PinchWatch: A Wearable Device for One-Handed Microinteractions

Christian Loclair, Sean Gustafson, Patrick Baudisch
Hasso Plattner Institute
Potsdam, Germany
{christian.loclair, sean.gustafson, patrick.baudisch}@hpi.uni-potsdam.de

ABSTRACT
Gesture-based wearable devices, such as *Gesture Pendant* or *Sixth Sense*, leave the user's hands empty. This allows users to abandon interaction with the device *instantly* when the physical environments requires it, e.g., if the user stumbles. In theory, this gives this class of devices the potential to be used during a mission critical "primary" task, such as driving or rock climbing.

In this paper, we explore this class of devices. In particular, we investigate how to *engage* in interaction quickly—because such scenarios typically allow users to take their eyes off the primary task only briefly (*microinteractions*). We present PinchWatch, a one-handed device with a wrist-worn display. Users invoke functions by *pinching* (i.e., thumb presses against finger or palm); they enter parameters by performing sliding or dialing motions with the thumb on the palm; or they move the hand as a whole (peephole interaction). All user input is tracked by a chest-worn camera observing the user's hand and wrist. Complementing this camera with depth sensing capability allows for tracking additional gestures, for use in sunlight, and for recognizing pinch gestures when the hand is not directly facing the camera.

Our design minimizes interference with the user's primary task by (1) building on large number of *single purpose* gestures, thereby avoiding modes and menus (2) by offering tactile feedback and eyes-free use for most interactions through pinching, and (3) requiring only a single hand, allowing the other hand to stay on the primary task at all times.

Categories and Subject Descriptors
H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General Terms
Design, Human Factors

Keywords
Wrist, watch, wearable, computer vision, peephole, mobile, spatial, gestures, microinteractions.

1. INTRODUCTION
When driving a car, playing sports, walking on a busy street, or rock climbing, users cannot take focus away from the physical world for more than an instant. If any interaction with a computer has to take place, it must therefore happen in very short bursts. Such short interactions (< 4sec), i.e., *microinteractions* [2], typically involve users checking transit schedules, traffic reports, weather advisories, or their email on their mobile or wearable device.

MobileHCI 2010 September 7-10, 2010, Lisboa, Portugal.
ACM 978-1-60558-835-3/10/09.

We investigate what types of devices and interaction techniques might be best-suited for these types of microinteractions. We find that wearable devices that leave the user's hands empty have potential, such as *Sixth Sense* [18] or *Gesture Pendant* [21]. Unlike traditional mobile devices that must be fetched and put away, these gesture-based wearable devices keep the users' hands free, and thus allow users to *instantly* return to the primary task at hand, such as driving or staying on the climbing wall.

In addition to interruptability, use with mission-critical tasks also requires minimizing the time that the device takes focus away from the primary task. Whenever possible, the device should allow for eyes-free use; it should use fast, single-purpose functions rather than menu navigation; and it should engage as few limbs as necessary; for rock-climbing, for example, this means one-hand use.

In this paper, we discuss the design space of such a device and present our preliminary implementation.

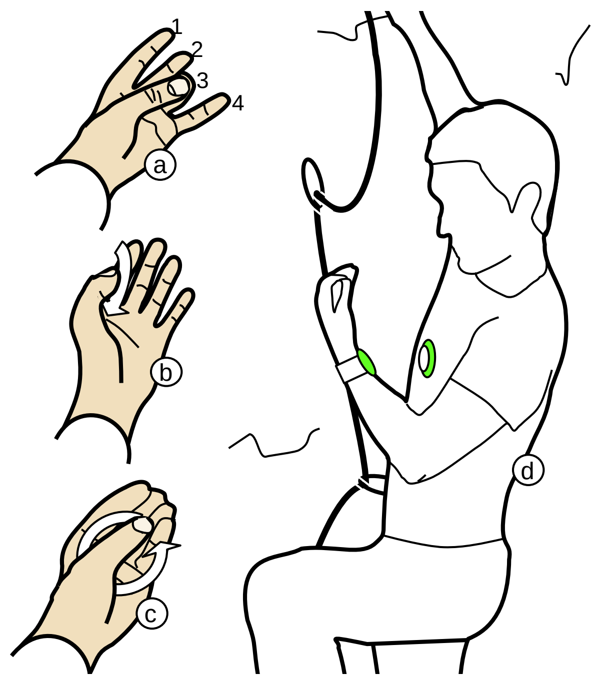

Figure 1. The PinchWatch device supports (a) discrete input (pinching different fingers and finger segments), (b) sliding along the fingers and (c) dialing on the palm. All interactions involve users touching their palm or fingers with their thumb, which offers physical feedback and thus allows for eyes-free use. (d) Because the user's hands are empty, the user is, at anytime, able to immediately abandon interaction and use both hands in the physical world.

2. PINCHWATCH WALKTHROUGH

PinchWatch (shown in Figure 1) consists of a display (here a wrist-worn LCD) and a chest-worn camera that tracks the hand wearing the display, recognizes its motion and gestures.

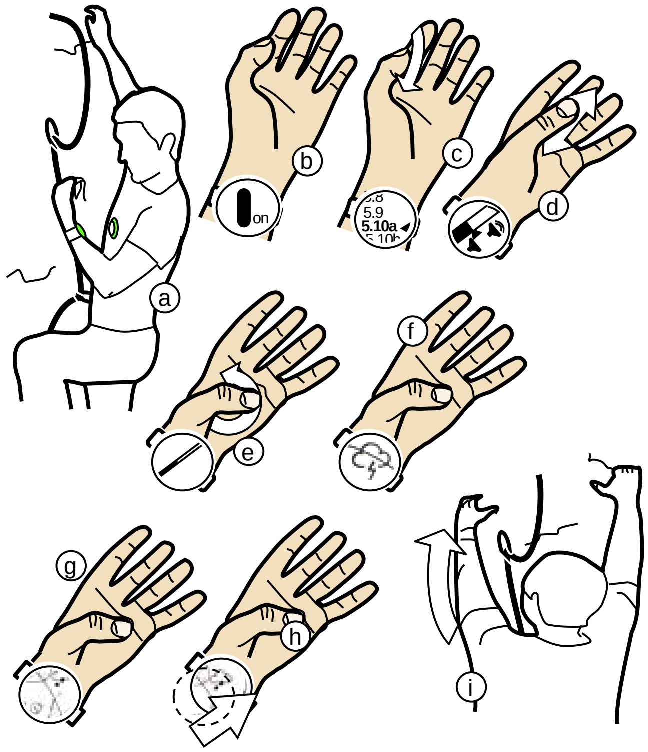

Figure 2. Rock climbing with our proposed device. Joe uses the device to record annotation of his climb, adjust his music and check the weather.

Figure 2 illustrates one possible scenario. Joe is rock climbing with his friends and wearing the PinchWatch device (running his climbing software). (a) As Joe brings the display towards his chest, the camera sees the display (b) causing it to turn on. Joe pinches his index finger and the device recognizes this as the "annotate" command. (c) By sliding his thumb down his finger to second crease, Joe rates the difficulty of his last maneuver as "5.10a". He does not look at his hand while doing this—all the feedback he needs is contained in the tactile features of his hand. (d) Next he adjusts the volume of his music player by sliding along his middle finger and (e) scrubs through the current song by rotating his thumb on his palm. As he reaches back for a climbing hold, the display piece leaves the camera's field of view, causing the camera to stop looking for gestures and the display to turn off.

Further up the wall Joe hears thunder rumbling in the distance. He presses his thumb into his palm (the gesture used to invoke the weather map) and moves his arm into view. (f) When the display turns on it is already showing the weather map over his current location. (g-h) By moving his hand over an imaginary map, Joe pans through the weather map (*peephole interaction* [24]). (i) At this moment, Joe's foot breaks free. Reflex-like, he grabs a hold of the wall with his device hand. His hand has been empty the entire time so this is not a problem. He regains his footing (and his composure), takes his hand off the wall and continues to study the weather conditions.

3. BENEFITS AND CONTRIBUTION

In this scenario, PinchWatch offered the following benefits:

(1) The user's hands remain empty at all times—unlike a hand-held device. This allows users to return to their primary task immediately and at any time, here when Joe's foot slips.

(2) A large part of the interaction is performed eyes-free, here by performing pinch gestures when invoking the display. The key is that all gestures are derived from a pinch, which causes both thumb and the touched part of the hand to *feel* the interaction. This allows users to prepare the interaction before starting it, minimizing the duration of visual interaction and maximizing the time users can focus on the physical world. In addition, pinch-style gestures deliver an unambiguous beginning and end of the gesture. The function is executed immediately upon unpinching the fingers—there is no need to wait for a dwell time to expire.

(3) The surface of the hand contains many prominent tactile features (creases and bones). Each of these could be assigned discrete values (such as in the annotation example above) or as landmarks to orient to when, for example, the finger surface is used as a volume slider.

(4) Five specialized interaction modes allow users to control functions using the best modality: button (pinch [23]), slider (slide finger along thumb, see Figure 3 for the *extended* slider.), dial (thumb spinning on palm), 2D pointing (thumb on palm), and panning (moving hand, peephole interaction [24]).

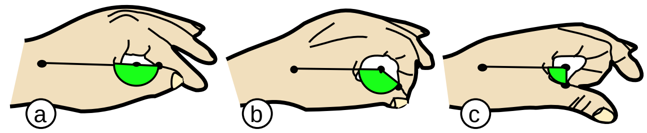

Figure 3. Extended slider: an extra degree of freedom is added to the pinch by sensing the angle from the center of the pinch hole to the pinch contact. This is an *extended* slider because the user can slide along their thumb and/or index finger in one fluid motion.

(5) Multiple single-purpose elements, such as a slider for each finger, allow each device function to be controlled via its own control element, similar to the design concept behind studio mixers and airplane cockpits. This eliminates the need for a menu, making the interaction faster and mode-less, thus less error-prone.

(6) By offering high expressiveness with a single hand, the other hand can remain in use for the primary task at all times, here holding on to the wall.

(7) By allowing users to place their hand close to the space that requires their attention in the physical world, the device limits the loss of attention.

Combined, these seven features allow for interactions in scenarios where other devices would be cumbersome or even dangerous to operate. While these features make PinchWatch particularly appropriate for crucial tasks, such as the one discussed above, they are also useful in day-to-day situations, such as when holding a drink or finger food at a party; holding onto the railing when riding bus or subway, etc.

4. DEVICE HARDWARE

Using a wrist-mounted display and a chest-mounted camera we are experimenting with different mechanisms for tracking the user's hand and finger gestures. Figure 4 shows the computer vision pipeline for an image captured with this setup.

Although somewhat promising, we found the approach based on a standard camera to be subject to three limitations:

(1) It is sensitive to the lighting situation. In order to allow the device to work in a broader range of lighting situations, others have proposed adding illumination to the device, e.g., a ring of IR LEDs (Gesture Pendant [21]). While we found that although this works well indoors, it fails outdoors (especially in direct sunlight) where mobile/wearable devices are quite often used.

(2) It cannot determine which finger is being pinched to the thumb. A pinch between the thumb and a finger can be robustly sensed (using the computer vision algorithm proposed by Wilson [23]) by identifying a small contour within the larger contour of the whole hand. We can add one degree of freedom to this gesture by sensing where the finger and thumb touch (Figure 4c,d), but obtaining more information than that (such as which finger is being pinched or where the thumb is touching the palm) is exceptionally difficult on a bare hand.

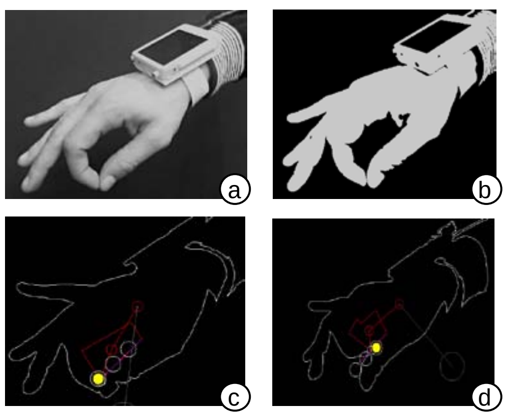

Figure 4. (a) Wrist display image taken by chest camera. From the (b) thresholded image we (c) determine the inner and outer contours and find the pinch location. (g) Detecting a pinch location further up the thumb.

(3) In order to recognize the finger pinching, the hand must be held at a specific orientation. If the hand is slightly rotated the algorithm does not work. To recognize a pinch the hand must be orientated so that camera can peer directly through the pinch hole (see Figure 4). The required posture depends on the location of the camera (Figure 5). Regardless, users must maintain a specific orientation which is difficult to do eyes-free.

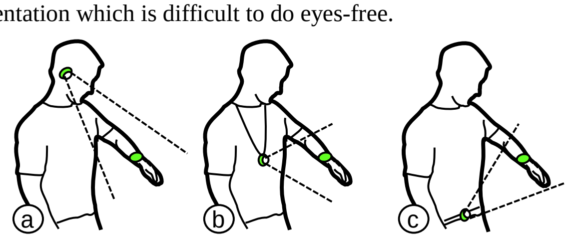

Figure 5. Possible locations for the camera: (a) ear; (b) chest; (c) belt.

To address these issues we are currently experimenting with a depth camera. We are using the *PMD[vision] CamCube 2.0* time-of-flight depth camera that offers 204×204 pixels at 25fps (see Kolb et al. [15] for an excellent survey of available technology).

The first results are encouraging. In particular, the depth camera shows promise in four areas:

(1) It reliably separates the foreground from the background, even in direct sunlight (Figure 6a,b).

(2) With a chest-mounted camera pointed at the palm of the hand and the wrist display on the inside of the wrist (as in all images in Figure 6), the arm can rest in a neutral position during the interaction.

(3) The depth information can be used to separate the parts of the hand that are in front of the palm. For instance, as shown in Figure 6c, by looking directly at the palm, bent finger tips can be seen in front the palm of the hand—something that is very difficult with a conventional camera. This allows us to detect pinches between the thumb and all fingers as well as thumb placement anywhere on the palm of the hand.

(4) We can reproduce the pinching gesture discussed above but now from more viewing angles—there is no longer the need to see through the pinch hole to create an inner contour (see Figure 6d). The user is freed from having to worry about the precise viewing angle of the camera—a considerable help for eyes-free interaction.

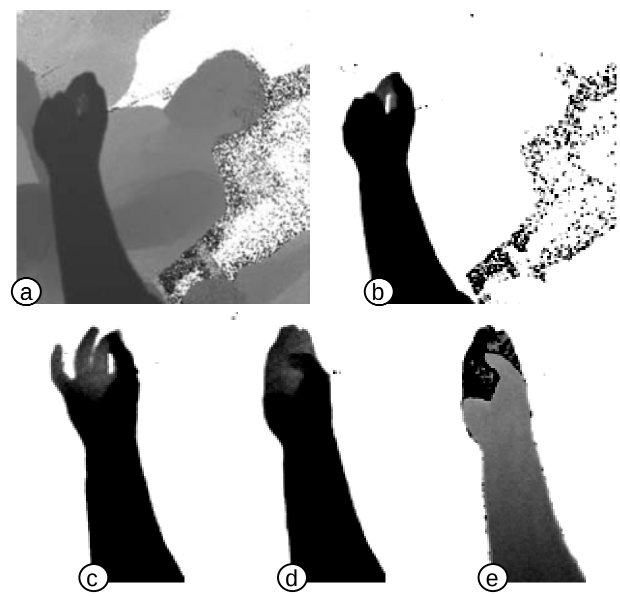

Figure 6. Depth camera images: (a) raw image with depth encoded as grayscale; (b) threshold applied to separate hand from background; (c) pinch on middle finger; (d) thumb on hand; (e) thumb on hand with extra processing to increase local contrast.

Depth cameras are not without their limitations. Depth resolution for our camera is limited (approx. 1cm in our tests) and subject to substantial noise. The size and cost of a suitable depth camera means they are impractical for deployment outside of a research lab. However, commercially available depth cameras are rapidly improving and it is reasonable to assume that some of these limitations will be overcome in the next couple years.

5. RELATED WORK

This work is influenced by hand gesture research in many areas. We find the hand gestures of classic systems, such as Charade [4], to be well suited for mobile interaction, particularly the pinch gestures used in menu techniques such as the TULIP menu [6] and rapMenu [19]. This is because pinching includes implicit tactile feedback when completing the gesture. Similarly, the Haptic Hand [14] uses the surface of the non-dominant hand as a tactile surface that can be interacted upon.

However, in mobile scenarios it is not appropriate to require the user to wear a specialized glove. To overcome this need, pure vision-based hand-gesture recognition methods have been developed that detect a bare hand. The Gesture Pendant [21] is one notable system, from which we borrow our chest mounted form factor. Wilson [23] provides a computationally lightweight method of detecting a pinch that we base our implementation on. The WristCam [22] senses simple finger gestures with a camera mounted on the user's wrist. The Sixth Sense [18] system can, with colored markers on the fingers, detect presses on buttons projected onto the non-dominant hand. Stereo [10] and depth cameras [7],[16] have also been used to sense hand gestures.

Other systems provide bare-hand gesture interaction in other ways. SideSight [8] and HoverFlow [17] detect movement around a handheld device with an array of proximity sensors. The Body Coupled FingerRing [9] uses accelerometers mounted to fingertips to allow (mostly) bare hand typing whereas Skinput [12] senses vibration on the skin to detect taps on various locations of the hand and forearm. GestureWrist [20] infers hand gestures indirectly by sensing changes in the shape of the wrist.

Wrist-based mobile devices have been a popular research platform. Ashbrook et al. [1] and Blaskó et al. [3] among others investigated direct input on the face of the watch. Baudisch and Chu presented the idea of a wrist-watch operated by touching the underside of the watch band [5]. Other systems, such as Gesture-Watch [13] and Abracabra [11] sense gestures performed by the hand near the wrist-worn device.

6. CONCLUSIONS

In this paper we presented PinchWatch, a one-handed modeless gestural wearable device designed specifically to support microinteractions.

Many situations, such as when rock climbing or driving a car, require a user to effortlessly switch attention between their mobile/wearable device and their primary task. The PinchWatch supports such a situation by leaving the users hands free at all times and by leveraging the implicit tactile sensation the thumb touching the palm to allow for expressive eyes-free gestures. Furthermore by assigning all functions of the device to be controlled by different parts of the hand, we create a completely modeless and menu-free interface where all functions can be immediately invoked. By forming the gesture first then bringing the hand into view of the camera users maintain the maximum attention on their primary task at hand.

7. ACKNOWLEDGMENTS

We thank Daniel Bierwirth and Lars Schneider for their helpful discussions and Thorsten Ringbeck from PMDTec for supporting us with a TOF camera.

8. REFERENCES

[1] Ashbrook, D., Lyons, K., and Starner, T. An investigation into round touchscreen wristwatch interaction. In *Proc. MobileHCI*, (2008), 311–314.

[2] Ashbrook, D. *Enabling Mobile Microinteractions*. PhD Thesis. (2010).

[3] Blaskó, G. and Feiner, S. An interaction system for watch computers using tactile guidance and bidirectional segmented strokes. In *Proc. ISWC*, (2004), 120–123.

[4] Baudel, T. and Beaudouin-Lafon, M. Charade: remote control of objects using free-hand gestures. *Communications of the ACM 36*, 7 (1993), 28–35.

[5] Baudisch, P. and Chu, G. Back-of-device interaction allows creating very small touch devices. In *Proc. CHI*, (2009), 1923–1932.

[6] Bowman, D.A. and Wingrave, C.A. Design and evaluation of menu systems for immersive virtual environments. In *Proc. VR*, (2001), 149–156.

[7] Breuer, P., Eckes, C., and Müller, S. Hand gesture recognition with a novel IR time-of-flight range camera–a pilot study. In *Proc. MIRAGE*, (2007), 247–260.

[8] Butler, A., Izadi, S., and Hodges, S. SideSight: multi-"touch" interaction around small devices. In *Proc. UIST*, (2008), 201–204.

[9] Fukumoto, M. and Tonomura, Y. "Body Coupled FingerRing": wireless wearable keyboard. In *Proc. CHI*, (1997), 147–154.

[10] de la Hamette, P. and Troster, G. Fingermouse-architecture of an asic-based mobile stereovision smart camera. In *Proc. ISWC*, (2006), 121–122.

[11] Harrison, C. and Hudson, S.E. Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices. In *Proc. UIST*, (2009), 121–124.

[12] Harrison, C., Tan, D., and Morris, D. Skinput: appropriating the body as an input surface. In *Proc. CHI*, (2010), 453–462.

[13] Kim, J., He, J., Lyons, K., and Starner, T. The Gesture Watch: a wireless contact-free gesture based wrist interface. In *Proc. ISWC*, (2007), 15–22.

[14] Kohli, L. and Whitton, M. The haptic hand: providing user interface feedback with the non-dominant hand in virtual environments. In *Proc. GI*, (2005), 1–8.

[15] Kolb, A., Barth, E., Koch, R., and Larsen, R. Time-of-Flight Cameras in Computer Graphics. *Computer Graphics Forum 29*, 1 (2010), 141–159.

[16] Kollorz, E., Penne, J., Hornegger, J., and Barke, A. Gesture recognition with a time-of-flight camera. *International Journal of Intelligent Systems Technologies and Applications 5*, 3 (2008), 334–343.

[17] Kratz, S. and Rohs, M. HoverFlow: expanding the design space of around-device interaction. In *Proc. MobileHCI*, (2009), 1–8.

[18] Mistry, P., Maes, P., and Chang, L. WUW - wear Ur world: a wearable gestural interface. In *CHI Ext. Abstracts*, (2009), 4111–4116.

[19] Ni, T., Mcmahan, R.P., and Bowman, D.A. rapMenu: remote menu selection using freehand gestural input. In *Proc. 3DUI*, (2008), 55–58.

[20] Rekimoto, J. Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. In *Proc. ISWC*, (2001), 21–27.

[21] Starner, T., Weaver, J., and Pentland, A. A wearable computer based American Sign Language recognizer. In *Proc. ISWC*, (1997), 130–137.

[22] Vardy, A., Robinson, J., and Cheng, L.T. The WristCam as input device. In *Proc. ISWC*, (1999), 199–202.

[23] Wilson, A.D. Robust computer vision-based detection of pinching for one and two-handed gesture input. In *Proc. UIST*, (2006), 255–258.

[24] Yee, K. Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers. In *Proc. CHI*, (2003), 225–234.