Making Bare Hand Input More Accurate

Table 1: Factors hypothesized to influence hand postures and motions

Proprioception:

- The position and orientation of fingers relative to the rest of the hand
- The position and orientation of two hands with respect to each other

Context:

- External reference frame, e.g., a pyhsical workspace layout
- Anticipated shapes and positions of physical devices when grasping
- Existence and intensity of haptic feedback
- Types of transfer functions, between input device and display, and degree of visual feedback

Chat Wacharamanotham
chat@cs.rwth-aachen.de
RWTH Aachen University
52056 Aachen, Germany

Abstract

To enable seamless interaction with both virtual and physical worlds, computers need to reliably distinguish intended input from other hand movements. Accurately classifying this input can reduce mode switches needed between interacting with computers and with physical objects in the environment. I aim to investigate the influence of proprioception and interaction context on hand postures and movements. Preliminary studies suggest consistent relationships that potentially allow for a more accurate prediction of users' intentions.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation Interfaces]: Input devices and strategies.

Context and Motivation

Registering bare hand postures and motions for spatial input (e.g., arranging or assembling virtual objects) has the benefit of manual dexterity, bimanual frame of reference, and high degree of freedom of manipulation. Such input is proposed for interactive surfaces [8] and in the space above desk surfaces [16]. However, a challenge for bare hand spatial input is to distinguish movements that are intended to control the computer from those that are not. Ideally, this classification should require minimal mode-switching gestures, additional devices, and cognitive load. I believe that sensing additional information from the interactive environment can help improving this classification.

In this thesis, I investigate the influence of proprioception and interaction context on hand posture and motion, and how this can be used to accurately classify *users' intentions* behind their hand movements.

Related Work

Buxton characterizes three elementary states for graphical input devices: out-of-range, tracking, and dragging [2]. Using bare hands for input increases difficulties in discerning intended input states. To minimize errors, we need to reliably recognize users' intentions to dwell in or make a transition among these states. Previous works use

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2559955

Figure 1: Indirect multitouch setup: horizontal surface for expressive input, vertical surface for ergonomic display.

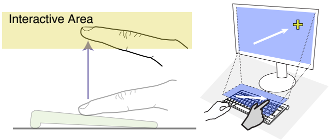

Figure 2: The area above the desktop surface provides an additional input dimension.

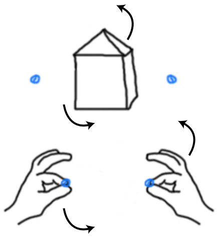

Figure 3: 3D indirect input setup. The user grabs and rotates the object with her bare hands.

explicit gestures [12], fixed invisible boundaries in midair [10], or head directions [9] for state transitions in midair input. These solutions either require equipping users with additional devices or users' awareness of gestures and invisible boundaries.

Hand postures and motions have been reported to be consistent within particular tasks such as describing 3D objects (miming motion and hand posture) [6], aiming at a target on touchscreens (visible features of fingernails) [5], and writing orientation [4]. These attributes, however, are influenced by the context of the interaction. For example, Kattinakare et al. found that distances and movement constraints influenced the accuracy of near-surface stylus movements [7]. Cockburn et al. found a trade-off between the allowed degree of freedom and the extent of visual feedback [3]. For the prehension[1] of physical objects, Barrett et al. reported that hand opening and kinematic profile were influenced by the orientation, shape, and size of the objects [1]. HCI has yet to use these relationships to improve the accuracy of predicting user intentions.

Thesis Statement

I hypothesize that the *postures and motions* of users' hands are consistently influenced by *proprioception*[2] and *context* (as defined in Table 1) and that the latter can be tracked with present-day sensors. Understanding how these factors influence hand motion will allow us to predict user intentions more accurately.

Research Goals

I aim to investigate how hand postures and motions are influenced by features in Table 1, and how these features interplay. I am interested in how these relationships allow us to improve accuracy in (1) classifying intended manipulation actions vs. spurious movements, (2) accurately determining the magnitude of the manipulation, and (3) compensating for systematic errors influenced by closed-loop visual or haptic feedback.

I focus my investigation on three desktop workspace settings that allow ergonomic interactions while being augmented with expressive input using bare hands.

Setting 1: Indirect multitouch surface: This setting consists of a horizontal multitouch surface for input coupled with a vertical screen for output (Fig. 1). Here, users benefit from both the expressiveness of the multitouch input and an ergonomic upright sitting position that allows the arms to rest on the desk surface [11].

Setting 2: Near-surface finger input: The second setup extends the typical desktop computer workspace with interactive midair layers. Users interact with desktop devices such as mouse, keyboards, or touchscreens in combination with midair input (Fig. 2). Each midair layer has a limited thickness, and they can be stacked to add additional degrees of freedom for input, e.g., by lifting the finger upward to reveal auxilary information layers [10].

Setting 3: 3D indirect input: In this setup, finger position in 3D is mapped to the virtual world shown on the screen to manipulate 3D objects (Fig. 3) [16].

Methods

For each of the three settings above, my research plan consists of three phases:

Phase 1 measures how the hands behave when asked to perform a set of interactions with the computer setting

---

[1] the action of grasping

[2] the perception of stimuli generated within the user herself

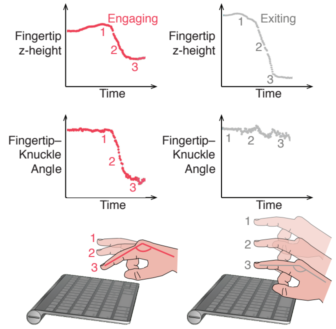

Figure 4: Hand posture changes can be used to classify between air-tapping and leaving [15].

and with the physical environment with the factors from Table 1 statically predefined.

Phase 2 uses the data from phase 1 to design classifiers for users’ intentions and evaluate their precision and recall in tasks that require frequent input state transitions.

Phase 3 investigates how dynamic factors such as transfer functions and closed-loop visual feedback influence hand posture and motion.

Results to Date

Indirect multitouch input: With another Ph.D. student, we identified four tracking-dragging switching methods from the literature: lift-and-tap, pressure switch, pressure hold (quasi-mode), and hold. We elicited the changes of the touch ellipsis over time for these methods from five users (Phase 1). We then implemented recognizers for these methods and compared them in single-finger, multiple-finger, and bimanual object manipulation tasks (Phase 2). We found that the lift-and-tap technique allows users to maintain and switch to intended input states more reliably than other methods throughout all conditions (Full paper at CHI ’13: [11]).

Near-surface finger input: With two undergraduate students, we elicited how hands behave when users access, stay inside, and leave the near-surface area (Phase 1). We found that users can maintain their fingers reliably within an area of 4 cm thickness, even when the movement is as large as 10 cm without the arm resting on a surface. The results also show that when the finger is lifted to access the near-surface area, each user has a consistent individual height to lift the finger up. We additionally found that when the user leaves the near-surface space towards a keyboard, this hand shape stays consistently flat (Fig. 4). From these results, we proposed two algorithms for recognizing users’ intentions (First half of Phase 2): (1) an algorithm that analyzes the velocity profile of the finger and dynamically places the center of the tracking area at the height determined by the first movement stroke to minimize the drifting outside the interaction area; and (2) an algorithm that distinguishes leaving the near-surface area from air tapping actions (Full paper at CHI ’14: [15]).

3D indirect input: With an undergraduate student, we investigated how accurate users specify rotation axes on a 3D object. We found that users were accurate when specifying rotation axes that aligned with the horizontal screen axis and with the axis perpendicular to the screen. Accuracy dropped significantly when they had to specify axes parallel to the vertical screen axis. We surmised that this error is caused by a slight upward perspective view. This suggests that there is an interplay between the continuous visual feedback and the proprioception of the bimanual reference frame (pending publication).

Thesis Status

I previously designed and evaluated Swabbing, an input technique that allows users with hand tremor to accurately select targets on touchscreens (Short paper at CHI ’11: [13], Workshop paper at CHI ’13: [14]. I measured the tremor in finger movements, derived an algorithm that filters the noisy input signals of tremor for touchscreen target selection, and evaluated Swabbing in two lab studies and one longitudinal study. This work inspired me to investigate bare hand input for able-bodied users.

The overview of my current thesis plan are shown in Table 2 and 3. I would love some feedback from the doctoral consortium to help evaluating and refining the direction of my remaining work.

Table 2: Progress overview

Input Settings	Phases		
	1	*2*	*3*
Touchscreen for hand tremor (special case)	✓	✓	✓
Indirect multitouch input	✓	✓	NO
Near-surface finger input	✓	☐	☐
3D indirect input	☐	NO	NO

Table 3: Planned work

Near-surface input

- Verify the near-surface state classification algorithm with users. (Phase 2)
- Extend the study to multiple midair layers; longitudinal study. (Phase 3)
- Study the influence of differently shaped objects on hand shapes during reaching, and create a classifier. (Phase 3)

3D indirect input:

- Study how much the error in 3D rotation systematically varies according to the change in parameters in the perspective projection. (Expanding phase 1)

Expected Contributions

This thesis will contribute a quantitative understanding of how proprioception and context influence hand posture and motions of bare hand input. This knowledge will help develop interaction techniques, algorithms, and guidelines to more accurately predict users' intentions.

Acknowledgements

This work was funded in part by the German B-IT Foundation.

References

[1] T. M. Barrett, E. Traupman, and A. Needham. Infants visual anticipation of object structure in grasp planning. *INFANT BEHAV DEV*, 31(1):1–9, 2008.

[2] W. Buxton. A Three-State Model of Graphical Input. In *Proc. INTERACT '90*, 449–456.

[3] A. Cockburn, P. Quinn, C. Gutwin, G. Ramos, and J. Looser. Air Pointing: Design and Evaluation of Spatial Target Acquisition with and without Visual Feedback. *INT J HUM-COMPUT ST*, 69(6):401 – 414, 2011.

[4] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *J MOTOR BEHAV*, 19(19):486–517, 1987.

[5] C. Holz and P. Baudisch. Understanding Touch. In *Proc. CHI '11*, 2501–2510.

[6] C. Holz and A. Wilson. Data miming: inferring spatial object descriptions from human gesture. In *Proc. CHI'11*, 811–820.

[7] R. S. Kattinakere, T. Grossman, and S. Subramanian. Modeling steering within above-the-surface interaction layers. In *Proc. CHI '07*, 317–326. 2007.

[8] K. Kin, T. Miller, B. Bollensdorff, T. DeRose, B. Hartmann, and M. Agrawala. Eden: a professional multitouch tool for constructing virtual organic environments. In *Proc. CHI '11*, 1343–1352.

[9] L.-P. Morency and T. Darrell. Head gesture recognition in intelligent interfaces: the role of context in improving recognition. In *Proc. IUI '06*.

[10] S. Subramanian, D. Aliakseyeu, and A. Lucero. Multi-layer Interaction for Digital Tables. In *Proc. UIST '06*, 269–272.

[11] S. Voelker, C. Wacharamanotham, and J. Borchers. An evaluation of state switching methods for indirect touch systems. In *Proc. CHI '13*, 745–754.

[12] D. Vogel and R. Balakrishnan. Distant freehand pointing and clicking on very large, high resolution displays. In *Proc. UIST '05*, 33–42. 2005.

[13] C. Wacharamanotham, J. Hurtmanns, A. Mertens, M. Kronenbuerger, C. Schlick, and J. Borchers. Evaluating Swabbing: A Touchscreen Input Method for Elderly Users with Tremor. In *Proc. CHI '11*, 623–626.

[14] C. Wacharamanotham, D. Kehrig, A. Mertens, C. Schlick, and J. Borchers. Designing a touchscreen web browser for people with tremors. In *Wksp on Mobile Accessibility, CHI '13*.

[15] C. Wacharamanotham, K. Todi, M. Pye, and J. Borchers. Understanding finger input above desktop devices. In *Proc. CHI '14*, (To appear).

[16] R. Wang, S. Paris, and J. Popović. 6D hands: markerless hand-tracking for computer aided design. In *Proc. UIST '11*, 549–558.