A Study of Hand Shape Use in Tabletop Gesture Interaction

Abstract
Although manual gesture has long been suggested as an intuitive method of input for horizontal human-computer systems, little research has been conducted into observing user preferences for tabletop gesture interaction. This is particularly the case for computer vision-based gesture input, where the recognition of different hand shapes opens up a new vocabulary of interaction. In this paper, results from an observational study of manual gesture input for a tabletop display are discussed. Implications for tabletop gesture interaction design include suggestions for the use of different hands shapes for input, the desirability of combined touch screen and computer vision gesture input, and possibilities for flexible two-handed interaction.

Keywords
Horizontal human-computer systems, gesture, hand shape.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Input devices and strategies, Interaction styles, User-centred design.

Introduction
Forming an understanding of the issues surrounding interaction with horizontal human-computer systems is a research area that has attracted increasing interest from the HCI community in recent years. This has
coincided with the maturing of input technologies for gesture interaction, for example touch screens, gloves and computer-vision based tracking and recognition. Together, these developments have brought about a proliferation of manual gesture-based tabletop interface prototypes, e.g. [2, 3, 4, 5, 6, 7, 8].

Multi-touch surfaces [4, 5, 6, 7, 8] afford opportunities for users to employ a range of hand shapes to perform single-handed and two-handed manipulation tasks, provided sufficient contact is maintained with the table surface. Wu and Balakrishnan [6] defined a gesture set that included index finger tapping or double-tapping for selection, index finger dragging for moving an object, two-finger movement for rotation or parameter adjustment, horizontal hand shape for displaying object properties, vertical hand motions for ‘sweeping’ objects to one side, two-corner shaped hands for defining an editing area. Their two-handed gestures were all symmetrical, however others have proposed asymmetric two-handed gestures for multi-touch surfaces such as selecting a file with the preferred hand while rearranging windows with the non-preferred hand to find the target window in which to drop the selected file [8].

Computer vision-based hand shape recognition for tabletop interaction [2, 3] provides a different style of interface. While it does not possess the directness and robustness of a touch screen, a computer vision-based interface can track and recognize the hand even when it is not touching the table surface. Ishii et al. use different numbers of fingers and hand movements to indicate different types of inputs, for example cutting or moving different movie clips, forwarding, rewinding, inserting or undoing in a video editing application [2].

They also recognize sequences of hand movements such as grab and release that would be difficult to achieve on a touch screen.

In this paper, we report on an observational study of hand shapes for tabletop interaction, for a range of atomic tasks. In contrast to previous work, which has defined a gesture set (and in some cases, tested aspects of it on users [6, 7, 8]), we allow users to decide how they prefer to perform the gesture interaction.

The main objective of the study was to ascertain what kinds of hand shapes users preferred for different types of tasks. It was anticipated that users would reuse simple, primitive hand shapes for a variety of different tasks, as suggested in [7], so it was also of interest to observe to what extent this occurred and for what kind of tasks.

Methods
Twenty volunteers participated in the study, thirteen males and seven females, between the ages of 19 and 50 (the majority between 20 and 40). All participants were right-handed.

The experimental apparatus comprised a large, rear-projected horizontal screen at about waist height [1]. For each task, a static image (a screen shot) occupying the full screen was displayed, and subjects were requested to perform an action on one or more objects displayed in the image. This experimental design has the disadvantage of not providing feedback to the user, however the objective of the study was to observe user behavior without imposing an arbitrarily defined gesture set upon them. Without knowing a priori what
hand shapes and movements a subject will use, it is very difficult to provide suitable feedback. The images displayed to the participants were screen shots of a standard Windows interface, with one or more icons, geometric shapes, words, or photos displayed within comfortable reach. Subjects were told that the tabletop was not touch-sensitive, and that they could assume that the overhead camera was watching what their hands were doing.

In all 36 tasks were performed by each subject, and these can be summarized as selection (of an icon(s) or text), opening (an application), moving (of an icon or a slider), scrolling, drawing, cut,copy (of an icon), rotation (of a geometric shape), zooming (of a photo), and instantiation of a floating menu. In 20 of these tasks, subjects were instructed to use one hand only (subjects were not told which hand to use). In the remaining 16 tasks, subjects were instructed to use both hands, and in four of these, they were told to perform one action with one hand and another with the other hand (subjects were not told which hand to use for which action). Participants’ responses were recorded using a video camera mounted above the tabletop and later analyzed manually, classifying each response into a particular category of hand shape and/or gesture.

Results

Preferred hand shapes
Overall, the index finger was by far the most commonly used hand shape, and was used in more than two-thirds of all 720 responses. For seven subjects, 80% or more of their responses included this shape, while only two subjects used this hand shape in less than 50% of responses. Variants on the ‘index’ shape exhibited by some subjects included use of the middle finger, or use of both the index and middle finger. One reason for the prolific use of the index finger may be the Windows-centric screen shots presented to the user during the study. Other commonly used hand shapes included a flat hand with fingers apart (‘spread’), a flat hand with fingers together (‘flat’), grabbing and releasing, and a vertical hand shape, as seen in Table 1.

Variability and consistency of responses
In terms of how many different hand shapes were used for a single type of task, selection, opening, text selection, drawing and slider operation all showed very little variation (mainly the index finger was used), while cutting, copying, rotation, zooming and instantiation of a floating menu exhibited a wide range of different shapes.

In terms of how consistent subjects were within their own choices for a single type of task, overall, subjects selected the same type of hand shape for the same type of task 81% on average. This figure varied between 69% and 99% among the various subjects, indicating that most subjects were generally consistent within their preferred method of input. One observation was that some subjects modified their strategies a little as the tasks progressed, reflecting the fact they this was the first time they had attempted this type of interaction. In terms of task types, subjects were 100% consistent within tasks involving single icon selection and opening an application from an icon, more than 90% consistent within drawing and slider movement tasks, and 65% to 85% consistent among the remaining tasks.
Table 1. The most commonly used hand shapes and their frequency of usage. The percentages shown in parentheses are the percentages of all tasks of that type which employed that hand shape, across all subjects.

<table>
<thead>
<tr>
<th>Hand shape</th>
<th>Example</th>
<th>% of total</th>
<th>Main uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index finger</td>
<td>[Image]</td>
<td>70.1%</td>
<td>Single selection, opening, drawing (&gt;90%), text selection, slider moving (&gt;80%), scrolling (&gt;70%), multiple selection, moving, rotation, zooming, floating menu (&gt;60%), cut/copy (&gt;50%)</td>
</tr>
<tr>
<td>Spread hand</td>
<td>[Image]</td>
<td>20.0%</td>
<td>Rotation (38%), multiple selection (36%), zooming (32%), scrolling (30%), floating menu (28%), drawing (25%)</td>
</tr>
<tr>
<td>Flat hand</td>
<td>[Image]</td>
<td>11.4%</td>
<td>Scrolling (25%), copying (20%), rotation, moving (15%), drawing (13%), zooming (12%), text selection, cut, floating menu (10%)</td>
</tr>
<tr>
<td>Grab/Release</td>
<td>[Image]</td>
<td>4.6%</td>
<td>Cut (25%), copy (13%), moving icons, moving slider (8%)</td>
</tr>
<tr>
<td>Vertical hand</td>
<td>[Image]</td>
<td>1.8%</td>
<td>Cut (8%), text selection, copy (3%)</td>
</tr>
<tr>
<td>Fingers together</td>
<td>[Image]</td>
<td>1.7%</td>
<td>Opening, zooming (5%), text selection, moving, cut, copy, slider moving (3%)</td>
</tr>
<tr>
<td>Fist</td>
<td>[Image]</td>
<td>1.5%</td>
<td>Floating menu (20%), zooming (3%)</td>
</tr>
<tr>
<td>‘L’ shape</td>
<td>[Image]</td>
<td>1.0%</td>
<td>Floating menu (5%), multiple selection, copy (3%)</td>
</tr>
</tbody>
</table>
**Height of hand above tabletop**
Surprisingly, considering the instruction that the tabletop was not touch-sensitive, all participants touched the table frequently. All but one subject double-tapped the table during the ‘open an application from an icon’ task. Touching the table surface presumably not only provides important feedback to the user (this may have been exacerbated in our study, where no other feedback was provided), but is a physically convenient place to rest the hand. Although the hand height from the tabletop was not explicitly measured, subjects also clearly raised their hands off the surface frequently while responding to the tasks. Notably, the ‘grab/release’ gesture, which comprised 4.6% of all responses, is not feasible using a touch screen alone. Similarly, 30% of all zooming responses were based predominantly around the subject lifting their hand(s) off the tabletop.

**Two-handed interaction**
Across all tasks that required subjects to perform exactly the same action with one and then later with two hands, a consistency of 79% was found, with the most consistent task between one-handed and two-handed operation being zooming (88%).

**Implications for Interface Design**

*Hand shape recognition for tabletop input*
Overall, the index finger was by far the most commonly used, therefore one design approach would be to require use of the index finger for all tasks, employing context to differentiate how it should be interpreted in each case. This approach, despite its perfect robustness

| 'C' shape | 0.6% | Floating menu (3%), zooming (2%) |
| Curved hand | 0.6% | Cut (5%), copy (3%) |
with respect to hand shape misrecognition, quickly finds limits in some applications. For example, dragging an index finger across an image could potentially represent drawing, scrolling or zooming according to this study. An alternative approach, more akin to previous prototypes [2, 3, 4, 5, 6, 7] would be to employ multiple hand shapes as suggested by Table 1. This permits a range of different types of input to be used without resorting to other approaches to decide on the context of a manual gesture input, however it does require the user to learn a range of shapes. If this approach were employed, the best uses for the index finger are drawing and selection (these were its two most consistent uses). Of course an interface design could employ both an index finger-based and a multi-shape approach, allowing the user to switch between them.

Two-handed interaction
If a multi-shape interface design approach is adopted, this suggests that both hands can be employed to improve the naturalness and perhaps the speed of various manipulation tasks. For example, it would be easy to imagine using the non-preferred hand to zoom in and out of an image while quickly switching between scrolling and drawing modes with the preferred hand, to rapidly mark up a large map or aerial photo.

Gesture input approaches
This study suggests the need for both a touch screen and computer vision-based gesture tracking and recognition for applications that require a wide range of commands.

Conclusion
This study offers a user-centric perspective on manual gesture-based input for horizontal human-computer systems, and suggestions for the use of different hand shapes in various commonly used atomic interface tasks. Future research includes implementing these suggested hand shapes followed by a live user evaluation, and exploration of less Windows-focused manual gesture input paradigms.

Acknowledgements
Our thanks to Prof. Bruce Thomas for his helpful suggestions and general guidance.

References
[6] Wu, M., and Balakrishnan, R., "Multi-finger and whole hand gestural interaction techniques for multi-
