accuracy of the system might further lead to a general reduction of driver demands, since the correction of wrong selections would draw more of the users' attention. In this context, it has been shown that the incorporation of additional vehicle data enhances mid-air selections, which finally resulted in reduced driver demands [2].

The results also show that the benefit of our approach is depending on different factors. Inexperienced users, who had a lower gesture pointing accuracy, benefit more than experienced users that are more likely to have a better pointing performance. Moreover, we observed that pointing performance decreases for elements that are further away from the driver's line of sight. The average difference of pointing accuracy was 23.59% between the leftmost and the rightmost element. These results are in line with findings in previous work [5, 17]. The gaze-added selection reduced the decrease by improving the accuracy for elements on the right, so that the difference between leftmost and rightmost element was only 14.14%.

In the end, this comes down to one observation: the benefit of gaze-added pointing increases, when initial gesture accuracy decreases e.g. for inexperienced users, or for elements that are further away from the users' line of sight. The significant negative correlation between gesture accuracy and Δa statistically backs this observation. The gesture accuracy for pointing is greatly influenced by the difficulty to target the requested element, which Fitts describes as a function of the width and the distance of selectable target items. [8]. In this regard, our algorithm is likely to provide greater improvements of selection accuracy when the selectable elements are more difficult to target with pointing gestures, e.g. when they are smaller (which will be necessary, if developers want to put more than four items on the screen), or if the position of the display in future vehicle concepts moves further away from the user.

Limitations

The algorithm presented in this paper is based on a simple approach, that aimed to support gesture selections by integrating gaze information in a clear and comprehensible way. Our results indicate that this approach generally works and leads to a significant improvement of accuracy, but they also reveal downsides of the algorithm. A more elaborate fusion algorithm is needed, e.g. by integrating gaze information based on a probabilistic model. Similar approaches have presented when incorporating vehicle data to optimize pointing performance [2]. Furthermore, inaccuracies of the used gestures sensor might contribute to bad pointing performance. Gesture recognition technologies in current vehicles also suffer from recognition problems. Even more, additional gaze data could help to reduce drawbacks from gesture sensor noise. Although eye-tracking technology faces similar accuracy problems, especially in the automotive domain, our experiment indicates how the fusion of both sensors can result in an overall increased accuracy. We also point out that the results are limited to a relatively specific setup, namely four large elements on a wide screen. The driving simulator did not support movements, which is likely to degrade pointing gesture performance even more [3]. Therefore, further studies will be needed to investigate the generalizability of the results for a greater variety of tasks, different setups and more realistic driving conditions.

CONCLUSION

We presented a user experiment that examined the benefits of integrating gaze input to improve the accuracy of pointing gestures while driving. In comparison to earlier research that combined these two modalities, we integrated gaze input in a passive way. Participants did not consciously use gaze input, but we only took naturally occurring eye-gaze data into account. We presented a simple algorithm how to fuse both modalities. The results of our experiment showed that the benefit of this approach is depending on the drivers' initial gesture accuracy, which is influenced by various factors, such as the experience with gesture interaction, or the size and position of target elements. The benefit of the presented approach grew when the difficulty of accurate gesture pointing increased. This led to major improvements for novice users and for those elements that are more difficult to select. On the other hand there was also a lack of support, or even a decline of accuracy, for those people that made very accurate pointing gestures. Despite some limitations, we argue that the gaze-added pointing approach can lead to an increase of selection accuracy, without posing additional demands on the driver. Based on these first promising results, our future work will focus on the development of a more elaborate fusion algorithm to better assess the drivers' focus of attention.

REFERENCES

1. Bashar I. Ahmad and Patrick Langdon. 2018. How Does Eye-Gaze Relate to Gesture Movement in an Automotive Pointing Task? *Advances in Intelligent Systems and Computing* 597, August 2017 (2018), 423–434. DOI: http://dx.doi.org/10.1007/978-3-319-60441-1
2. Bashar I. Ahmad, Patrick M. Langdon, Simon J. Godsill, Richard Donkor, and Rebecca Wilde. 2016. You Do Not Have to Touch to Select : A Study on Predictive In-car Touchscreen with Mid-air Selection. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 113–120. DOI: http://dx.doi.org/10.1145/3003715.3005461
3. Bashar I. Ahmad, Patrick M. Langdon, Simon J. Godsill, Robert Hardy, Lee Skrypchuk, and Richard Donkor. 2015. Touchscreen usability and input performance in vehicles under different road conditions. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2015), 47–54. DOI: http://dx.doi.org/10.1145/2799250.2799284
4. Bashar I. Ahmad, James Kevin Murphy, Simon Godsill, Patrick M. Langdon, and Robery Hardy. 2017. Intelligent Interactive Displays in Vehicles with Intent Prediction: A Bayesian framework. *IEEE Signal Processing Magazine* 34, 2 (2017), 82–94. DOI: http://dx.doi.org/10.1109/MSP.2016.2638699
5. B Biguer, M Jeannerod, and C Prablanc. 1982. The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale* 46, 2 (1982), 301–304. DOI: http://dx.doi.org/10.1007/BF00237188
6. Daniel Brand, Alexander Meschtscherjakov, and Kevin Büchele. 2016. Pointing at the HUD : Gesture Interaction Using a Leap Motion. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. DOI: http://dx.doi.org/10.1145/3004323.3004343
7. Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM Press, New York, New York, USA, 131–138. DOI: http://dx.doi.org/10.1145/2818346.2820752
8. Paul M. Fitts and James R. Peterson. 1964. Information capacity of discrete motor responses. *Journal of Experimental Psychology* 67, 2 (1964), 103–112. DOI: http://dx.doi.org/10.1037/h0045689
9. Dagmar Kern, Angela Mahr, Sandro Castronovo, Albrecht Schmidt, and Christian Müller. 2010. Making use of drivers' glances onto the screen for explicit gaze-based interaction. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, 110–116. DOI: http://dx.doi.org/10.1145/1969773.1969792
10. Sven Mayer, Katrin Wolf, Stefan Schneegass, and Niels Henze. 2015. Modeling Distant Pointing for Compensating Systematic Displacements. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* 1, 3606 (2015), 4165–4168. DOI: http://dx.doi.org/10.1145/2702123.2702332
11. David McNeill. 1992. Guide to Gesture Classification, Transcription and Distribution. (1992). DOI: http://dx.doi.org/10.2307/1576015
12. S F Neggers and H Bekkering. 2001. Gaze anchoring to a pointing target is present during the entire pointing movement and is driven by a non-visual signal. *Journal of Neurophysiology* 86, 2 (2001), 961–970.
13. R Neßelrath, Mohammad Mehdi Moniri, and Michael Feld. 2016. Combining Speech, Gaze, and Micro-gestures for the Multimodal Control of In-Car Functions. In *International Conference on Intelligent Environments*. 190–193. DOI: http://dx.doi.org/10.1109/IE.2016.42
14. Sharon Oviatt, Phil Cohen, Lizhong Wu, Bernhard Suhm, Josh Bers, Terry Winograd, and James Landay. 2009. Designing the User Interface for Multimodal Speech and Pen- Based Gesture Applications : State-of-the-Art Systems and Future Research Directions. *Human-Computer Interaction* 15, May 2014 (2009), 37–41. DOI: http://dx.doi.org/10.1207/S15327051HCI1504

15. Carl A. Pickering, Keith J. Burnham, and Michael J. Richardson. 2007. A research study of hand gesture recognition technologies and applications for human vehicle interaction. *Institution of Engineering and Technology Conference on Automotive Electronics* (2007), 1–15.

16. Katrin Plaumann, Matthias Weing, Christian Winkler, Michael Müller, and Enrico Rukzio. 2017. Towards accurate cursorless pointing: the effects of ocular dominance and handedness. *Personal and Ubiquitous Computing* (2017), 1–14. DOI: http://dx.doi.org/10.1007/s00779-017-1100-7

17. Florian Roider and Konstantin Raab. 2018. Implementation and Evaluation of Peripheral Light Feedback for Mid-Air Gesture Interaction in the Car.. In *International Conference on Intelligent Environments. Accepted for publication.*

18. Dario D. Salvucci and John R. Anderson. 2000. Intelligent gaze-added interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM Press, New York, New York, USA, 273–280. DOI:http://dx.doi.org/10.1145/332040.332444

19. Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 246–253. DOI: http://dx.doi.org/10.1145/302979.303053

20. Yanxia Zhang, Sophie Stellmach, and Abigail Sellen. 2015. The Costs and Benefits of Combining Gaze and Hand Gestures for Remote Interaction. (2015). DOI: http://dx.doi.org/10.1007/978-3-319-22698-9