ABSTRACT
Researchers are exploring touchless interactions in diverse usage contexts. These include interacting with public displays, where mouse and keyboards are inconvenient, activating kitchen devices without touching them with dirty hands, or supporting surgeons in browsing medical images in a sterile operating room. Unlike traditional visual interfaces, however, touchless systems still lack a standardized user interface language for basic command selection (e.g., menus). Prior research proposed touchless menus that require users to comply strictly with system-defined postures (e.g., grab, finger-count, pinch). These approaches are problematic because they are analogous to command-line interfaces: users need to remember an interaction vocabulary and input a pre-defined symbol (via gesture or command). To overcome this problem, we introduce and evaluate Touchless Circular Menus (TCM)—a touchless menu system optimized for large displays, which enables users to make simple directional movements for selecting commands. TCM utilize our abilities to make mid-air directional strokes, relieve users from learning posture-based commands, and shift the interaction complexity from users’ input to the visual interface. In a controlled study (N=15), when compared with contextual linear menus using grab gestures, participants using TCM were more than two times faster in selecting commands and perceived lower workload. However, users made more command-selection errors with TCM than with linear menus. The menu’s triggering location on the visual interface significantly affected the effectiveness and efficiency of TCM. Our contribution informs the design of intuitive UIs for touchless interactions with large displays.

Categories and Subject Descriptors
H.5.2 [User Interfaces]: Interaction styles.

General Terms
Design, Experimentation, Human Factors.

Keywords
Motion-base interaction, touchless interaction, menu systems, mid-air gestures, large displays, NUI, intuitive interaction.

1. INTRODUCTION
With the advent of markerless motion-tracking sensors, touchless interaction has been explored in different contexts from

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

AVI '14, May 27 - 29 2014, Como, Italy
Copyright 2014 ACM 978-1-4503-2775-6/14/05 $15.00.
http://dx.doi.org/10.1145/2598153.2598181
Therefore, it relieves users from both recalling a vocabulary of precise postures and complying with those pre-defined poses.

In a two-part, controlled experiment, we first investigated how different triggering locations of TCM affect user performance. Then, we compared between TCM and contextual linear menus with grab gestures. Our work contributes the following:

1. A command-selection technique that solely builds upon human sensorimotor abilities. Although the menu structure, the menu-triggering mechanism, and the menu-selection delimiter already exist in practice, a combination of these to harness our motor abilities is a novel approach toward designing touchless menu systems.

2. We further provide important empirical evidence applicable to the design of touchless user interfaces for large displays:
   a. Our results show that the performance of TCM depends significantly on their triggering locations on the visual display, suggesting an effect of our asymmetric motor abilities on touchless interactions with large displays.
   b. Our experiments also suggest that TCM is more efficient and causes less workload than command-selection techniques using strict postures, such as grab.

2. RELATED WORK

We build upon three research areas: large display interaction, touchless interactions, and command-selection techniques.

2.1 Large Display Interaction

Large display interaction can happen near-the-display [23] or at-a-distance [7, 20]. While multitouch techniques are commonly used for interacting near the display, at-a-distance interactions can be device-based [6] or device-free [1, 29]. Moreover, users may be standing or sitting (Figure 2).

Device-free interaction with large, distant displays frees users from employing any intermediate technology and almost has a universal appeal. However, researchers [23] caution that unguided mid-air gestures are generally less efficient and more fatiguing than device-based gestures. This suggests that touchless interaction with large displays may not be suitable for prolonged periods of intense activity. Nevertheless, existing research shows its usefulness for certain usage contexts, such as:

- Public displays, where users interact for a brief amount of time and may not spend the time and effort to connect a hand-held device with the display [25, 29].
- Sterile operating rooms, where surgeons need to browse medical images [25] and cannot touch devices.
- Interactive TVs, where users can use touchless interactions (popularly called smart interactions) to browse sporadically multimedia or access favorite commands [22, 28].

Interestingly, existing research on device-free interaction with large displays have only (empirically) investigated settings, where users are standing in front of the display. However, a sitting posture limits hand-movements more than a standing posture, and is equally relevant but largely unexplored.

2.2 Touchless Interactions: Natural?

Freehand input techniques are increasingly becoming popular due to the recent advances in markerless motion tracking and improved gesture-recognition techniques. The growing popularity of touchless interactions stems from its expectation as something natural to use. While critics have repeatedly refuted such a claim of inherent naturalness to this modality, researchers have explained that naturalness of touchless modality lies in the actions enabled and settings (or communities of practice) that give meaning to such actions [25]. In a similar line, designers have been encouraged to find naturalness in users, rather than in interaction techniques or interface components [32].

Another research domain, which investigates how to design interfaces that are intuitive to use, has proposed the intuitive interaction model [3]. Their model explains how different levels of prior knowledge – from innate abilities to expertise – and their unconscious application define an interface’s intuitiveness. For example, any interface that uses motion to attract attention (e.g., inertial scrolling) taps into our innate abilities to respond toward movement; while advanced software features often require a certain level of expertise. Until now, all touchless interaction techniques have been proposed as an extension of what has proven efficient for mouse-based, pen-based, or multitouch interfaces [1, 18]. Our design approach for touchless interactions uses human abilities to inform interface components.

2.3 Command-Selection Techniques

Command-selection techniques have been studied for decades (Table 1). Different menu techniques have been proposed for point-and-click [5, 16, 17, 26] and multitouch systems [19]. The major difference between other interactive systems and touchless systems is the device-free nature of the later. Due to the absence of a device, freehand interaction lacks control and precision [23]. Hence, it becomes important to consider the strength and limitations of human motor abilities while extending any device-based menu-techniques to touchless systems.

Prior research [1, 11, 18] proposed touchless menus by extending successful device-based menus (Table 2). From their evaluation, researchers report interesting findings on how our motor abilities limit touchless interactions. In an informal testing [11], researchers found that a 3D marking menu was most efficient, when users were not required to make accurate 3D marks: Users found it difficult to gauge a 3D angle. In [1], researchers reported that most users had difficulties constraining their gestures in a 2D plane. These observations suggest human limitations to perform 3D movements accurately in mid-air. Most importantly, this emphasizes our premise that designing touchless menus require more than a mere extension of device-based menus.

It is also important to identify the features of touchless menu techniques that require different considerations than device-based techniques. For example, with pen-input or multitouch surfaces, triggering a menu is straightforward: Users put the pen down or
touch the surface with fingers. Similarly, command-selection is delimited by breaking contact with the interface. In device-based paradigms, both linear and radial menus are common. Now without the guidance of a device, we are faced with the obvious questions: What would be an efficient triggering mechanism or a menu selection delimiter? Can we accurately make directional movements in mid-air to operate a radial menu?

All existing touchless menu techniques (Table 2) employ hand-postures (e.g., grab, finger-count, or pinch) for menu-invocation and menu-selection. Only [11] investigated scale-invariant marks as an alternative menu-selection delimiter, but reports its limitations due to 3D angular movements. [1] reports no significant difference in accuracy for linear and marking menus. Alternative to these existing techniques, we propose a touchless menu system that relieves users from both recalling a precise vocabulary of hand-postures and strictly complying with them.

3. Touchless Circular Menus (TCM)

During our qualitative exploration phase, we looked for human capabilities that could relieve users from the burden of complying with pre-defined hand-postures. It would save users from recalling a fixed vocabulary of gestures and from maintaining positions with pre-defined hand-postures. It would save users from recalling capabilities that could relieve users from the burden of complying optimally for the pose-recognizer. We found that users can reliably make directional gestures in mid-air, a sensorimotor ability that we frequently use in our everyday lives, such as during conversations or to give directions. Since such everyday movements happen unconstrained in 3D space, we observed the same problem as reported earlier [1, 11, 13]: users’ obvious difficulty in gauging 3D angles accurately. We mitigated this problem by shifting the burden of users’ input to the interface—interpreting users’ 3D translation by its orthographic projection on the 2D display. Based on our ability to make directional strokes in mid-air, and informed by some of the successful features of device-based menus, we designed iteratively a contextual menu system for large displays: Touchless Circular Menus (Figure 3).

3.1 Menu Invocation

To trigger the contextual menu, a user must cross the region-of-interest (ROI) of a display object. The ROI can be of any symmetrical shape around the center of the target, with its size directly proportional to the technique’s sensitivity. To support rapid exploration without accidental invocation of the menu, the menu appeared against the users’ direction of movement. So if users would reach the ROI of a display object from the top, or left, the menu would appear against their direction of approach: at the top-left corner of the target. Users can then make a directional

Table 1. Different features of some device-based menu techniques that have been widely studied.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>one-handed</td>
<td>one-handed</td>
<td>one-handed</td>
<td>one-handed</td>
<td>one-handed</td>
<td>two-handed</td>
</tr>
<tr>
<td>Menu triggering mechanism</td>
<td>not applicable</td>
<td>radial</td>
<td>radial</td>
<td>radial</td>
<td>radial</td>
<td></td>
</tr>
<tr>
<td>Menu selection delimiter</td>
<td>release the mouse button</td>
<td>release the mouse button</td>
<td>release the mouse button/stylus</td>
<td>moving a threshold distance from the menu-center (no crossing of any interface element)</td>
<td>re-entering the menu-center</td>
<td>none</td>
</tr>
<tr>
<td>Gesture semantics</td>
<td>none</td>
<td>none</td>
<td>scale-invariant directional strokes</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Menu breadth</td>
<td>4 (later studies suggest 8)</td>
<td>8</td>
<td>12 (later studies suggest 8)</td>
<td>8</td>
<td>8</td>
<td>~70 x 70 pixel Toolglass sheet</td>
</tr>
<tr>
<td>Expert mode</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 2. Different features of touchless menu techniques for distant and near-surface interactions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>one-handed</td>
<td>one-handed</td>
<td>two-handed</td>
<td>one-handed</td>
<td>one-handed</td>
<td>one-handed</td>
</tr>
<tr>
<td>Triggering mechanism</td>
<td>opening the hand toward the display</td>
<td>radial</td>
<td>radial</td>
<td>radial</td>
<td>radial</td>
<td>radial</td>
</tr>
<tr>
<td>Menu selection delimiter</td>
<td>closing the hand</td>
<td>closing the hand</td>
<td>closing both hands at the same time</td>
<td>releasing the pinch</td>
<td>releasing the pinch</td>
<td>moving passed the boundary of any interface element (crossing)</td>
</tr>
<tr>
<td>Gesture semantics</td>
<td>opening and closing hand</td>
<td>strokes, and closing hand</td>
<td>finger combinations with both hands, and closing hand</td>
<td>rolling the wrist, and pinching with fingers</td>
<td>pinch, and 3D directional strokes (non-dominant hand)</td>
<td>none</td>
</tr>
<tr>
<td>Menu breadth</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>8/12</td>
<td>26/48/52</td>
<td>5</td>
</tr>
<tr>
<td>Expert Mode</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

1Interactions from a distance; 2Interactions near surface
stroke toward the command (see Figure 3) and select it by crossing; but if they continue in their direction of movement the touchless circular menu would disappear.

3.2 Command Selection by Crossing
To select a command after triggering the menu, users cross it using a stroke in the command’s direction. Device-based marking menus are scale invariant, and a mark’s angle is interpreted to select commands. However, when extended to touchless techniques [1, 11], this implies the need of a posture-based menu-invocation, and a posture-based menu-selection delimiter, which negates the typical advantages of marking menus. Hence, it is not surprising that [1] reports similar accuracy for touchless implementations of linear and marking menus.

Until the crossing happens, users can cancel TCM by moving in any direction away from the triggered menu. To allow easy escape routes, we designed the structure of TCM as a semicircular array of options appearing at the top-left or the bottom-right corner of the target. As users approach the menu, to give them orientation, a trace is drawn connecting the target and the users’ hand position. Based on Fitts’ law [8], to improve users’ pointing performance, we designed the menu options to increase in amplitude as users approached them. To provide further feedback, menu options changed color when selected by crossing.

3.3 Accessing Submenus
Currently, our menu design scales up to two levels (5 x 5), with users performing continuous strokes (Figure 4). When users cross a command in the root menu, a submenu appears opposite to it, pivoted around the center of the selected command. To operate the submenu, users then change their track and cross another command. In device-based hierarchical menus, submenus appear in the same direction of the root menu. Due to the lack of precision and control of freehand movements, TCM require users to make inflections in their continuing trajectories, and thereby avoid accidental command-selections. Users can dismiss a submenu by continuing in their direction of movement after selecting a command from the root menu. In the following sections, we discuss our experiments with single-level TCM.

4. EXPERIMENT: PART I
TCM are contextual menus for large displays, and ideally they are expected to perform optimally across the entire display canvas. Hence we conducted a controlled experiment to investigate how effectiveness and efficiency of TCM is affected by their triggering locations on the visual interface.

4.1 Hypotheses
Our menu design was motivated by our abilities to make directional strokes in mid-air. When we move our arms in mid-air, biomechanical properties of the human body (such as position of the forearm relative to the upper body) affect how accurately and quickly we can make arm-movements. Certain arm-postures result in more static equilibrium of the body and hence are more comfortable than others. The absence of any guidance device, such as a remote [23] or a wand [6], further aggravates the control and the precision of such mid-air movements [23]. Based on these theories, we made the following hypotheses:

H1: Triggering location will affect the efficiency of TCM.
H2: Triggering location will affect the effectiveness of TCM.

Furthermore, in our experimental setup, based on our sensor’s tracking specifications and pilot testing, we ensured that the tracking performance was optimal across all triggering locations.

4.2 Apparatus
The high-resolution large display (Figure 1) integrated by Fakespace Systems comprises of eight 50” projection cubes laid out in a 4 x 2 matrix. It is driven by a single computer. Each cube has a resolution of 1600 x1200 pixels, resulting in a 160” wide by 60” high display with over 15.3 million pixels. Our goal was to evaluate TCM as a potential user interface component using off-
the-shelf motion-capture sensors. We used a Kinect for Windows to track users’ hand position, and recognize gestures. Though this system is limited from a technological perspective, we wanted to evaluate user performance with a commodity-range camera. The experiments were written in C# running on Windows 7, and were implemented with OpenNI 1.4 SDK and PrimeSense’s NITE 1.5.

4.3 Task

To test our hypotheses, we designed a menu selection task informed by the ISO standard 9241-9 [15]. On a large interactive display (Figure 1), participants were shown a circular arrangement (594-pixel diameter) of 9 equally sized (320-pixel) squares, aligned to the horizontal and the vertical center of the background (Center, N, NW, NE, S, SW, W and E). Participants’ task was to invoke TCM for a (randomly generated) white square and select the ‘Remove’ command by crossing (Figure 1). The ROI was set to 256 pixels, and TCM’s diameter was set to 400 pixels.

4.4 Procedure

We recruited 15 right-handed participants (4 females) from a university campus, with 8 participants having prior familiarity with touchless gestures, and 11 participants below 30 years of age.

Participants sat in a comfortable couch (Figure 1) at 2.25 m away from the display (~1 m away from the sensor), and took 20-30 minutes to complete all trials. Prior to the experiment, all participants completed 3 blocks of practice trials. Throughout the experiment, participants were required to take at least a 10-second break in between each blocks. Trials were randomized within subjects. In summary, the study design was as followed:

9 triggering locations (trials)  
x 7 blocks  
x 15 participants  
= 945 total trials

Participants hovered over a ‘Start’ circle, to begin a block. Trials were defined as a successful selection of the ‘Remove’ command. We recorded performance time, command-selection errors, and encouraged participants to make comments about the menu. Time was measured from the target’s appearance to a successful command selection. A command-selection error was recorded, when participants selected a wrong command from the triggered menu. When a command-selection error occurred, ‘error’ was flashed on the display, and the trial restarted.

Participants received a $20 gift card for 2 hours of participation.

Successful Trigger Rate.

TCM are contextual menus. They are triggered when users reach the ROI of a target, and are dismissed if users move away from the triggered menu. During selecting commands, users may inadvertently dismiss the menu before selecting any command and re-trigger it again. To understand how unwanted menu dismissals affect users’ efficiency, we defined successful trigger rate as successful triggers / (successful + unsuccessful triggers). Successful triggers: When users trigger a menu, and continue to select a command from the triggered menu. Unsuccessful triggers: When users trigger a menu, but the menu is dismissed before any command is selected. Obviously a high successful trigger rate would increase a menu’s efficiency as users would not have to re-trigger it.

4.5 Results

Performance time was normally distributed, but error rate and successful trigger rate were not. We used repeated measures ANOVA (and its nonparametric version) for data analysis.

Triggering location {Center, N, NW, NE, S, SW, W and E} had a significant effect on task time, $F(6.9, 718.12) = 4.74, p < .001$. Planned contrasts revealed that both north (3466 ms) and south (3646 ms) locations took significantly more time than the center location (3095 ms), $p < .001$. We found a significant learning effect across blocks, $p < .01$. Participants were about half a second faster in the last block than the first block. Menu triggering location also significantly affected successful trigger rates, $\chi^2(8) = 18.83, p < .05$. Across all triggering locations, the average successful trigger rate per trial was 88.4%. $H1$ was supported.

Triggering location did not significantly affect error rate. The average error rate (participants selecting a wrong command from the menu) across all triggering locations was 2.7%. During 88.5% of the command-selection errors, users chose the nearest neighbor options (‘Send’ or ‘Share’, Figure 1). $H2$ was not supported.

Apart from the initial novelty effect that excited the participants, they appreciated the use of fewer muscles in the crossing gesture. However, participants also commented on the lack of control: “I felt I had to rush to select the menu option” and precision: “It was sometimes difficult to be precise.” Overall, users liked the feedback language of the menu: “It feels like the menu is a bow, and I am aiming an arrow to select one of the options.” Finally, some users were excited about their own performance: “I was surprised that I could do so well.”

4.6 Discussion

From our user study, we learned the following about TCM:

1. Depending upon the menu’s triggering location on the display users’ control on their hand movements varied significantly.

2. A visual comparison of successful trigger rates and time spent in command-selection across all triggering locations (Figure 5) reveals that unsuccessful triggers were not the sole reason behind the variability in efficiency of TCM. For example, at certain triggering locations (such as, N and S), users did not lose the triggered menu more than the average but spent more than average time in command-selection. One possible explanation is that participants had to put more physical effort, thereby spending more time at certain triggering locations.

Overall, our results suggest that touchless interaction with large displays is significantly affected by the asymmetric nature of human motor abilities (control and precision).
The average efficiency of TCM was 3.3s and accuracy 97.3%. [1] reported performance measures for linear menu as 6.6s (94.2%), marking menu as 7.2s (95.3%) and finger-count menu as 8.5s (93.4%). Our results cannot be directly compared to [1] because we used different experimental tasks and menu hierarchy (details in Table 4). However, this is an encouraging result. Although such performance time is higher than the menu-selection time in typical Xbox games, it is important to note that Xbox gamers are continually (visually) guided to position themselves in an optimal space—in front of the sensor (2-3 meters)—so that the sensor can track users’ entire body [21]. TCM was implemented using hand tracking algorithms that did not require whole body tracking.

Limitations. Due to sensor limitations, when participants moved their arms very fast, tracking points were lost, thereby causing unwanted menu dismissals. This may have decreased the successful trigger rate for TCM. As TCM do not require any static poses, their invocation and selection suffers from certain limitations. To provide users escape routes, the breadth of TCM is limited to 5. Moreover, menu invocation is not tolerant to target overshoing (when hand movements trail the eye gaze), and may cause accidental invocations if users decide to change the direction of movement for target acquisition. One possible approach to mitigate these limitations is using explicit dynamic gestures (e.g., lassoing or pigtails [14]) as a menu-selection delimiter. As a delimiter, dynamic gestures would be more efficient than static poses, as users would not have to halt-and-execute a pose, but fluidly end the selection. Furthermore, we do not foresee a large number of commands in large-display touchless interfaces, as they are not fitted for intense editing but suited for exploratory data browsing. As the location of menu options in TCM depends on users’ direction of movement, users cannot exploit spatial memory to locate them. However, TCM appear at either the NW- or the SE-corner of a target in a symmetric layout (as mirror images of one another). Further research is required to understand if users can exploit this symmetry to locate menu options in TCM.

External Validity. Our findings can be generalized to settings, where users are sitting away from a large display, facing the display, and within the sensor’s tracking range. Since sitting posture already constrains our arm movements to a certain extent (e.g., when leaning back or resting the elbow), we expect similar or better user performance of TCM in a standing posture.

Part I of our experiment suggested an encouraging performance of TCM. However, it was unclear how this performance would compare with menu systems that employ static postures, especially in similar settings.

5. EXPERIMENT: PART II

Part I of our experiment focused on investigating the performance of TCM across different triggering locations. In part II, we investigated how the overall user experience of TCM compares with contextual linear menus using grab gestures.

Contextual Linear Menus. With linear menus, participants could point-and-select a display object by doing a grab gesture. They would do a grab gesture by making a fist, and opening their hand again (Figure 6). To trigger the linear menu, users would do a grab gesture on a target, and the menu would appear to its right. Then users would select a command by doing another grab. In this technique, gesture registration happens with the first grab; then gesture relaxation follows, where users point to a command, and then grab gesture is reused to select that command [33].

Figure 6. To trigger linear menus, users made a grab gesture on the target by closing (left) and opening their hand (center). A command was then selected by another grab gesture (right).

5.1 Hypotheses

Based on previous research and our pilot studies, we made the following two hypotheses:

H3: Compared with TCM, the linear menu design uses more muscle groups [31] and involves reuse of gesture primitives [33].

H4: We hypothesized that TCM would be easier to use than linear menus because of the use of more muscle groups [31] in grab pose than in a crossing gesture.

5.2 Task and Procedure

For part II of our study, we ran the linear menu experiment to compare its user experience with TCM. Thus, part II used the same experimental task, procedure and evaluation metrics as part I. However, due to sensor limitations, we designed the command-selection task for linear menus only at six different locations (Center, N, NW, NE, W and E). A successful grab gesture on the target triggered the linear menu 200 pixels right and 700 pixels top from the top-right corner of the target. The menu consisted of five equally sized (256-pixel) squares (Figure 6), and the participants’ task was to select the ‘Remove’ command by a grab gesture. Users would dismiss the linear menu if they performed a grab gesture anywhere outside the menu. We recorded menu dismissals to calculate the menu’s successful trigger rate. Self-reported system usability scores were recorded using SUS, and perceived workload using NASA-TLX. Participants responded to SUS [4] after using each menu (except questions 1, 2 and 6). After using both the menus, they completed the NASA-TLX scale [12].

Since we conducted both parts of our experiment on the same day, and with the same participants, the menu condition was counter-balanced. Participants took a break of about 10 minutes in between sessions. Trials were randomized within subjects.

Apparatus. The linear menu experiment was written in C running on Windows 7 and was implemented with OpenNI 2.2 SDK, NITE 2.2 and Windows SDK 1.7. For the grab gesture recognition, we used PrimeSense’s Grab detector library [27].

5.3 Results

5.3.1 TCM are More Efficient than Linear Menus

TCM ($M = 3.3s$, $SD = .7$) were more than twice as fast as the linear menus ($M = 7.4s$, $SD = 2$), $t(14) = 7.43$, $p < .001$, $r = 0.89$. H3 was supported. However, there was no significant difference in successful trigger rates between TCM ($Mdn = 89\%$, $IQR = 9.55$) and linear menus ($Mdn = 92\%$, $IQR = 7.62$).
5.3.2 TCM are Less Effective than Linear Menus

TCM (Mdn = 1, IQR = 3) were significantly less effective than linear menus, Z = 2.68, p < .01, r = .69. With TCM, on average, users made about 3 errors per 100 trials. Given the lack of precision and control associated with freehand movements, 97.3% accuracy is an encouraging result. Leaving out the outliers, users made no command-selection errors with linear menus.

5.3.3 TCM elicit Less Workload than Linear Menus

System usability scores were not significantly different between TCM (M = 82.86, SD = 13.58) and linear menus (M = 72.62, SD = 19.96). However, overall workload was significantly higher for linear menus (Mdn = 20.83, IQR = 9.17) than TCM (Mdn = 19.17, IQR = 19.17), Z = 2.89, p < .01, r = .75. When the NASA-TLX scale was analyzed separately, we found significant differences between the menus regarding physical demand, temporal demand, and effort. H4 was partially supported.

5.3.4 User Comments

Compared with TCM, linear menus received mixed user reactions. A male participant younger than thirty was enthusiastic: “This is how I envision using touchless gestures.” A female participant over fifty said: “It was a lot of effort.” She pointed out that Arthritis patients would find it difficult to do grab gestures.

Surprisingly, the linear menu had an accuracy of 100%, which means participants did not select any wrong command from the triggered menu. Nevertheless, participants lost the triggered menu in 8% of the trials. Our videos revealed that while grabbing a menu command, participants often moved their hands horizontally away from a specific command (right or left); thereby dismissing the menu. As they did not move their hands up or down, and the linear menu options were stacked vertically (Figure 6), command-selection errors did not occur. Compared with linear menus, TCM had an accuracy of 97.3%. This maybe because:

1. The options in linear menus were 256-pixel squares and more than eight times wider than the options in TCM (306 pixels in length, 30 pixels in breadth).
2. In linear menus, users triggered the menu with a grab gesture. They also selected a command using another grab gesture. Between these two gesture registrations, users could move their hands freely around the display. However, for TCM, after the menu is triggered, users could inadvertently move their hand and select a wrong command. Unlike linear menus, TCM required users to constrain strictly their freehand movements after triggering the menu.

Overall, we learned the following from part II of our study:

1. In our experimental settings, TCM were more efficient, but less effective than linear menus. TCM elicited significantly less workload than linear menus.
2. Compared with linear menus, participants were more than two times faster with TCM, but there was no significant difference in successful trigger rates between them. This suggests that menu-triggering by reaching the ROI (88% accuracy) performed on par with menu triggering by grab (92%). Moreover, participants seemed to spend more time with linear menus due to more effort required in performing a grab gesture than a crossing gesture.

Limitations. Capabilities of our tracking sensor limit our results. An ideal gesture recognition algorithm may have made the linear menus more efficient than TCM. In this work we proposed a touchless menu system that does not employ any pose-recognition techniques, but performed on par with current available menu techniques (Table 4). Furthermore, with future improvements in tracking capabilities, we expect that TCM will outperform linear menus because it builds on users’ previously learned skills of making in-air directional gestures. Aimed at a preliminary understanding of a touchless menu system that does not employ any pose-recognition techniques, both our visual interface and task was simple (always selecting the ‘Remove’ command from a single-level menu). Future research is required to assess the user experience of TCM in more realistic usage scenarios.

External Validity. Large displays are becoming popular in consumer electronics (e.g., interactive TVs), healthcare settings and public spaces. Touchless gestures offer a promising interaction modality for these novel devices. Our proposed touchless menu system uses dynamic gestures for selecting commands on large displays while interacting from a distance.

Table 3. Contrasting characteristics of touchless circular menus vs. contextual linear menus

<table>
<thead>
<tr>
<th>Menu selection delimiter</th>
<th>Touchless circular menus</th>
<th>Contextual linear menus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossing the boundary of an interface element</td>
<td>Grab gesture</td>
<td></td>
</tr>
<tr>
<td>Reaching a pre-defined ROI of a display object</td>
<td>Grab gesture</td>
<td></td>
</tr>
<tr>
<td>(Dynamic) stroke</td>
<td>(Static) grab</td>
<td></td>
</tr>
<tr>
<td>Hand tracking</td>
<td>Hand pose recognition</td>
<td></td>
</tr>
<tr>
<td>Radial</td>
<td>Linear</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Performance measures across touchless menus.

<table>
<thead>
<tr>
<th>Touchless system</th>
<th>Menu options</th>
<th>Average time</th>
<th>Average accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear menu [1]</td>
<td>8 x 8</td>
<td>6.6s</td>
<td>94.2%</td>
</tr>
<tr>
<td>Marking menu [1]</td>
<td>8 x 8</td>
<td>7.2s</td>
<td>95.3%</td>
</tr>
<tr>
<td>Finger-count menu [1]</td>
<td>5 x 5</td>
<td>8.5s</td>
<td>93.4%</td>
</tr>
<tr>
<td>Contextual linear menu</td>
<td>5</td>
<td>7.4s</td>
<td>100%</td>
</tr>
<tr>
<td>TCM</td>
<td>5</td>
<td>3.3s</td>
<td>97.3%</td>
</tr>
</tbody>
</table>

1Participants standing; 2Participants sitting.
6. CONCLUSION AND FUTURE WORK

Prior work on touchless interaction with large displays contributed interaction techniques that require users to comply with pre-defined postures. Our research suggests that dynamic gestures—such as simple crossing—when coupled with human sensorimotor abilities—such as making directional strokes—is more efficient than posture-based techniques. Specifically, whereas existing touchless menu systems for selecting commands from a distance are posture-based [1, 18], we introduced a novel touchless menu system (TCM) for large displays, which solely uses our ability to make directional strokes in mid-air and relieves users from recalling a vocabulary of gestures.

Our comparative study suggests that TCM are more than two times efficient than contextual linear menus using grab gestures. Users also perceived less workload with TCM. However, TCM caused 3% more errors than linear menus. This may happen because, unlike linear menus, TCM required users to constrain strictly their freehand movements after triggering the menu. To address this problem, we are exploring ways to minimize these errors by guiding users’ mid-air strokes during menu selection.

We also found that the asymmetric nature of human motor capabilities significantly affected the efficiency of our proposed touchless circular menus. We expect this effect to be pervasive in touchless interactions with large displays, which requires further investigation. The design of future touchless interfaces can be informed by identifying these asymmetric motor abilities.

Our future work will expand on touchless menu designs in two ways. First, we are conducting empirical studies with hierarchical TCM to evaluate their effectiveness. Second, while this work introduced TCM as a one-handed technique, we are investigating bimanual gestures to operate menu levels: the non-dominant hand crossing the 1st level, and the dominant hand crossing the 2nd.

7. ACKNOWLEDGMENTS

We thank the study participants and UITS AVL. This research is partially supported by an IUPUI Research Support Funds Grant.

8. REFERENCES