ProxiWatch: Enhancing Smartwatch Interaction through Proximity-based Hand Input

Florian Müller
Telecooperation Lab
TU Darmstadt
Darmstadt, Germany
mueller@tk.tu-darmstadt.de

Sebastian Günther
Telecooperation Lab
TU Darmstadt
Darmstadt, Germany
guenther@tk.tu-darmstadt.de

Niloofar Dezfuli
Telecooperation Lab
TU Darmstadt
Darmstadt, Germany
dezfuli@tk.tu-darmstadt.de

Mohammadreza Khalilbeigi
Telecooperation Lab
TU Darmstadt
Darmstadt, Germany
khalilbeigi@tk.tu-darmstadt.de

Max Mühlhäuser
Telecooperation Lab
TU Darmstadt
Darmstadt, Germany
max@tk.tu-darmstadt.de

CHI'16, May 07-12, 2016, San Jose, CA, USA.
Copyright © 2016 ACM ISBN... 978-1-4503-4082-3/16/05$15.00.
http://dx.doi.org/10.1145/2851581.2892450

Abstract
Smartwatches allow ubiquitous and mobile interaction with digital contents. Because of the small screen sizes, traditional interaction techniques are often not applicable. In this work, we show how the degree of freedom offered by the elbow joint, i.e., flexion and extension, can be leveraged as an additional one-handed input modality for smartwatches. By moving the watch towards or away from the body, the user is able to provide input to the smartwatch without a second hand. We present the results of a controlled experiment focusing on the human capabilities for proximity-based interaction. Based on the results, we propose guidelines for designing proximity-based smartwatch interfaces and present ProxiWatch: a one-handed and proximity-based input modality for smartwatches alongside a prototypical implementation.

Author Keywords
Human Factors; Design; Measurement; Smartwatch.

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

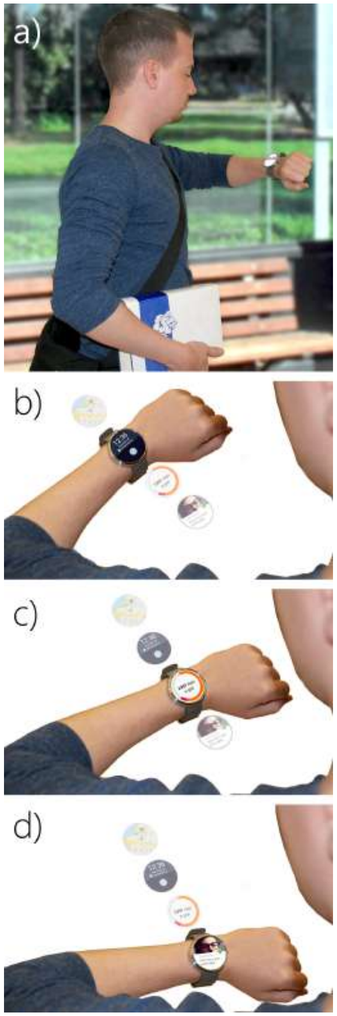

Figure 1: A launcher application as an example for a) one-handed proximity-based interaction: The user can open an application by raising his arm at different distances (b-d).

Introduction and Related Work

Highly capable smartwatches have become an emerging class of wearable devices that allow ubiquitous and mobile interaction with digital contents. Such devices usually consist of small multi-touch displays, bundled together with computing and sensing hardware from a mid-range smart phone, worn on the user's wrist. Therefore, smartwatches allow users to see, access and modify information right at their wrist, anytime and anywhere. The screen size of such devices is a tradeoff between mobility and interaction space. The evolution of wearable devices shows a trend towards small and elegant devices with small interaction space [9].

Therefore, traditional interaction techniques are not directly applicable to smartwatches. Current consumer devices (e.g. Apple Watch, Android Wear, Pebble Watch) mainly focus on 1) touch-based interfaces, 2) physical input controls on the frame of the device (e.g. a digital crown, buttons) and 3) off-device input modalities such as voice input. While practical and useful, those styles of interaction have various drawbacks in the context of smartwatches. Traditional touch-based interaction techniques suffer from small screen sizes as the user's interacting finger occludes a big part of the screen. Physical input controls such as a digital crown allow to interact with the content without occluding the screen. However, these approaches do not support for direct targeting and selecting of UI elements. In addition, touch interfaces as well as physical input controls require both hands of the user and, thus, may diminish the user experience in situations where the user is encumbered [8]. Voice input lacks direct manipulation and is difficult to use in noisy environments. Despite the advantages of using natural language, voice input is not suited for continuous interactions and is not socially appropriate in some situations [12].

In recent years, work is emerging that addresses these interaction challenges: Research presented on-device input modalities beyond traditional touch-based interfaces using a finger-mounted stylus [14] or tapping gestures [10] on the device. As another approach, off-device input modalities have been proposed that increase the interaction space of such devices through leveraging the space around the device: Using infrared, acoustic or magnetic sensors embedded into the frame, the device can track the location of the fingertip of the user's dominant hand. This allows users to perform off-screen (multi-)touch [1, 3] or air gestures [2, 5, 6] around the device or on surrounding surfaces [13] and, thus, without occluding the content on the screen. Despite the advantages, the presented approaches still require both hands of the user. As another approach, one-handed interfaces have been proposed that allow users to trigger a set of actions by performing gestures with their finger or hand [4, 6, 11, 15] of their non-dominant arm. While such interfaces can be operated with one hand, they do not support for continuous interactions.

We propose a novel interaction modality to operate smartwatches. We leverage the proximity of the hand relative to the user's body as an additional input dimension for one-handed interaction. We focus on the degree of freedom offered by the elbow joint, i.e., flexion (moving the hand towards to body) and extension (moving the hand away from the body). This allows users to interact with the smartwatch by moving the hand alongside their line of sight. For instance, users can trigger shortcuts just by raising their hand at a specific distance (c.f. Figure 1).

In this paper, we explore the concept of proximity-based interaction for smartwatches. We 1) contribute the results of a user study focusing on the human capabilities for a proximity-based input modality for smartwatches. Based on

Figure 2: The setup of the controlled experiment with two retro-reflective apparatuses mounted on the participant's head and wrist and the display showing the current task.

our lessons learned from the study, we 2) propose guidelines for the design of proximity-based interfaces for smartwatches. Last, we 3) present ProxiWatch: an interaction concept and prototypical implementation for one-handed and proximity-based input for smartwatches.

Controlled Experiment

We conducted a controlled experiment to investigate how efficiently and accurately users can raise the hand to a given target position in the space in front of them without any visual feedback. For this, we recruited 15 participants (5 female, 2 left-handed), aged between 19 and 30 years, using the University's mailing address. None of them had prior experience with smartwatches. Eight were users of regular watches in everyday life. We chose a within-subject design. No compensation was provided.

Design and Task

We defined a basic information space alongside the participants line of sight, evenly split into multiple layers and numbered in ascending order. The participants task was to raise their arm at a specified target layer without any visual feedback. We varied the number of layers as an independent variable with integer values from 2-8. We defined the maximum boundary of the interaction space with the participant's individual arm-length and the minimum boundary as the near point of the human's eye (not closer than 12.5cm to the user's face). However, we told the participants to use the space that is most comfortable for them as an interaction space.

Setup

We used an optical tracking system (OptiTrack) to measure the distance alongside the line of sight between the participant's wrist and eyes. As shown in Figure 2, participants wore a wristband at the typical watch position on their non-dominant hand and a pair of glasses, each augmented with a set of retro-reflective markers, during the experiment (c.f. Figure 2). We placed a display in front of the participants that showed the current task (layer subdivision and target layer within this subdivision). Additionally, we mounted a button within the reach of their dominant hand. For each trial, we measured the **distance** between wrist and eyes and logged it together with the **target layer** and the **condition** (total amount of layers).

Procedure

We used a repeated measure design with 7 levels for the number of layers (2, 3, . . . , 7 and 8). For each level, the participants targeted each layer with 5 repetitions. This resulted in a total of $(2 + 3 + 4 + 5 + 6 + 7 + 8) * 5 = 175$ trials per participant. The order of conditions as well as the order of targets within each condition was counterbalanced using a Balanced Latin Square design.

After welcoming the participants, we introduced them to the concept and the setup of the study. We mounted the two trackable apparatuses and calibrated the system to adapt it to the respective arm length. Before each condition, we told the participants about the layer subdivision for this task. Each trial was started by asking the participant to stand relaxed and lower his non-dominant arm. Once ready, the user pressed the button to start the trial. After that, the system showed the target layer as a number from 1 (nearest layer to the body) to the maximum layer of the current condition (2-8). Then, the participants raised their hand at the position where they imagined the respective layer with the backside of the hand towards their body. We told the participants to look at the center of the trackable apparatus on their wrist as if it was a watch. After raising their hand, participants had to confirm their action by pressing the nearby mounted button with their non-interacting hand.

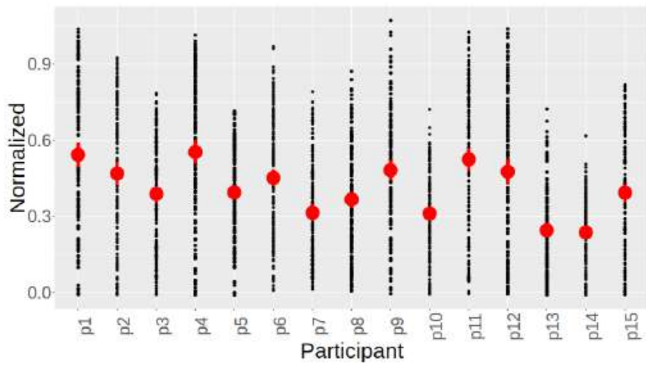

Figure 3: The interaction space of the participants. Black dots show the recorded distances normalized to the arm size, red dots show the center point of the interaction space. The used space differs significantly between participants.

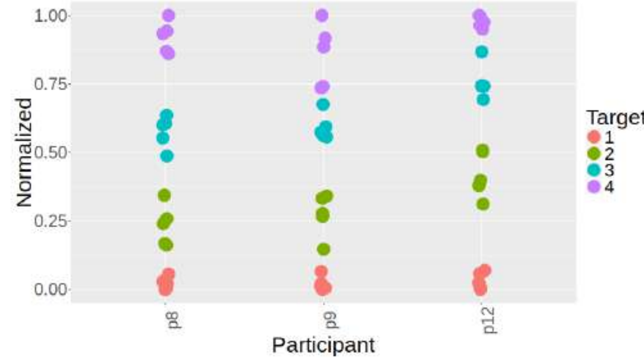

Figure 4: The distances for three exemplary participants for layer subdivision $n = 4$, scaled to their personal interaction space. A global model to classify points to a target layer is not possible. However, individual models per user are possible.

In the following, the system asked the user to take their hand down and enforced a 5 second break before starting the next trial.

We told the participants to focus on the accuracy instead of the speed. Participants did not receive any feedback during the study. After each condition, participants took a 30 second break. The complete experiment took about 30 minutes for each participant.

Results

We normalized the recorded data to the respective arm length into a scale from $0 \ldots 1$. In this scale, 0 refers to the near point of the human eye $(12.5\ cm)$ and 1 to the arm length of the user from shoulder to wrist. This maximum arm reach was measured in the calibration process with the same optical tracking system. We analyzed the data using repeated measures ANOVA and applied Bonferroni corrected pairwise t-tests for the post-hoc analysis.

Personal Interaction Space

We asked the participants to separate the space into layers in a way that is convenient for them. We found that the interaction space used by the participants as well as the center point of all interactions differs significantly ($F_{14,84} = 8.645;\ p < 0.001$) between participants (Min: P14, $0.0 - 0.51$ with one outlier, center point of interaction μ=0.24, Max: P4, $0.0 - 0.98$, center point of interaction μ=0.54). Figure 3 shows the interaction space of all participants normalized to their arm length. This personal interaction space for each participant remained constant for different layer subdivisions. For all further evaluations, we scaled the data based on the personal interaction space of each participant with 0 as the closest and 1 the most distant data point.

Directly Accessible Layers

We found that the size and the location of the center points of the layers differ significantly (Size: $F_{14,98} = 2.582;$ $p < 0.01$, Location: $F_{14,98} = 6.893;\ p < 0.001$) between participants even after scaling the data to the personal in-

Figure 5: Example application for discrete interaction: An application launcher that allows users to launch applications by raising their arm.

Figure 6: Example application for continuous interaction: A brightness control that lets the user modify the value by moving the arm alongside the line of sight.

teraction space of each participant. A general model that is able to map points from every user into the respective target layer is, therefore, not feasible for layer subdivisions > 2 as the layers overlap between participants. Within the data of individual participants however, a more finegrained differentiation between the layer zones is possible with no overlapping layers for a subdivision of at least 4 for all participants. Figure 4 shows the data points for three participants for a subdivision of four layers. The analysis further showed smaller layers for the outer regions (i.e. close to and far away from the body) compared to the inner regions (Inner: $\mu = 0.16,\ \sigma = 0.06$, Outer: $\mu = 0.12,\ \sigma = 0.07$). This is not influenced by the personal interaction spaces of the participants.

Implications

Respect the personal interaction space

Our experiment showed that every user has a personal convenient interaction space that is not generalizable over multiple users. Thus, a system should not force the user into a fixed set of layers that spans larger or smaller than the user's personal interaction space.

Provide bigger layers in the inner regions

The outer layers of the mental model of participants are smaller than the inner layers. Therefore, a system might use smaller layers on the outside and increase the size of the inner layers to support the user.

Use a personal model to achieve the best recognition rate

We found that the location and size of the layer in the mental model of the participants differed within their individual interaction spaces. Therefore, a general model over all users is not feasible and, thus, a personal model is necessary to achieve the best recognition rate for higher subdivisions.

ProxiWatch

The results discussed above together with prior work focusing on the human capabilities for continuous interaction [7] showed that proximity-based interaction with smartwatches is a viable concept. Based on the findings, we developed ProxiWatch: A one-handed proximity-based hand input modality for smartwatches along with two main interaction techniques. In addition, we implemented two example applications to show the usefulness for varying scenarios.

Discrete Interaction

Discrete interaction is based on the directly accessible layers found in the study. Within this set of layers, each layer can be mapped to a functionality in terms of a shortcut. Such a system can be used to directly select various items in a fixed set of objects (e.g. launcher, media player controls) by raising the arm within the boundaries of one of the target zones in front of the body. As a result of the presented experiment, we found that 4 layers are easily distinguishable for users. We showcase this interaction technique using an application launcher for the smartwatch. This launcher allows users to open up a favorite application by raising the arm (c.f. Figure 5). This technique supports quick and immediate interaction with a smartwatch.

Continuous Interaction

Continuous interaction allows the user to adjust a continuous value by moving the hand within the bounds of the personal interaction space. Through visual feedback on the smartwatch, the user is able to quickly and efficiently adjust a value (e.g. slider) just by moving the arm. We also propose this interaction technique as an extension to discrete interaction for lists with more elements. We illustrate this technique with a brightness control for the smartwatch that allows users to adjust the value by moving the hand. The user can select a value by lowering the hand.

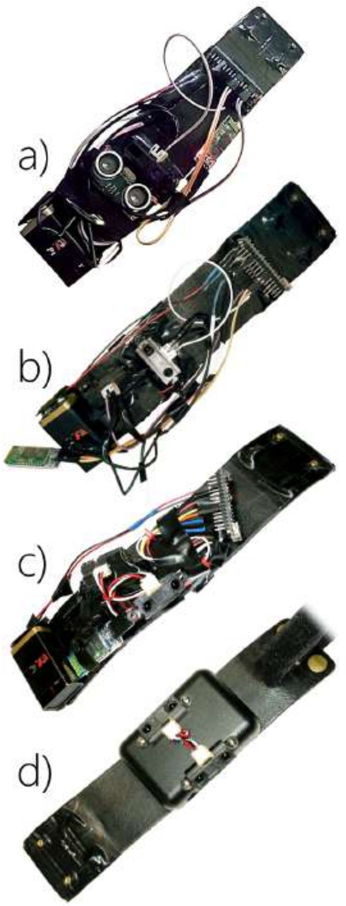

Figure 7: Iterative design of the prototype: a) First version using a generic ultrasonic and infrared sensor b) second version using two infrared sensors c) third version with tilted infrared sensor alignment d) current version with boxed design

Both presented interaction techniques can be combined (i.e. discrete interaction to launch an application, continuous interaction to adjust a value). Furthermore, traditional touch-based input is possible on each layer.

Technical Overview
We built a stand-alone, wireless prototype to enable proximity-based interaction concepts on a consumer smartwatch (Motorola Moto 360) in an iterative process (c.f. Figure 7).We used a battery powered Arduino Nano with two infrared distance sensors (Sharp GP2Y0A21YK0F). When holding the hand in a rotation that allows to read the display of the watch, both sensors are directed towards the body of the user. However, the angle of both sensors towards the body is slightly different (c.f. Figure 7 f). The system reports the value of the first sensor as long as this sensor has the body of the user within its field of view. Otherwise, the value of sensor two is reported (c.f. Figure 8). Depending on the body structure of the user, the handover point is found around 20cm distance. Because of the limited processing capabilities of the Arduino, we transmit the raw sensor data to a processing application on a mobile phone (Samsung Galaxy S4) via Bluetooth. The phone handles the incoming raw data and selects the appropriate sensor. In addition, we use a Kalman filter to reduce the statistical noise of the returning sensor values. After processing, we send the estimated distance to the smartwatch. To detect if a user has raised his arm, we use the acceleration sensor of the smartwatch. This allows to support the presented discrete interaction technique. Furthermore, the acceleration sensor is used to register a shake-wrist gesture which can be used for secondary actions (i.e. select item).

We compared the estimated distance values of our prototype to the real distance and found that our prototype robustly recognizes the distance to the user for the complete

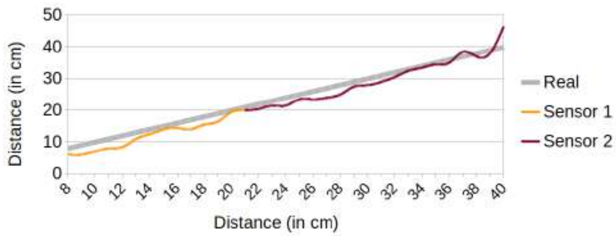

Figure 8: The estimated distance values reported by the two sensors compared to the real distance. The handover between both sensors is at ∼21cm.

interaction space ($\mu = 1.9cm,\ \sigma = 1.19cm$, c.f. Figure 8). Future work is needed to improve the quality of the measurements and to further miniaturize the prototype.

Conclusion and Future Work
In this paper, we explored how the proximity of the hand can be leveraged as an additional input dimension for smartwatches. We reported on the results of a controlled experiment focusing on the human capabilities. The results confirmed the applicability of our concept. Based on the results, we proposed guidelines and presented ProxiWatch: A set of concepts and prototype implementation of a proximity-based input enabled smartwatch. As a next step, we plan to investigate more deeply how proximity-based interaction can enhance real world smartwatch usage using the presented prototype. Furthermore, we want to explore further proximity dimensions of the hand beyond the degree of freedom provided by the elbow joint.

Acknowledgments
This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.B0101-15-1292, Development of Smart Space to promote the Immersive Screen Media Service).

References

[1] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight. In *Proceedings of the 21st annual ACM symposium on User interface software and technology - UIST '08*. ACM Press, New York, New York, USA, 201. DOI: http://dx.doi.org/10.1145/1449715.1449746

[2] Chris Harrison and Scott E. Hudson. 2009. Abracadabra. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology - UIST '09*. ACM Press, New York, New York, USA, 121. DOI: http://dx.doi.org/10.1145/1622176.1622199

[3] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, New York, New York, USA, 453. DOI: http://dx.doi.org/10.1145/1753326.1753394

[4] Frederic Kerber, Philipp Schardt, and Markus Löchtefeld. 2015. WristRotate. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia - MUM '15*. ACM Press, New York, New York, USA, 218–222. DOI: http://dx.doi.org/10.1145/2836041.2836063

[5] Jungsoo Kim, Jiasheng He, Kent Lyons, and Thad Starner. 2007. The Gesture Watch: A Wireless Contact-free Gesture based Wrist Interface. In *2007 11th IEEE International Symposium on Wearable Computers*. IEEE, 1–8. DOI: http://dx.doi.org/10.1109/ISWC.2007.4373770

[6] Jarrod Knibbe, Diego Martinez Plasencia, Christopher Bainbridge, Chee-Kin Chan, Jiawei Wu, Thomas Cable, Hassan Munir, and David Coyle. 2014. Extending interaction for smart watches. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*. ACM Press, New York, New York, USA, 1891–1896. DOI: http://dx.doi.org/10.1145/2559206.2581315

[7] Florian Müller, Mohammadreza Khalilbeigi, Niloofar Dezfuli, Alireza Sahami Shirazi, Sebastian Günther, and Max Mühlhäuser. 2015. A Study on Proximity-based Hand Input for One-handed Mobile Interaction 1 ¨. In *Proceedings of the 3rd ACM symposium on Spatial user interaction - SUI '15*. DOI: http://dx.doi.org/10.1145/2788940.2788955

[8] Alexander Ng, Stephen A. Brewster, and John H. Williamson. 2014. Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 1981–1990. DOI: http://dx.doi.org/10.1145/2556288.2557312

[9] Tao Ni and Patrick Baudisch. 2009. Disappearing mobile devices. In *Proceedings of UIST 2009*. ACM Press, New York, New York, USA, 101–110. DOI: http://dx.doi.org/10.1145/1622176.1622197

[10] Ian Oakley, DoYoung Lee, MD. Rasel Islam, and Augusto Esteves. 2015. Beats. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 1237–1246. DOI: http://dx.doi.org/10.1145/2702123.2702226

[11] J. Rekimoto. 2001. GestureWrist and GesturePad: unobtrusive wearable interaction\ndevices. *Proceedings Fifth International Symposium on Wearable Computers* (oct 2001), 21–27. DOI: http://dx.doi.org/10.1109/ISWC.2001.962092

[12] Thad E. Starner. 2002. The role of speech input in wearable computing. *IEEE Perversive Computing* 1, 3 (jul 2002), 89–93. DOI: http://dx.doi.org/10.1109/ISWC.2008.4911608

[13] Wouter Van Vlaenderen, Jens Brulmans, Jo Vermeulen, and Johannes Schöning. 2015. WatchMe. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. ACM Press, New York, New York, USA, 2091–2095. DOI:http://dx.doi.org/10.1145/2702613.2732789

[14] Haijun Xia, Tovi Grossman, and George Fitzmaurice. 2015. NanoStylus. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. ACM Press, New York, New York, USA, 447–456. DOI:http://dx.doi.org/10.1145/2807442.2807500

[15] Chao Xu, Parth H. Pathak, and Prasant Mohapatra. 2015. Finger-writing with Smartwatch. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications - HotMobile '15*. ACM Press, New York, New York, USA, 9–14. DOI:http://dx.doi.org/10.1145/2699343.2699350