IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction, Kharagpur, India, December 27-29, 2012

Freehand Pose-based Gestural Interaction: Studies and Implications for Interface Design

Dustin Freeman
University of Toronto
Toronto, Canada
dustin@cs.toronto.edu

Ramadevi Vennelakanti, Sriganesh Madhvanath
Hewlett-Packard Laboratories
Bangalore, India
{rama.v, srig}@hp.com

***Abstract*—Most work on Freehand Gestural Interaction has focused on high-energy expressive interfaces for expert user. In this work, we examine the use of hand poses in laid-back freehand gestural interactions for novice users and examine the factors that impact gesture performance. Through two Wizard-of-Oz studies, one leading to the other, we observe how novice users behave under relaxed conditions. The first study explores the ease of use of a pose-based hand gesture vocabulary in the context of a photo-browsing task, and examines some of the key factors that impact the performance of such pose based gestures. The second explores pose-based interaction techniques for widget manipulation tasks. These studies reveal that while hand poses have the potential to expand the vocabulary of gestures and are easy to recall and use, there are a number of issues that show up in actual performance related to inadvertent modifications in hand pose and hand trajectories. We summarize the implications of these findings for the design of pose-based freehand gestural interfaces, which we believe would be useful for both interaction designers and gesture recognition researchers.**

***Index Terms*—user interfaces, interaction styles, freehand, hand pose, 3D, gesture, wizard of oz, laid-back.**

I. INTRODUCTION

Freehand gestural interfaces are making the transition from academic research to deployment in people's homes. However, most modern interfaces, including gestural interfaces, require our entire mental and physical attention. We must approach them if we wish to interact, removing ourselves from our current posture. The use of freehand gestures for interaction from a distance is not a new idea. However, most of the research work so far in this area has been on expert interfaces in formal environments, or gesture recognition methodologies and algorithms. Rarely has work examined “real” gestural behavior by novice users, in particular why users would provide invalid input.

This work examines issues with laid-back freehand gestural interaction for novice users. When we use the word novice, we are referring to a user that is not willing to invest a large amount of time into learning a 'virtuoso' interface. We are particularly interested in the user’s ability to recall and perform a gesture, and the factors that impact performance, in a home environment.

Freehand gestural interfaces have received a fillip in recent times due to the availability of relatively inexpensive depth cameras such as the Microsoft Kinect [9]. The most convenient design choice for freehand interaction has been to support gross hand movements such as swipes, or use the tracked hand as a whole as a cursor on the interface, with a gesture (often, just hovering for a fixed time) that serves as a click metaphor. While this is approachable for novice users, especially those familiar with computers, it presents a low ceiling for efficiency. Gesture recognition researchers have reported some success using the available low-quality depth data to discern from a distance, hand poses (e.g. palm, pointing, thumb-up/down, fist, etc.) made by the user, on the fly, even as the user moves his or her hands in a predefined gesture trajectory [13]. With improving sensor depth resolution, it will soon become possible to discern even finer hand poses. We are therefore interested in understanding the potential role of hand poses in freehand gestural interaction for laid-back interaction scenarios, suitable for novice users.

This adds to the design space of gross-movement freehand gestures an extra dimension of hand pose. In order to explore this space, we conducted two studies, one leading to and informing the other. We describe related work, and how our studies build on current gaps in knowledge.

II. RELATED WORK

There are a few commercial systems that use range cameras to detect hand position but not pose [7, 9, 10]. Previous work on interfaces that do use hand pose has been, very early proof-of-concept [1], domain-specific [6], or design instances without an attempt to generalize principles [19]. Complex hand pose has been thoroughly explored in the context of multitouch [5, 8, 17, 18]. Despite several examples in the above of compelling use of hand poses in freehand interaction, there is no high-quality solution to detect pose of an unaugmented hand from range camera data, despite ongoing work [2, 11, 16]. We do not present a technical solution in this work, but are instead interested in determining if it is worth solving, and when.

Kicker Studio [7] and OZWE [10] present design work on interface using range cameras. Microsoft’s Xbox Kinect [9] represents the first widespread commercial instance of a gestural interface, using a range camera provided by PrimeSense [12]. The three interfaces above are targeted to “live” in a home environment. However, none utilize hand pose as an input parameter; the only input possible is rather energy-inefficient movement of the whole arms and body.

978-1-4673-4369-5/12/$31.00 ©2012 IEEE

Hinckley et al. [3] identified the clutching problem as one of the most difficult problems in 3D user interfaces. Vogel and Balakrishnan's work on Distant Freehand Pointing and Clicking [14] explores hand motion plus hand pose interaction at a distance with an marker-augmented hand, They invent the AirTap gesture as a solution to the clutching problem for freehand gestural interfaces, although they use a finger, not the whole hand as we do.

Wu et al. [18] explored the use of hand poses in a tabletop interface. They came up with the idea of gesture relaxation, where the interface maintains the mode triggered by the gesture after the user has relaxed the shape of their hand into a more comfortable shape. We explore the application of this concept to freehand interaction.

III. STUDY OVERVIEW

We observe that the use of hand poses in interaction is compelling, but none of the interfaces designed thus far have explored the limitations and behavior of real users. We ran two studies, the first exploratory and the second examining interaction techniques at a lower level. Both of these studies use the Wizard-of-Oz protocol to detect hand position, hand pose and higher-level gestures. We justify the use of the Wizard-of-Oz as, at the time of the study, there was no method for detecting hand poses without augmenting the hand. Additionally, we wanted to observe user behavior with a lenient detector (the Wizard). A gesture detector developed with current technology would be too restrictive to observe casual user behavior.

IV. STUDY 1: EXPLORATORY STUDY

Our goal with this exploratory study was to explore the learnability and usage of pose-based freehand gestures by novice users for lean-back media consumption tasks. We designed a set of eight pose-based freehand gestures and evaluated their recall and usability in the context of a co-present photo-sharing scenario. During the course of the study, we observed how participants' gestures changed over time. In the analysis, we identified the factors that impact the performance or enactment of a freehand gesture. Our insights are based on observations and are qualitative in nature. The factors that impact the performance or enactment of the gestures are the focus for this paper.

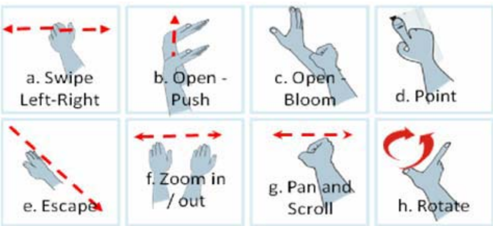

Figure 1. The Eight Freehand Gestures for Study 1.

A. Interaction Techniques

A set of eight freehand gestures were designed for the experiment: left/right, open, escape, rotate, pan/scroll and zoom in/out (Figure 1). These gestures were considered to be minimally adequate to browse and navigate a photo collection. The participant had to hold a specific hand pose while making a movement in a specific direction. All gestures were demonstrated with the arm outstretched.

B. Design Principles

The eight gestures used in this study are the outcome of a design process. We first identified all the actions a user would need to browse through a photo collection – find an album, open an album, play slide show, return or jump to the first photo while in the middle of an album etc. We then designed a small set of gestures to perform these actions. Keeping the constraints of current technology in mind, we required the gesture to be performed away from the body and distinct from one another. The gestures were limited in number but designed to be extendible by combining them with different speech "qualifiers" (not part of this study).

C. Study Setup

We used photo sharing on a large display (e.g. TV) in a living room as the scenario to represent laid-back interaction, using a simple photo browsing application built in vvvv [15]. Since photo sharing is a social task, we studied users, in pairs, as they interacted with the prototype photo application in a re-creation of a living room from a distance of 6 feet, using the set of freehand gestures. The Wizard was seated at an elevation behind the participants' couch so as to have a clear view of their gesturing hands. The moderator was seated to the participant's left and the observer to their right. All proceedings were videotaped.

D. Procedure

The experiment had three stages. Stage 1 was a picture puzzle game played to acclimatize users with gesture-only communication. One of the two participants was asked to take on the role of the User and the other, that of the Doer

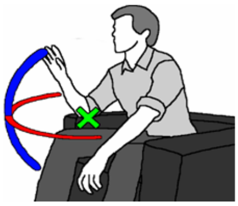

Figure 2. A resting elbow (green X) creates a pivot point and can only access 2 degrees of freedom (red and blue arcs), even through a 3-d space.

(the System). The User then instructed the Doer using only gestures to solve the puzzle. No communication was permitted in the opposite direction. Participants were free to use any gestures to communicate and were not exposed to the designed gesture set. In Stage 2, aimed at evaluating the recall and ease of use of the gesture vocabulary, the User in Stage 1left the room and the Doer in Stage 1 watched a training video (once), demonstrating the designed gesture vocabulary. They then completed a set of simple tasks using the demonstrated gestures such as browsing to a particular photo, rotating a photo, zooming in, and finding the name of the last album. In Stage 3, the User was invited back into the room and instructed by Doer on using the gestures (without the aid of the tutorial video), after which the User performed a different task - finding five copies of a particular photo placed in different albums - using the gestures taught. This task provided the participant opportunity to use all the gestures.

The Wizard manipulated the photo application in response to the participants' gestures. Task completion time was not of consequence at any stage; however each session lasted for 45 minutes to an hour.

(a) (b)

Figure 3. Changing arm positions while gesturing: (a) outstretched arm (b) elbow rested.

E. Participants

We recruited 20 `lead' users, an equal number of male and female lead users in the age group of 18 to 30 years. They were asked to bring 'a friend – someone like you'. There was no specification on the gender of the friend, so we had mixed pairs in some cases. None of the participants had experienced a Kinect, though some had read and heard about it and only a few had used a touch phone.

F. Observations

1) Gesture Recall: We observed that participants needed some practice to become familiar with gesture-only communication during Stages 2 and 3. Most users were able to correctly recall and perform the gestures without assistance. The two gestures that users had a problem recalling and needed help with were 'escape' and 'pan and scroll'. The training video was adequate instruction.

2) *Posture Relaxation:* Participants typically started with an attentive "lean-forward" stance - their torsos in an upright seating posture, their arm outstretched and the hand pose held correctly as demonstrated in the training video. However over time and with their increasing comfort level with system response, the arm slackened so that the elbow rested on the couch's armrest or the knee (Figure 3 b). Only when the system (the Wizard) did not respond did the participants correct their posture and gestures, but in most cases participants could operate the system adequately with the laid-back posture. With the observed "fixed-elbow" laid-back posture, upper arm movements are restricted and the range of motion of their hand is limited to a sphere with the radius of their forearm, as illustrated in Figure 2. This restriction had an impact on how the gesture was enacted in comparison to the original.

3) Gesture Performance: Participants seldom performed gestures consistently, as practiced, using the interface. Gestures underwent a series of modifications during repeated use. We observed different degrees of variation in hand pose, translation distance and translation direction. The modification is traceable to restrictions in movement from elbow-fixed, relaxed posture and the participants' increasing comfort with the system. Gestures used repeatedly in a sequence, e.g. 'Swipe Left / Right', and gestures used often in a task sequence, e.g. 'Escape', were more likely to undergo modification on the hand pose and the translation distance , resulting in multiple variations even from the same user (Figure 3 (a) & (b)).

Hand poses became more relaxed and ambiguous, and translation reduced compared to the original. Gestures requiring translation towards-and away from the interface (e.g. `Push') were restricted even more, especially due to the fixed elbow, and were difficult to discern even for the Wizard. Through these variations in hand pose and translation distance the translation direction was likely to be maintained (see Video). Gestures that were more discrete in nature and did not by design require much translation, for example 'Point', 'Rotate' were less likely to be modified. We noticed that participants did not treat vertical starting point as a significant feature; gestures that were similar except for a minor variation in the starting point, i.e. a horizontal movement (Swipe) vs. a diagonal movement (Escape) were unlikely to be distinguishable from each other, especially with a fixed elbow.

V. STUDY 2: POSE-BASED INTERACTION

A. Objective

Having examined the use of pose-based freehand gestures in the initial study, we sought to examine application-agnostic pose-based interaction techniques. We also sought to evaluate the learnability of "arbitrary" hand poses. Specifically, we decided to focus on the use of pose for *clutching*. For simplicity, the interaction techniques in this study use one (dominant) hand.

B. Design Principles

Given the findings of the first study, we focused on gesturing that can be performed with the elbow fixed, within the space described by Figure 2. We also determined that hovering over widgets (i.e. "dwell") as a selection or clutching technique though widely used [11], is not desirable, as the hover time puts an upper limit on how quickly the user can operate the interface. Also, with hover-based clutching, it is very easy for the user to accidentally clutch or de-clutch, especially if they are pausing to consider the state of the system. Holding the arm elevated in one position for very long may be uncomfortable.

C. Interaction Techniques

We examine one non-pose (NP) and two pose (P) clutching techniques. For the non-pose clutching technique, we chose to use AirTap [16]. For the pose-based clutching techniques, we introduce and evaluate two different pose-based clutching methods, differing only in how the user de-clutches. The first is to have the user form a pose with their hand to clutch, hold it to maintain clutch, and un-form it to stop clutching. We call this the Pose Hold (PH) technique. However, while forming a pose works well to clutch into an interaction, it is unusual for users to maintain that pose for the duration of the clutched action, as is described in the earlier study. Further, it may be difficult for the system to accurately determine if the hand is still in the same pose due to motion blur. The user may also move in such a way that the angle of their hand relative to the camera changes significantly and the system is unable to recognize the pose. The alternative we invent is Pose Tap Out (PTO), where the users relax their hand pose after forming it initially. The user de-clutches with an AirTap. PTO resembles Gesture Relaxation as described by Wu et. al [6], except with a tap to finish instead of lifting the hand off the interface.

D. Interface

The two tasks were designed to mimic many low-level interface activities. For the Sliders task, we use pose instead of movement to access different widgets across the screen. For the Manipulation task we use pose instead of a context menu to access different types of manipulations.

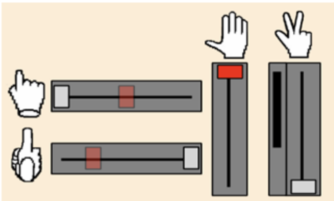

Figure 4. Sliders Task. Icons show the pose to make to capture the slider. In non-pose (NP), the icons are hidden and user must AirTap on the slider to capture it.

1) *Sliders Task:* In this task, we explored the idea of accessing a widget by forming a hand shape, as opposed to navigating a cursor to it. We had 4 sliders, 2 horizontal and 2 vertical, labeled with hand pose icons (Figure 4). The hand poses were chosen so that they were unnatural enough that the user would not make them accidentally, but not so unnatural that they would be difficult to make. To clutch a desired slider, the user simply needed to make the indicated hand pose. For the NP condition, hand pose icons are not displayed on the sliders.

In the evaluation, the goal was always to move a certain slider to within 10% of a target value. We implement 3 types of slider targets, inspired by the discussion of representative scrolling tasks by Hinckley et al. [4]. In Marked Value, we marked the target value directly with a red rectangle. In Search Value, the knob of the slider turned red. The participant had to adjust the slider until the knob turned from red to green. In Optimize Value, a bar appears next to the slider. The participant must adjust the slider until they find the maximum length of the bar; by searching the entire length of the slider.

2) *Manipulation Task:* We wanted to see if hand poses could provide a way to quickly access alternate features on a single widget, analogous to a left-click, thus reducing the number of actions required. We implemented a program where a rectangle could be translated, rotated and scaled (See Figure 5). In most modern object editing programs, these three manipulations are accessible by different handles on the object, or by clicking in a certain area near the object, i.e. near the corners to rotate. However we cannot expect the user to click on pixel-sized objects accurately. In the NP condition, when the user AirTaps on the rectangle, a context menu appears where they may select the desired manipulation type. In the P conditions, hand pose icons are shown when the user hovers over the object. If the translate is selected, the object will move relative to the translation of the users' hand. Since scale and rotate actions control a single dimension, we map horizontal movement of the hand to scaling and rotate respectively. Previous work has found horizontal movement to be more accurate than vertical [10].

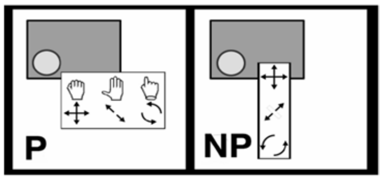

Figure 5. Manipulation Task. The user must match the grey rectangle to a red target rectangle (not shown) through translation, scaling and rotation. On the left is the pose (P) technique and on the right is the non-pose (NP) technique. In NP, a context menu appears after the user AirTaps.

E. Study Setup

To simulate a home environment participants sat in a large, comfortable leather chair facing a monitor, and adjacent to the monitor was webcam pointed at the user. We used a Wizard of Oz system for the experiment; both for hand tracking and for pose recognition. The wizard's interface was a window with the view of the user from the webcam. The wizard would indicate the location of the user's hand using the mouse, and with their other hand they would hold a key on the keyboard to indicate the hand pose. With practice during pilots, this worked effectively during the evaluation. There was admittedly some lag in the cursor movement when the user moved quickly. As most depth cameras operate at a frame rate of 30 fps, we could expect lag in a real system to be at least 30 ms.

F. Participants

We had 16 participants between the ages of 25 and 46, none of whom were from the first study. Participants were recruited by word-of-mouth at a corporate office, as well as through friends and family. We did not want technical people, so we restricted our participants to not have any programming experience, or any experience with a freehand gestural interface such as a Kinect.

G. Procedure

We opted for a mixed-design study. The P techniques, PTO and PH, are only subtly different from each other, so we could not compare them in a within-subjects design. We chose a design where each participant would do one task (S or M), using one of the pose techniques (PTO or PH) in one of two possible orders (NP-P or P-NP). Thus, there were

2x2x2 = 8 possible conditions. For each participant, the NP and P portions of the evaluation were the same: the participant would first watch a video of the technique, followed by a practice phase of 5 trials, followed by 12 "official" trials. We told participants to complete the evaluation at a comfortable pace. In the video, the demonstrator performed gestures at an unhurried pace, with a laid-back, fixed-elbow posture. The study would total slightly over half an hour.

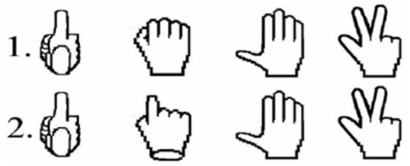

Figure 6. Poses Spaces for the Sliders task. 1 the first iteration. 2. the second iteration. We include the fourth pose, which was not in danger of conflicting as the index and middle fingers are separated.

H. Observations

1) Quantitative Results: We compared trial duration between Pose and Non-Pose interaction, and found a significant difference for the Manipulation task, ($t95 = 2.679$, $p < 0.05$), with mean trial completion times of 37.5 and 50.7 seconds respectively. We did not find a significant difference for the Sliders Task. Between Pose Hold and Pose Tap Out interaction, we did not find a significant difference in terms of trial time ($t94 = 0.211$,$p > 0.05$).

2) *Pose and Gesture Recall:* As in the first study, most users were able to form the correct pose and use the interface without errors. While poses were arbitrary, they were always displayed along the relevant widget, so users didn't have to remember the poses as in the first study.

3) *Pose Relaxation:* For the eight participants in the PTO condition, pose relaxation was optional. Three participants (#1, 8 & 15) took advantage of relaxation as expected: after clutched, their hand pose would slowly relax. Participant 7, a very tech-savvy user, would always relax his hand immediately after the interface indicated it recognized the pose. Three other participants (#3, 5 & 10) almost always held the pose, only relaxing in unusual circumstances. Participant 3 relaxed when he needed to move his hand in a way that made holding the current pose (hand flat towards screen) uncomfortable, producing a large angle at the wrist. Participant 10 always relaxed for the tapping motion at the end, perhaps perceiving a button metaphor, as it is awkward to tap a button with an unusual hand shape. Participant 13 never relaxed.

4) *Pose Space Collisions:* We had issues when we initially tested the interaction techniques. In the Sliders task, we noticed a lot of pose errors using the initial poses (Figure6-1) as participants intended to make the "thumbs up" pose, yet, while forming that shape, accidentally made the "closed hand" pose. We changed the poses to the set Figure 6-2. Now, each pose can be made from a "relaxed" pose, without accidentally forming an intermediate pose that is a defined part of the pose language.

5) *Relaxation (PH) versus AirTap (PTO):* There were two techniques to de-clutch; relaxing a pose (PH) or AirTap (PTO). The goal of the user when de-clutching is to end the manipulation within a certain target in time and space. We observed that declutching by relaxing a pose was slower, but more spatially accurate. De-clutching by AirTap was faster, but more spatially inaccurate because of the harsh movement. These techniques represent a trade-off between spatial and temporal accuracy.

Transforming the hand from one pose to another is a continuous process, so, even for our human wizard, it is difficult to tell the exact moment when the user wanted to change poses. Many of the participants had very fluid movements and changed poses while moving, leading to a lack of spatial accuracy as well. By contrast, de-clutching with an AirTap is much more discrete. It is clear when the de-clutch occurs in time, but since the de-clutch is a violent movement in space, it is often difficult to determine where the AirTap was intended on the interface.

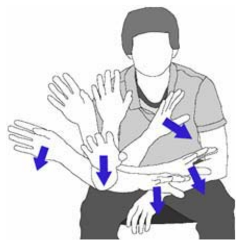

Figure 7. The direction of an AirTap depends on the direction of the arm, and could be significantly different from typically expected.

6) *AirTap Direction Variation:* As a consequence of the fixed-elbow posture, the main component of participant's AirTap direction was often tangential to the view of the camera, not towards-and-away. This is despite that in the gesture demonstration video; AirTaps were directed towards the camera. This effect was strongest when participants' forearms were angled significantly away from vertical (Figure 7).

VI. FINDINGS AND DISCUSSION

The observations from the two studies we carried out have some interesting and significant implications for the future of both pose-based free hand gesture design and gesture recognition. Here we discuss these implications along with topics for further exploration.

We found that users can easily recall pose-based gestures from a limited set. However, gestures need to be distinct, as gestures get modified during performance, and easily become undistinguishable. The ways that gestures become ambiguous may not be obvious at first. We observed in the first study that gesture parameters such as translation direction, hand pose and translation distance all undergo modification to different degrees over a period of use. In the second study, we had to change our hand pose language because collisions occurred when the user was transforming their hand from one pose to another. It is an interesting design challenge to come up with a hand pose language that has a distribution of hand poses that do not collide, and match the distribution of command frequency seen in the interface.

When users become laid-back over time, their posture changes in a way that significantly restricts their ability to perform gestures. In our studies, we universally observed an elbow-fixed posture, which reduced the degrees of freedom of hand movement. Gestures are performed significantly differently when performed with a locked elbow or an elbow rested on the knee or armrest. When users moves their hand in a way that they perceive as towards the camera, they are also moving it downwards, from the point of view of the camera.

We took suggestions for the poses used in our prototype from researchers in our gesture recognition group. The flat palm is easiest hand pose for the system to detect, but painful to use. Poses need to be designed for what users can perform, not what is easy to recognize, despite the latter being tempting. The use of speech along with gestures may help towards disambiguating a gesture more easily, but we considered it out-of-scope for this work.

Gesture recognition algorithms also need to become more robust to address variation in gesture performance by users. Since the appearance of the hand pose can change significantly in the view of a single frontal camera, more robust recognition may require additional camera sensors.

Given the current state of art in hand pose recognition using a single depth camera, and the observed variation in gesture performance, the use of hand pose relaxation, and translation direction as the key gesture parameter thereafter, seems the prudent choice for pose-based hand gesture design.

The major finding in the second study was the difference in spatial and temporal accuracy for pose changes versus hand movements. To reiterate, pose changes are spatially accurate and temporally inaccurate. Hand movements are temporally inaccurate and spatially inaccurate. Given the task, an interface designer may choose the appropriate gesture type for each action.

VII. Conclusions and Future Work

In this work, we have examined the use of static hand poses in laid-back freehand gestural interaction for novice users. The first study we conducted explored the use of a pose-based hand gesture vocabulary and examined some of the key factors that impact the performance of such gestures. The second study explored pose-based interaction techniques for generalized widget manipulation tasks. We have summarized the findings from these studies and the implications of these findings for the design of pose-based freehand gestural interfaces, which we believe would be useful for both interaction designers and gesture recognition researchers.

We have found that hand poses as shortcuts are suitable in some interaction scenarios. There are many other scenarios that should be explored. The feasibility of truly dynamic gestures and multimodal gestures should be investigated in the same style as these studies. Naturally, the implications from this study are limited due to our use of the Wizard of Oz technique. However, we have given a sense of how users behave when not restricted by a rigid recognizer. The expressivity of static hand poses also needs to be studied with a working recognizer. Are there enough intuitive but differentiated hand poses for gesture design that may be discernable both by the system and the user? Like many studies with novel interfaces, participants are very forgiving and eager. However, longitudinal studies that include ergonomics of gesture enactment are in order.

We intentionally did not include speech in our design. Despite this, we observed many unintentional meaningful vocal utterances in the first study. These are worthy of further study as part of a multimodal interface for laid-back interaction.

References

[1] T. Baudel and M. Beaudouin-Lafon, “Charade: Remote control of objects using free-hand gestures,” Communications of the ACM 36, 7 (1993), pp. 28-35.

[2] G. Hackenberg, R. McCall, and W. Broll, “Lightweight palm and finger tracking for real-Time 3D gesture control,” Proc. of IEEE VR 2011.

[3] K. Hinckley, E. Cutrell, S. Bathiche, and T. Muss, “Quantitative analysis of scrolling techniques,” Proc. of ACM CHI 2002, pp. 65– 72.

[4] K. Hinckley, R. Pausch, J.C. Goble, and N.F. Kassell, “A survey of design issues in spatial input,” Proc. of ACM UIST 1994.

[5] D. Freeman, H. Benko, M. Ringel-Morris, and D. Wigdor, “ShadowGuides: Visualizations for in-situ learning of multi-touch and whole -hand gestures,” Proc. of ACM ITS 2009.

[6] T. Grossman, D. Wigdor, and R. Balakrishnan, “Multi-finger gestural interaction with 3-D volumetric displays,” Proc. of ACM UIST 2004, pp. 61-70.

[7] Kicker Studio Design Session, http://www.kickerstudio.com/blog/2009/03/case-studygestural-entertainment-center-for-canesta/

[8] G.J. Lepinski, T. Grossman, and G. Fitzmaurice, “The design and evaluation of multitouch marking menus,” Proc. of ACM CHI 2010, pp. 2233- 2242.

[9] Microsoft Kinect, http://www.xbox.com/kinect/

[10] OZWE QB1 interface, http://www.ozwe.com/catalogue/qb1/

[11] P. Premaratne and Q. Nguyen, “Consumer electronics control system based on hand gesture moment invariants,” Institution of Engineering and Technology Computer Vision 1, 1 (March 2007), pp. 35-41.

[12] PrimeSense, http://www.primesense.com/

[13] P. Suryanarayan, A. Subramanian, and D. Mandalapu, “Dynamic hand pose recognition using depth data,” Proc. of IEEE ICPR 2010, pp. 3105-3108.

[14] D. Vogel and R. Balakrishnan, “Distant freehand pointing and clicking on very large high resolution displays,” Proc. of UIST 2005. pp. 33-42.

[15] vvvv - A multi-purpose toolkit. http://vvvv.org/

[16] R.Y. Wang and J. Popovic, “Real-time hand-tracking with a color glove,” ACM Trans. Graph. 28, 3, Article 63 (August 2009).

[17] J.O. Wobbrock, M.R. Morris, and A.D. Wilson, “User-defined gestures for surface computing,” Proc. of ACM CHI 2009, pp. 1083-1092.

[18] M. Wu, C. Shen, K. Ryall, C. Forlines, and R. Balakrishnan, “Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces,” In Proc. of IEEE TABLETOP 2006, pp. 185–192.

[19] J. Zigelbaum, A. Browning, D. Leithinger, O. Bau, and H. Ishii, “g-stalt: A chirocentric, spatiotemporal, and telekinetic gestural interface,” Proc of. ACM TEI 2010, pp. 261-264.