Around-Body Interaction: Interacting While on the Go

Florian Müller, Sebastian Günther, and Max Mühlhäuser
Technische Universität Darmstadt

Head-mounted displays (HMDs) have come a long way since Sutherland introduced the concept 50 years ago.[1] This brings us closer to the vision of seamlessly molding interactive digital information onto the physical world—be it with HMDs or a future replacement technology (which we will subsume under HMD below for simplicity). As bits and atoms will continue to merge, corresponding interaction will increasingly happen everywhere, hence often on the move. However, today's interaction techniques, ranging from touch input on the frame of the HMD (e.g., Google Glass) or accessories[2] to voice commands,[3] are not tailored to the specific requirements of on-the-go interactions, limiting their viability for such situations. Besides the missing emphasis on mobile interactions, we see that these interaction techniques largely ignore the skills and dexterity we show in our day-to-day interactions with the physical world: Throughout our lives, we have trained extensively to use our upper and lower limbs to interact with and manipulate the physical world around us.

In this article, we provide a concise overview on how the skills and dexterity of our upper and lower limbs, acquired and trained in interacting with the physical world, can be transferred to the interaction with HMD. We present the vision of *around-body interaction*, in which we use the space around our body, defined by the reach of our limbs, for fast, accurate, and enjoyable interactions. In the remainder of this article, we first introduce a design space for such *around-body interactions* and, second, present two examples of such interaction techniques, which use the degrees of freedom of the lower limbs for on-the-go interactions. Finally, we conclude with future research challenges.

DESIGN SPACE

The envisioned ubiquitous interaction with a computer system using our limbs presents our body with additional tasks, entailing additional

Digital Object Identifier 10.1109/MPRV.2020.2977850
Date of current version 14 May 2020.

burdens. This becomes a particular challenge when an interaction with digital information coincides with an interaction with the physical world: Any part of the body involved in the interaction with the HMD is no longer (or, at least, no longer entirely) available for interaction with the physical world. As a consequence, a single interaction technique cannot support all situations a user might encounter during the day without interfering with the interactions with the physical world. To overcome the limitations of individual interaction techniques, future interaction with HMDs will, therefore, require a set of several interaction techniques using different limbs and different modalities of these limbs. From such a set, the most suitable yet least interfering technique can then be selected based on the situation.

There exists already a variety of interaction techniques that leverage movements of our limbs as an input modality. To classify these interaction techniques and help to select appropriate interaction techniques for different situations, we present a design space for on-the-go interactions using our limbs. This design space can help to identify yet unexplored—and, thus, unsupported—interaction situations for future work.

Building on an in-depth analysis of the design space for on-the-go interactions using our limbs, we propose a classification based on the following characteristics.

Limbs used for Input: Interaction techniques can use 1) the upper limbs, 2) the lower limbs, or 3) both, upper and lower limbs.

Output Reference: Different tracking technologies of HMD allow different types of visualization that 1) appear at a fixed position in the user's field of view (*head-stabilized*), 2) move as the user moves (*body-stabilized*), or 3) are anchored in the real world (*world-stabilized*) and, thus, allow the user to move independently of the visualization.[4]

Mobility: Interaction takes place during 1) stationary periods, and 2) while the user is in motion.

Number of Users: By design, HMDs are radically private devices, tailored to 1) single-user operation. To support collaborative situations, additional 2) multiuser interaction techniques are necessary.

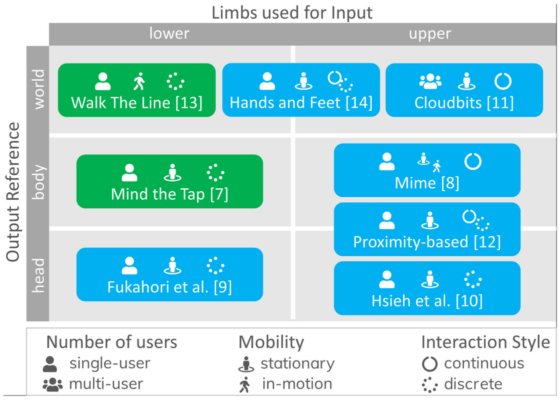

Figure 1. Design space for *around-body interactions* based on the 1) limbs used for input, 2) output stabilization, 3) mobility, 4) number of users, and 5) interaction style. This article discusses the two entries highlighted in green.

Interaction Style: For fast and direct interaction both, 1) continuous (e.g., gradual adjustment of a slider) and 2) discrete (e.g., traverse cascading menus) interaction techniques are necessary.

Figure 1 depicts a selection of *around-body interaction* techniques from the body of related works, classified in the presented design space. In the following section, we discuss two exemplary interaction techniques supporting on-the-go interaction using the lower limbs to support users during stationary and in-motion periods, as highlighted in Figure 1.

FOOT-BASED INTERACTION

While there is a long history of research on foot-based interactions in various areas from industrial machines[5] to smartphones,[6] foot-based interactions have seen little systematic evaluation for interaction with HMDs. At the same time, interaction using our feet can offer great advantages in mobile situations: While our hands and arms are often—in particular in a mobile context—busy interacting with the physical world, our feet are free to contribute to the interaction. Therefore, we explore two interaction techniques using the lower limbs, which allow hands-free interactions in highly mobile situations.

Mind the Tap: Interaction While Standing

Mind the Tap investigates foot-based interaction with an augmented interface that is fixed to

Figure 2. *Mind The Tap* allows hands-free interaction with information through foot taps on the floor in front of the user.

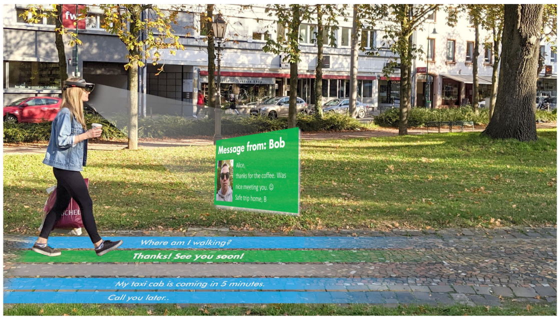

Figure 3. *Walk the Line* visualizes options as lanes on the floor. Users select options by shifting the walking path sideways.

the dominant foot of the users standing position and, thus, always available to the user. The interface consists of a semicircular interaction wheel, which is divided into a grid by several rows and columns. Each cell of the grid represents an option that users can select by tapping it with their foot (see Figure 2).

We investigated two different styles of interaction with this interface. In both variants, users operate the interface by foot taps, but the visualization of the interface is different:

Direct Interaction: The semicircular grid is visualized within leg reach on the floor. The user can interact with the system by looking to the ground and tapping the location, where the respective grid cell is visualized. Therefore, there is a direct connection between the location of input and output.

Indirect Interaction: The indirect interface moves the visualization to the air while keeping the input location on the floor. Despite the missing direct connection between the location of input and output, users still can interact with the system using foot taps. This ability is based on the sense of proprioception, which gives us an innate understanding of the relative position and orientation of our body parts without looking at them.

We compared the two styles of interaction in a controlled experiment and evaluated their accuracy and efficiency. The results show the suitability of both interaction styles for accurate ($> 99\%$ for two and four targets using both interface styles) and fast (1.0–1.2 s for selecting a highlighted target, depending on the condition) interactions through foot movements. The comparison of the interaction types revealed that direct interactions are better suited for short and fine-granular interactions, whereas indirect interactions are better suited for longer (e.g., multistep) interactions.

Walk the Line: Interaction While Walking

Besides the advantages of foot-taps as an input modality such as hands-free usage, there is one major limitation: The user has to stand in a fixed position in order to operate the interface. This limitation is based on the problem of differentiating between intentional movements aimed to contribute to the interaction and movements that occur naturally (e.g., while walking), cf., the *Midas Tap Problem*.[7] If, however, the vision of ubiquitous interaction with information in a digitally augmented physical world is to become a reality, a substantial part of the interaction with such devices will happen on the go. This highlights the necessity for truly mobile interaction techniques that not only support interaction while being at different places but also during the process of getting there, e.g., while walking.

We consider a system that displays multiple lanes parallel to the walking path of the user. Each lane represents an option the user can select. The lanes can be arranged on both sides of the users walking path (see Figure 3). To interact, users shift their path sideways until they walk on the desired option lane. The system highlights the lane the user is currently walking on by changing the visualization (e.g., the color) of the respective lane. This change affects the entire lane, which is also visible in front of the user. Therefore, users do not have to look to the ground to interact with the system but can keep their heads up. By walking along one of the lanes for a certain period (i.e., the selection time), the respective option can be selected, analogously to the concept of selection by dwell time in eye-gaze interaction.

In addition to the option lanes, we propose a nonactive *null lane* that covers the path directly in front of the user, which remains free. Therefore, if users continue walking straight ahead, the system does not interpret this as an interaction and does not trigger any actions.

We evaluated this interaction technique in a controlled experiment focusing on accuracy and efficiency. We found a selection time of 2/3 s with ~ 11cm wide lanes to be the best compromise between accuracy ($\sim 94\%$) and efficiency ($\sim 1.8\,\mathrm{s}$ for selecting a highlighted target).

CHALLENGES AND LIMITATIONS

As discussed above, interaction techniques can be rendered unsuitable by the context of use. When we carry things, we cannot use our hands for interaction. When we walk, we cannot use foot-tapping for interacting and, vice versa, when we stand, we cannot use a walking-based interaction technique. Besides the need to support all situations addressed in the design space using appropriate interaction techniques, this also raises the question of how to select the appropriate interaction technique for the current situation.

Simultaneous availability of multiple interaction techniques as the most straightforward solution leads to many problems similar to the Midas Touch Problem in gaze-based interaction: How can natural body movement be distinguished from intended interaction? The constant availability of all interaction techniques would, thus, lead to a large number of false actions by the system. Another simple solution could be the explicit selection of an interaction technique by the user. However, this selection must also be performed somehow—possibly by operating a switch, which again requires the use of the hands and is, therefore, not feasible in all situations.

Further work is needed in this area in order to arrive at a truly integrated system that can—based on the context and the situation of the user—make a decision on which interaction techniques should be available in the current situation.

CONCLUSION

We are closer than ever to the era of digitally augmented physical worlds—not just inside rooms, but increasingly *everywhere*. Interacting with the digital world on the go, while coping with the physical world, is a crucial challenge on that evolutionary path.

In this article, we presented a vision for a body-centric framework for the future *interaction* in the space around our body by harnessing the degrees of freedom offered by our bodies for more natural, pleasant, and fun interactions.

REFERENCES

1. I. Sutherland, “A head-mounted three dimensional display,” in *Proc. Fall Joint Comput. Conf., Part I*, 1968, p. 757. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1476589.1476686
2. D. Ashbrook, P. Baudisch, and S. White, “Nenya: Subtle and eyes-free mobile input with a magnetically-tracked finger ring,” in *Proc. SIGCHI Annu. Conf. Human Factors Comput. Syst.*, May 2011, p. 2043. [Online]. Available: http://dl.acm.org/citation.cfm?id=1978942.1979238
3. Y. He *et al.*, “Streaming end-to-end speech recognition for mobile devices,” Nov. 2018. [Online]. Available: http://arxiv.org/abs/1811.06621
4. M. Billinghurst, J. Bowskill, N. Dyer, and J. Morphett, “Spatial information displays on a wearable computer,” *IEEE Comput. Graph. Appl.*, vol. 18, no. 6, pp. 24–31, Nov./Dec. 1998. [Online]. Available: http://ieeexplore.ieee.org/document/734976/
5. K. H. E. Kroemer, “Foot operation of controls,” *Ergonomics*, vol. 14, no. 3, pp. 333–361, May 1971. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00140137108931255
6. M. Fan, Y. Ding, F. Shen, Y. You, and Z. Yu, “An empirical study of foot gestures for hands-occupied mobile interaction,” in *Proc. ACM Int. Symp. Wearable Comput.*, 2017, pp. 172–173. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3123021.3123043
7. F. Müller, M. Schmitz, J. Mcmanus, M. Mühlhäuser, S. Günther, and M. Funk, “Mind the Tap: Assessing foot-taps for interacting with head-mounted displays,” in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2019, pp. 1–13. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3290605.3300707

8. A. Colaco, A. Kirmani, H. S. Yang, N.-W. Gong, C. Schmandt, and V. K. Goyal, “Mime: Compact, low power 3D gesture sensing for interaction with head mounted displays,” in *Proc. 26th Annu. ACM Symp. User Interface Softw. Technol.*, 2013, pp. 227–236. doi: 10.1145/2501988.2502042.
9. K. Fukahori, D. Sakamoto, and T. Igarashi, “Exploring subtle foot plantar-based gestures with sock-placed pressure sensors,” in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, 2015, pp. 3019–3028. doi: 10.1145/2702123.2702308.
10. Y.-T. Hsieh, A. Jylha, V. Orso, L. Gamberini, and G. Jacucci, “Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses,” in *Proc.CHI Conf. Human Factors Comput. Syst.*, 2016, pp. 4203–4215. doi: 10.1145/2858036.2858436.
11. F. Muller, S. Gunther, A. H. Nejad, N. Dezfuli, M. Khalilbeigi, and M. Mühlhäauser, “Cloudbits: Supporting conversations through augmented zero-query search visualization,” in *Proc. 5th Symp. Spatial User Interact.*, vol. 17, 2017, pp. 30–38. doi: 10.1145/3131277.3132173.
12. F. Muller, M. Khalilbeigi, N. Dezfuli, A. Sahami Shirazi, S. Gunther, and M. Muhlhauser, “A study on proximity-based hand input for one-handed mobile interaction,” in *Proc. 3rd ACM Symp. Spatial User Interact.*, 2015, pp. 53–56. doi: 10.1145/2788940.2788955.
13. F. Muller, D. Schmitt, S. Gunther, M. Schmitz, M. Funk, and M. Muhlhauser, “Walk the line: Leveraging lateral shifts of the walking path as an input modality for head-mounted displays,” in *Proc. CHI Conf. Human Factors Comput. Syst.*, to be published, doi: doi:10.1145/3313831.3376852.
14. J. Schoning, F. Daiber, A. Kruger, and M. Rohs, “Using hands and feet to navigate and manipulate spatial data,” in *Proc. 27th Int. Conf. Extended Abstracts Human factors Comput. Syst.*, 2009, pp. 4663–4668. doi: 10.1145/1520340.1520717.

Florian Müller is a research assistant with the Technical University of Darmstadt, Darmstadt, Germany. Contact him at mueller@tk.tu-darmstadt.de.

Sebastian Günther is currently working toward the Ph.D. degree with the Technical University of Darmstadt, Darmstadt, Germany. Contact him at guenther@tk.tu-darmstadt.de.

Max Mühlhäuser is a full professor with the Technische Universität Darmstadt, Darmstadt, Germany. Contact him at max@informatik.tu-darmstadt.de.