User Preference for Smart Glass Interaction

Florian Bemmann

Abstract— Smart glasses are wearable devices providing the user always with information, using augmented reality techniques. In contrast to other devices such as smartphones they can be used without hiding the scene the user is in, so that it would be possible to use smart glasses in nearly every situation. Especially for on-the-go and working situations where smartphones can't be used, smart glasses are appropriate. To fully exploit these possibilities, new interaction concepts are required. This paper's aim is to first provide an overview of possible interaction concepts for smart glasses, independent of their technical feasibility of the currently available smart glass devices. Improving current devices is still required and ongoing, so currently impossible interaction concepts could become integrated in next versions if they turn out as providing a great user experience. I will evaluate which concepts might be preferred by users regarding (social) acceptance and performance. In the paper's second part I will for each gesture-based concept propose a use case suitable to its methods. Therefore my paper is based on existing studies examining acceptance and performance of interaction concepts on head-worn displays, such as smart glasses and augmented reality devices.

Index Terms—Smart glasses, Head-worn displays, HWD, interaction, input techniques, body interaction, mobile interfaces, Wearable, Augmented Reality

1 Introduction

After smartphones have revolutionized most people's everyday life within the last 10 years, the fast developing market of mobile computing devices offers more and more things. While tablets and smart watches are similar unappropriate on-the-go as smartphones, smart glasses are a completely different concept. They integrate in the user's life different, what could offer some new use cases. To gain the most benefit, other interaction concepts are required. In this paper I present some possible interaciton concepts for smartglasses and evaluate how they are preferred among the users. Promising the best user experience, I will focus on gesture based concepts.

2 Classification of interaction concepts for smart glasses

There exist several alternatives for structuring the possible interaction concepts. One is distinguishing the concepts into: free form and others. The former is defined as not requiring any extra device other than the smart glass to be performed and detected. Out of this group can further be selected a group of gesture based concepts, which I will focus on in the second part of this paper. For the first part, considering all possible interaction concepts for smart glasses, I will divide concepts into the groups touch, non-touch and handheld [5].

- handheld: interactions with any device that has to be held in hands, e.g. smartphone, controller, joystick
- touch: tapping and gesturing on body surfaces or wearable devices, providing tactile feedback. In the following are mentioned the target areas face, handpalm, wearable devices, the smart glass itself and at least other body parts
- non-touch: other movements or gestures. Mainly gestures performed with hands, also voice recognition, eye tracking, wink detection

3 Interaction concept's preference among users

This section I based on a user-elicitation study [5] where users was shown a effect of a game task and they were asked to perform a input action of their choice to cause that effect. Based on the percentages of which actions the user had chosen and a rating and interview afterwards, I determined which interaction concepts are the most preferred in each group.

3.1 Touch inputs

The most preferred touch input is using a finger to perform a gesture on the hand palm (chosen by 50% of the study participants [5]). Its similarity to touchscreens and trackpads leads users to the same input actions as on both aforementioned. Other on-body actions are finger, leg, handback and forearm. Interaction with the face had a quite low portion in this study (1%), but examining another sudy by Bertarini I would nevertheless recommend hand-to-face input. It promises a good level of acceptance and low intrusiveness [1]. Touching on the smart glass itself reached a 2% portion only in the study of Tung et al., even though it is one of the two primary input methods of Google Glass. As mentioned for hand-to-face input I would rate touching on the HWD a bit better as well. Especially its social acceptance is good (better than on face) [1] which is not a consequence of appearance, but of hygienic issues and meaning of face gestures in other ethnic groups [1]. On the other hand the performance on-device is lower than on-face, due to its small touching area [1]. A common wearable, the smart watch, was preferred by only 5% [5]. Interestingly 12% preferred a ring [5], a rather uncommon wearable. Another interesting concept is a digital belt, promising a good performance. Its quick and easy reachability was seen as benefit by the users- The social acceptance on the belt depends on the interaction length. For short interactions users did not feel very uncomfortable using all areas around the belt. When performing longer tasks, areas other than the front pockets were perceived as less suitable [3]. Although there aren't user preference scores comparing the belt with the other input concepts, belt is a promising one.

3.2 Non-touch inputs

In-air gestures are the by far most preferred non-touch input methods. 89% of the non-touch actions chosen were in-air gestures [5]. In-air gesture concepts, I will focus on in a later section. The methods eye tracking, wink detection and voice command are less preferred by users [5]. Even though voice command is one of both Google Glass' primary input methods, it reached only a 2% portion [5]. Anyway I would regard voice command as a good input method because its very intuitive. Its low score's reason might be a low social acceptance in public contexts, where the study was conducted in. Overall non-touch interaction was rated a little bit better than touch concepts [5].

3.3 Inputs using handheld devices

Handheld devices should only be a compromise solution. Their preference score was the lowest compared to the groups touch and non-touch

- *Florian Bemmann is studying Media Informatics at the University of Munich, Germany, E-mail: Florian.Bemmann@campus.lmu.de*
- *This research paper was written for the Media Informatics Proseminar, 2015.*

inputs [5], because users don't like that the device is not always available, it has to be taken out of the pocket first [5]. The worst fact in my opinion is that the interaction is not hands-free anymore, what destroys a main advantage of head-worn displays.

4 USE CASES FOR GESTURE BASED CONCEPTS

To assure a great user experience [1] I will now focus on gesture-based interaction. To evaluate whether a interaction concept is suitable to an operation I will in the following regard the concept's performance (performing time and the user exertion) and (user and social) acceptance. To find operations suiting to a task to be performed, I first separate into action and navigation tasks [4]. A action task can usually be performed by one action (e.g. answer a phone call, pause music player), whereas a navigation task can be more complex like navigating through a menu oder moving an object, e.g. a web browsers viewport.

4.1 On-body interaction

A factor for whether an on-body interaction is suitable is the area it is performed on. An area attracting attention when touching it or where touching is human unnatural has a low social acceptance [4]. The second important factor is the actions intrusiveness. Body movements which are to intrusive will not be accepted by users [4]. Aside from these limitations, on-body interaction offers lots of possibilities like coupling with on-body projection, and has the advantage of giving feedback through the human skins proprioception [4].

4.1.1 Hand-to-face

Hand-to-face input has an overall good performance. The most preferred areas for hand-to-face actions are cheek and forehead. Due to their large area users think they are the best parts of the face, especially the cheek which is perceived as a touchpad [4]. Performing actions on the cheek turned out as significantly faster and less exerting than the same action on the forehead and on the HWDs temple (chosen as direct alternative to hand-to-face input) [4] (Figure 11). The social acceptance in general is good as well, face contact is something natural [4]. Nevertheless the social acceptance for hand-to-face interaction is worse than for HWD interaction, escpecially in public context, but still on a good level and most people don't mind using the face. Some users show lower acceptance because of issues with facial cosmetics and dirt on the hands [4]. Users preferred hand-to-face for navigation tasks more than for action tasks. The performance is good for the typical navigation tasks panning and zooming due to the face' large areas [4]. Only for the navigation task "panning" the performance on the HWDs temple (oversized) is slightly better [4]. Moreover because of the HWDs higher acceptance, panning tasks should better be done on the HWD (provided that the HWD has an oversized temple). Coming to a conclusion I would recommend using the cheek for zooming tasks. The best suitable technique might be a linear zooming move. The alternative cyclo has low social acceptance because it could be perceived as the "you are crazy" gesture [4].

4.1.2 Palm based imaginary interfaces

Touching the palm is the users favorite touch interaction approach [5]. As reasons users mentioned that it is less intrusive, because it requires the least physical movement moving the right hand to the left hand palm [5]. Seaming similar to a smartphone touch display, the palm was often used as proxy touch-screen or trackpad. The palm offers haptical feedback both through finger and handpalm which helps navigating to the target, whereas a touchscreen can guide the user by e.g. drawing a grid and offers feedback only through the finger. As expected the touchscreen is of advantage, except when blindfolded. When blindfolded navigating on the palm is much faster, as an experiment conducted by Bertarini's shows [1] (figure 4). To find out whether the active (finger) or passive (palm) sense is most relevant, another experiment compared performance of palm, fake palm, and palm with finger cover. It came to the result that the passive tactile sense produces the most tactile cues [1] (figure 5). Summing up it can be said that using the palm has much better performance than using a real touchscreen when the user is blindfolded, what makes it suitable for on-the-go use-cases and impaired users. Because of the low preference score of handheld-devices mentioned in chapter "comparison among categories", the palm might be the better solution in not-blindfolded use cases as well.

Most suitable to be performed on the palm might be moving or drawing tasks using the palm's large surface [5]. E.g. moving an object to a specific position or just left and right; or drawing a path [5] (figure 7). For action tasks which are quite simpler the palm is suitable too, according to a user preference study. Nontheless, if the palm is still used for sophisticated tasks, I think it makes more sense to perform the action tasks on other surfaces to prevent occluding the palm with various different action types. Other input methods were preferred for action tasks as well [5].

4.1.3 In-air gestures

Due to the least attracted attention users prefer gestures performed in front of the chest. Also the exertion moving the hands to the chest is low. The second most chosen gestures are in front of the face, thereafter comes the area in front of the belly [5] (figure 9). The main reason for this preference order might be the social acceptance, which isn't as high when performing gestures in front of the face or the belly because it could look weird. Theoretically I can imagine in-air gestures for lots of tasks, but I suppose assigning navigation and selection in menus to in-air gestures. No other concept has shown suitable for this by now, and in a study Datcu et al. approved this in connection with a Augmented Reality system. The authors examined performance and users appreciation with a gesture interaction system used for navigating to a menu item (at a maximum menu depth of 4 levels) and came to the conclusion that spatial interaction is appropriate for AR [2]. Users were able to adapt to gesture interaction fast and only 20% did feel insecure, discouraged, irritated, stressed or annoyed while performing the menu task. [2]

4.1.4 Hand-to-body input: other body parts

Minus the so far considered body areas there are the areas finger, leg, handback, forearm and ring left. These areas could be used for action tasks requiring just one tap, each task or group of similar tasks dispersed to another area, like users did in the study of Tung et al. [5]. The concrete surface usually is irrelevant. Large surfaces like the chest can be used for lower precision requirements, such as selecting a single option from 4. Performed by a tap on one of 4 areas of the chest, a good performance can be reached [5]. The touch-area depending performance and acceptance might behave similar to the results examined for non-touch inputs. Areas which are hard to reach (very low areas like lower leg / foot or high areas on the head) have low performance scores due to the effort moving a hand towards this area. The acceptance might be low as well because it looks weird touching these hard reachable areas.

5 CONCLUSION

This paper explored possible interaction concepts for smart glasses, regardless of current smart glass version's technical capabilities. The main factors for whether a action is suitable are its performance, which consists of performing time and the user's exertion, and the user acceptance, especially in a public social context. In-air gestures in front of the chest and imaginary interfaces on the hand-palm turned out as the most suitable concepts. They allow blindfolded on-the-go use cases and hand-free interaction, two big advantages of smart glasses against other devices. Both aren't too intrusive to the user and attract little attention when performing in a public context. Future work has to focus on user studies in more realistic use cases in a real environment and with a real application. In addition it should be examined how much effort is required of the user when learning how to use the smart glasses. I think that might be harder than learning how to deal with a smartphone because of the huge variety of possible inputs and the missing guidance that touchscreen and button interaction offer. User guidance and learning concepts should be constructed and proved.

References

[1] M. Bertarini. Smart glasses: Interaction, privacy and social implications. https://www.vs.inf.ethz.ch/edu/FS2014/UCS/reports/MaricaBertarini_SmartGlasses_report.pdf, 2014. visited 24.04.2015.

[2] D. Datcu and S. Lukosch. Free-hands interaction in augmented reality. In *Proceedings of the 1st Symposium on Spatial User Interaction*, SUI '13, pages 33–40, New York, NY, USA, 2013. ACM.

[3] D. Dobbelstein, P. Hock, and E. Rukzio. Belt: An unobtrusive touch input device for head-worn displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2135–2138, New York, NY, USA, 2015. ACM.

[4] M. Serrano, B. M. Ens, and P. P. Irani. Exploring the use of hand-to-face input for interacting with head-worn displays. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 3181–3190, New York, NY, USA, 2014. ACM.

[5] Y.-C. Tung, C.-Y. Hsu, H.-Y. Wang, S. Chyou, J.-W. Lin, P.-J. Wu, A. Valstar, and M. Y. Chen. User-defined game input for smart glasses in public space. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3327–3336, New York, NY, USA, 2015. ACM.