An Evaluation of Discrete and Continuous Mid-Air Loop and Marking Menu Selection in Optical See-Through HMDs

Zhi Han Lim
Department of Engineering
University of Cambridge
zhl23@cam.ac.uk

Per Ola Kristensson
Department of Engineering
University of Cambridge
pok21@cam.ac.uk

ABSTRACT

This paper investigates discrete and continuous hand-drawn loops and marks in mid-air as a selection input for gesture-based menu systems on optical see-through head-mounted displays (OST HMDs). We explore two fundamental methods of providing menu selection: the marking menu and the loop menu, and a hybrid method which combines the two. The loop menu design uses a selection mechanism with loops to approximate directional selections in a menu system. We evaluate the merits of loop and marking menu selection in an experiment with two phases and report that 1) the loop-based selection mechanism provides smooth and effective interaction; 2) users prioritize accuracy and comfort over speed for mid-air gestures; 3) users can exploit the flexibility of a final hybrid marking/loop menu design; and, finally, 4) users tend to chunk gestures depending on the selection task and their level of familiarity with the menu layout.

CCS CONCEPTS

• **Human-centered computing → Gestural input**;

KEYWORDS

Menu selection; optical see-through head-mounted display

ACM Reference Format:
Zhi Han Lim and Per Ola Kristensson. 2019. An Evaluation of Discrete and Continuous Mid-Air Loop and Marking Menu Selection in Optical See-Through HMDs. In *21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19), October 1–4, 2019, Taipei, Taiwan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3338286.3340127

1 INTRODUCTION

Optical see-through head-mounted displays (OST HMDs) are under intense industry development and as a consequence some OST HMDs are now made available for reasonably affordable prices, such as the Microsoft HoloLens. While the mainstream OST HMD market is yet to emerge, OST HMDs are already being deployed into vertical markets, such as security and defense, construction, engineering and manufacturing. Current OST HMDs are limited in mass market appeal, primarily due to their bulky form factor, limited field-of-view and power consumption demands. However, devices, such as Microsoft HoloLens, are already excellent protoyping testbeds for investigating next-generation user interfaces for OST HDMs.

One primary challenge in such a future is the demand for interaction techniques that are robust in the presence in an inherently noisy sensor environment. Noise is unavoidable in any interface and arises from multiple sources such as the user's neuromuscular and cognitive systems, device sensors and inference algorithms (e.g., gesture recognition algorithms). While current limitations of OST HMDs, such as form factor and low field-of-view, are likely to be solved in the not too distant future, providing uninstrumented users with fluid interaction techniques that are resilient to noise is a fundamental obstacle blocking mainsteam adoption of OST HMDs. Unlike typical computing environments, such as laptops and capacitive touchscreen-enabled mobile devices, OST HMDs must infer the user's intention using noisy sensors, such as depth sensors, and probabilistic algorithms, such as random decision trees and deep neural networks. Further, OST HMDs interactions are situated in a particular physical environment, and the characteristics of this environment will introduce additional noise, which is difficult to control and model.

In this paper we evaluate discrete and continuous marking menu and loop menu selection. We investigate design considerations for how to transplant a known efficient and flexible control interface, the marking menu, to OST HMDs and ensure it is sufficiently noise resilient. Such a control interface is useful for many situations. First, it is based on a known efficient menu technique which provides users with a seamless transition from novice-to-expert behaviour. It is therefore a suitable component of standard OST HMD user interfaces, for instance to allow users to browse command hierarchies and trigger commands. Second, we envision a future where OST HMDs are used, among other things, to seamlessly interact with smart devices to remotely trigger a range of functions, such as using an OST HMD to remotely configure a heating control appliance in a home or to reprogram robot in a smart factory. A robust control interface is vital for such interactions, as errors may incur a substantial recovery cost.

However, the marking menu is not fully robust to noise due to its reliance on sequences of linear strokes, which are difficult to carry out reliably using OST HMD sensors, such as a front-mounted depth sensor. We therefore also explore a variant of the marking menu which in this paper will be referred to as the *loop menu*, inspired by FlowMenu [17].

Our key findings are 1) the loop-based selection mechanism provides smooth and effective interaction; 2) users prioritize accuracy and comfort over speed for mid-air gestures; 3) users can exploit the flexibility of a final hybrid loop and marking menu design; and, finally, 4) users tend to chunk gestures depending on the selection task and their level of familiarity with the menu layout.

2 RELATED WORK

Gesture design and gesture-based menu design have been extensively researched in the past. Dachselt and Hubner [12] surveyed previous 3D menu applications, Cheng et al. [11] reviewed recent research activities with 3D hand recognition using depth sensors and Bailly et al. [5] analyzed visual menu techniques. Zhai et al. [43] reviewed gestural stroke-based user interfaces while Delamare et al. [13] surveyed existing gesture guiding systems and provided an online tool assisting designers to quickly identify and compare current solutions that match their required design specifications. Below we review the literature from two perspectives. We first review prior research on gesture control techniques and thereafter we review prior research on gesture design theory.

2.1 Gesture Control Techniques

2.1.1 Gesture Discrimination by Hand Configuration. The finger-count menu [23] allows menu selection via recognition and matching of the number of fingers the user is extending to a labelled menu item. However, recognition of hand form does not allow smooth navigation through many hierarchies in a single cluster of actions, otherwise referred to as dynamic co-articulation [9]. Aoki et al. [2] argues that there will be wasted motion in switching one hand form to another in doing so.

2.1.2 Gesture Discrimination by Relative Hand Position/Orientation. Prior research has extensively researched pointing gestures [32, 40] and wrist tilt gestures, [30] respectively. Vogel and Balakrishnan [40] developed and compared different freehand techniques for pointing and clicking from a distance. They used features such as velocity, acceleration, absolute position and movement axis of the finger to recognize intended clicks. They tested raycasting (from tip of index finger), relative raypointing (with tracking hand positions on a vertical plane) and a hybrid of both techniques and found out that freehand raycasting faced high error rates in selecting small and medium targets. Freehand click-and-point can also lead to the "Heisenberg effect", a phenomenon where the action of clicking causes the hand position to be adjusted [8]. In lieu of the unconstrained freehand movements, which causes inaccuracy in raypointing, Shoemaker et al. [37] proposed that distance-independent techniques are preferable.

Instead of using the hand/finger to point-and-click, Pfeuffer et al. [32] explored using gaze-and-pinch interactions for manipulations in virtual reality. Nonetheless, like any pointing-based interaction technique, each selection is performed with a discrete action or gesture, which does not allow chunking of multiple selections or dynamic co-articulation. This reduces the suitability of such techniques in menu selection systems as they do not support quick navigation of menu hierarchies. Lastly, the rapMenu [30] is an orientation-based gestural menu system that uses wrist tilts and multiple distinct pinch gestures to make menu selections without requiring precise movements.

2.1.3 Gesture Discrimination by Shape/Path. The marking menu [25] is a 2D menu interface design with a similar layout to the pie menu [10]. In a marking menu, a menu selection is made by drawing a gesture (denoted *mark*) from the centre to the desired menu item. It allows scale-independent selection, open-loop and hence eyes-free operation for experts, while also providing support for novice usage with closed-loop visual menus and novice-to-expert transitions. It allows quick access to menu items in deeper hierarchies via dynamic co-articulation in a single continuous stroke.

Zone and Polygon [45] and Flower [4] menus are variants of marking menus with different layouts, providing users with either better accuracy or menu item capacity. Zhao et al. [46] also presented a variant using multiple discrete strokes (higher-order gestures [43]) known as "simple marks". Marking menus have also been adapted for text entry [39] and parameter controls [33]. Finally, introductions of new input devices have stimulated further explorations of marking menus. For instance, a two-handed marking menu for touchscreens has been presented [21]. Ren and O'Neill [34] proposed a 3D marking menu for freehand gestural interaction. They concluded that the additional z-axis provided by the 3D menu should be avoided as accuracy and movement speed along the z-axis is lower for users. Other examples include the wave menu [3] and the double crossing technique [29].

Besides a directional path, drawing different shapes in mid-air can also be used to select objects or menu items. The Vision-based Unicursal Gesture Interface (VUGI) [2] allows menu selection by drawing shapes. It is similar to QuikWriting [31] and Cirrin [27], where users can make text-entry by moving out of the central area to designated zones for selection and returning back to the central area for confirmation.

2.2 Gesture Interface Design Theory

Prior research [16, 26, 28] has indicated that human vision consists of a ventral stream for conscious perception of objects (identify a menu item) and a dorsal stream for unconscious, online control of visually guided action toward objects (reach for the menu selection). Furthermore, the execution (dorsal stream) of limb movements can be viewed as two distinct component phases: an initial, ballistic phase and a subsequent phase where refinement and error-correction takes place, usually relying on visual feedback to reduce the error between the effector and the target [15, 42]. Studies [14, 41] have also shown that dynamic visual cues in both the central and peripheral field-of-view affects motor behaviour. It is also generally viewed that providing information regarding the learner's success in meeting the environmental goal (knowledge of results) after a response is widely regarded as critical for motor learning [35]. Thus, mid-air menu selection systems should account for motor movement and visual feedback considerations. This includes the error correction phase in selections made with upper limb movements, potential visual distractions often present in optical see-through devices and lastly, visual feedback for error correction. A plausible approach would be to avoid target selection methods and explore trajectory based movements, such as goal-crossing [1].

Kurtenbach [25] improved learnability and memorability of gesture-based interaction systems by catering for novice exploration, expert operation and training to facilitate novice-to-expert transition. He also explored how an "unfolding interface" or self-revealing devices can help achieve this novice-to-expert transition. Other work, such as OctoPocus [7] and LightGuide [38], used self-revealing guidance to provide feedforward and feedback, teaching users new gesture commands and proper movements respectively. Bailly et al. [6] explored mid-air marking menus and Zheng et al. [47] explored the positive and negative aspects of continuous and discrete marking menus defined on a grid.

3 LOOP AND MARKING MENU SELECTION

We investigate two menu selection techniques: the marking menu and the loop menu, which is a refinement of the marking menu and closely related to FlowMenu [17] in its design.

A marking menu works as follows. A user makes a menu selection by first triggering the menu, which is typically a pie menu centred around the invocation point. In novice mode the underlying pie menu is visualized to the user. The user then navigates the hierarchical pie menu by articulating linear strokes towards the desired pie slices. When a slice has been selected it can either trigger a sub-menu, and the process repeats, or the selected pie slice can be a menu item, which will be triggered if selected by the user. In expert mode, the underlying pie menu is not shown to the user. Instead, the user selects the desired menu item by articulating a series of continuous linear strokes, mimicking the same movement pattern as in novice mode. Since the movement pattern is the same in both novice and expert mode, this allows users to implicitly learn the motor memory patterns for menu selections in novice mode and seamlessly transition from closed-loop visually guided novice behavior to open-loop direct recall from motor memory expert behavior. This principle is also underpinning the gesture keyboard mobile text entry method [22, 44].

To encourage users to transition to expert mode, the visual depiction of the marking menu is typically not revealed until a set timeout after the menu has been triggered. An often overlooked implementation detail of the marking menu is that if the user stops their movement for a set duration in expert mode, which indicates the user has forgotten the direction for the next selection, the marking menu switches back to novice mode. In addition to the original continuous marking menu it is also possible to implement as a discrete marking menu where each linear stroke is separately delimited.

The loop menu uses a selection mechanism with loops instead of linear strokes to approximate directional selections in the menu system. Our design is based on the hypothesis that moving in loops for menu selections is easier, faster and more accurate than producing linear motions in unsupported 3D space for an OST HMD.

To investigate loop and marking menu selection in an OST HMD we use a Microsoft HoloLens. The loop and marking menu gestures are delimited by the default built-in HoloLens 'click', or air tap gesture (touching the index finger with the thumb).

To make a selection:

(1) The user starts with a 'click' gesture to trigger hand tracking.
(2) While the user has the click gesture triggered (i.e., the thumb and index finger remain in contact), the user moves their hand in mid-air, in the direction towards the intended menu item and exits the inner octagon. A selection occurs upon re-entering the octagon. If the item selected is a parent item, the menu updates to reveal subsequent child items.
(3) The user either releases the 'click' gesture for selection confirmation, or continues to navigate deeper through the menu hierarchy by repeating step 2.

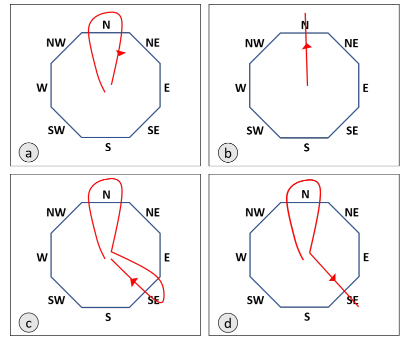

Figure 1: Illustration of making a menu selection: a) single selection with a loop, b) single selection with a line, c) multi-hierarchical selection with multiple loops, d) multi-hierarchical selection with loops and ending with a line.

For shorter gesture, and hence selection, duration, selections can also be made by drawing lines instead of loops. Simply release the 'click' gesture early, before re-entering the inner octagon. Though faster, this method of selection does not allow for chunking for navigation through the menu hierarchies. Users can also choose to combine loops and lines for multi-hierarchical selection, with lines used as the last selection.

The selection process is illustrated for all cases in Figure 1.

The system detects loop menu selections as follows. Each menu item is assigned to an angular range, with the range dependent on the number of menu items. The user selects an item by tracing a looped path in the general direction towards that menu item. The algorithm takes data points throughout the entire loop. The interpreted direction of the drawn loop is the line joining the average point of the loop, ($\frac{\sum_1^n X}{n}$, $\frac{\sum_1^n Y}{n}$) and the middle point between ($\frac{I_x+J_x}{2}$, $\frac{I_y+J_y}{2}$) and (0 , 0), where n is the total number of data points per loop, X and Y are the horizontal and vertical coordinate of a given point on the loop, I is the point where the loop exits the inner octagon and J is the point where the loop reenters it. Figure 2 illustrates the key aspects of the system's selection mechanisms.

This approximation method allows users to easily select off-axis directions without having to be precise with their gestures. Users would not need to focus on exiting and entering the exact zones or follow a specific trajectory path. This allows gesture articulation to be fluid, quick and less strenuous.

By default, we used a hierarchical menu layout of breadth 8 and depth 2, thus offering access to 64 commands, to ensure a fair and meaningful comparison with the marking menu (see Section 4.2).

4 EVALUATION

We evaluated the loop and marking menu in a within-subjects experiment with two phases. In Phase 1 we investigated the strengths

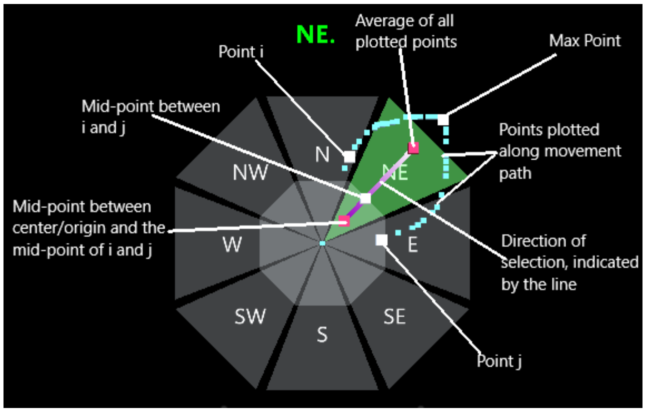

Figure 2: A breakdown of the system's selection mechanism.

and weaknesses of using loops versus marks (linear gestural strokes) for menu selection for continuous and discrete selection separately. Phase 1 therefore compared the continuous and discrete loop menu (with loops only) with the continuous [25] and discrete [46] marking menu (linear strokes only).

In Phase 2 we investigated whether users were able to exploit the flexibility of the loop menu to 1) enhance their selection performance and/or 2) cater for different selection tasks. We investigated if performance with the loop menu would improve when users were allowed the flexibility to combine both loops and linear strokes for menu selection, and if the selection tasks themselves would affect users' preference for any of the various selection methods available.

Phase 1 and Phase 2 were conducted in one sitting with the same participants and apparatus. We investigated the following hypotheses:

Hypothesis 1: Menu selection with loops is more accurate than linear gestural strokes.
Hypothesis 2: Menu selection with loops reduces workload compared to linear gestural strokes.
Hypothesis 3: Menu selection with loops is faster than linear gestural strokes.
Hypothesis 4: Using both loops and linear gestural strokes together improves selection performance.
Hypothesis 5: There will be a difference in preferred selection methods when using the loop menu for different selection tasks.

4.1 Participants and Apparatus

We opportunity-sampled 24 right-handed participants, 9 female and 15 male, in an age range of 22 to 31, with the exception of one 38 year old participant. None had previous experience with marking menus or the HoloLens. All participants took part in both Phase 1 and Phase 2.

Both Phase 1 and Phase 2 were conducted using a Microsoft Hololens, a Kinect 360 and an ASUS X556U laptop computer. The Hololens was used to capture the participants' gestural inputs and the Graphical User Interface (GUI) was viewed from the optical see-through display provided by the HoloLens. We use software by Hincapié-Ramos et al. [19], which uses the Kinect 360 to measure the Consumed Endurance (fatigue) of the participants during the trials. It is important to realize that while the HoloLens is an excellent practical and affordable choice for investigating OST HMD user interfaces, such devices are still in their infancy. As a consequence, the field of view is limited and there is a small lag in the hand tracking.

4.2 Task and Stimuli

Participants were given a sequence of labels (e.g., N-S, which meant first select North and then select South) and were requested to perform the appropriate gestures to generate the respective labels. After the gesture was completed, the participant was presented with a label based on their input. A gesture stroke that was unrecognized would require the participant to re-attempt the same gesture. Due to the limitations of the Hololens, should the participant move his/her hand beyond the device's sensor boundary, the gesture would be cancelled and the participant would need to re-attempt the gesture. When a selection did not match the stimulus, the participant was asked to rectify the mistake(s) until a correct selection is made. An incorrect selection could be undone with a double-click gesture. At the end of each test condition, participants were instructed to complete a NASA-TLX [18] form.

Similarly to previous research on marking menus by [24, 46], we used compass labels ('N', 'S', 'E' and 'W') for the menu layout. This simulates a user who is already well-accustomed to the menu layout and would therefore not require menu visualisations. For reference, a compass showing all directions was also displayed on the top-right corner of the HoloLens display.

Each participant was provided 8 gesture stimuli per trial block. There were two trial blocks for each test condition (menu selection technique). The 8 gesture stimuli and the order were selected randomly. Similarly to prior work [46], the gesture stimuli were non-ambiguous for continuous marking menus (e.g., N-N or NW-NW etc.).

4.3 Statistical Analysis

For all subsequent reported results, unless otherwise stated, we used the following methods for statistical analysis. We used General Linear Model (GLM) repeated measures analysis of variance to analyze Consumed Endurance scores and time durations. Time durations were log-transformed prior to analysis. NASA-TLX ratings were analyzed using Friedman's test. Multiple comparisons were adjusted using Holm-Bonferroni correction from an initial significance level of $\alpha = 0.05$. For some analyses we used a Generalized Linear Mixed-Effects (GLME) model with either a Poisson or Binomial kernel and a log link function as some of our measures technically consisted of count data and thus violated the assumptions of a GLM analysis of variance. The specific kernel used in these analyses is reported for each of the relevant dependent variables. GLME allows us to use more statistical power compared to non-parametric testing. Note that in reporting the statistical outcomes of GLME analyses we report t parameter values, which should not be confused with reporting results from ordinary t-tests. It is not possible to calculate meaningful effect sizes for the GLME models. Instead we report model fit as R^2, which is a proportion of how good the fit is (R^2 = 0: no fit; R^2 = 1: perfect fit). The word *significant* should be interpreted as *statistically significant* throughout the rest of this paper.

4.4 Phase 1: Menu Selection Technique

The objective of Phase 1 was to evaluate the effects of menu selection technique. We investigated two independent variables: 1) DiscreteMenu with two levels (loop menu vs. marking menu); and 2) ContinuousMenu with two levels (loop menu vs. marking menu). These independent variables were investigated as four individual conditions in a single experimental session per participant. To avoid ordering effects, the order of the test conditions was counterbalanced using a Latin Square design.

Prior research shows that in order to ensure a lower than 10% error rate for the marking menu, a menu breadth of 4 can have a maximum depth of 4 (4×4), a menu breadth of 8 can have a maximum depth of 2 (8×2), and any menu breadths exceeding 8 would not be practical accuracy-wise [24, 46]. However, when using the marking menu in an OST HMD, it is cumbersome to keep four continuous linear strokes within the camera field of view/boundary. Hence, we used the 8×2 menu layout for all four test conditions in Phase 1.

Participants did practice trials before each test condition. Participants had to make 7/8 correct selections before they were allowed to proceed with the actual trial. This allows us to study typical user behavior once a completely unfamiliar hypothetical user has familiarized themselves with menu selections in an OST HMD. Each participant went through Phase 1 in one sitting, including breaks between menu selection techniques, in approximately 1 hour.

The experimental design for Phase 1 was as follows (excluding practice trials): 24 participants × 4 test conditions × 8 items for each test condition × 2 trial blocks = 1536 menu selections in total.

The dependent variables were:

(1) **Accuracy:**
   (a) Correct Selection: Measured in number of menu selections that matched the gesture stimuli on the first attempt. A Binomial kernel was used for the GLME analysis.
   (b) Number of gestural attempts. A Poisson kernel was used for the GLME analyses on three of the following variables:
      (i) Exit Boundary Count: Number of re-attempts due to exiting the camera FOV/detection boundary.
      (ii) Repeated Incorrect Selections: Number of re-attempts as a result of repeated wrong selections made.
      (iii) Additional Attempts: Total number of re-attempts to complete all the menu selections.
(2) **Workload:**
   (a) Consumed Endurance: A metric for estimating a user's arm fatigue during mid-air interaction [19]. Hincapié-Ramos et al. [19] defined endurance as "the amount of time a muscle can maintain a given contraction level before needing rest", while Consumed Endurance is the ratio of the interaction time and the computed endurance time associated with the gestures. For instance, a Consumed Endurance level at 100% would mean the user is required to rest.
   (b) NASA-TLX survey results.
(3) **Selection Duration:** The articulation duration of the recorded, correct selection.
(4) **Trial Duration:** The total duration between the time when the trial stimulus is triggered and when the correct corresponding selection is completed.
(5) **Subjective Ratings:** 1–5 Likert scale ratings of the perceived accuracy, selection duration and comfort of the techniques.

4.5 Phase 1 Results

4.5.1 *ContinuousMenu Comparison (Figure 3).*

(1) **Accuracy.** We found a significant difference ($t = 2.3003$, $p < 0.05$, $R^2 = 0.0071$) between the menu selection techniques for the mean number of correct selections. The continuous loop menu was significantly more accurate than the continuous marking menu. We also found a significant difference for the mean number of additional gesture attempts ($t = -7.3039$, $p < 0.05$, $R^2 = 0.6428$) between the menu selection techniques. It took users more attempts on average to complete the 16 menu selections for the continuous marking menu. Out of the total additional attempts, the continuous marking menu also had more attempts due to repeated incorrect selections ($t = -2.9235$, $p < 0.05$, $R^2 = 0.3557$), so-called cascading errors [20]. There was no significant difference for the exit boundary count ($p = 0.11$).
(2) **Workload/Comfort.** There were no significant differences for the overall workload score for the NASA-TLX questionnaire ($p = 0.10$) or the mean Consumed Endurance score ($p = 0.17$) between the menu selection techniques.
(3) **Selection Duration.** There was a significant difference for the mean duration per gesture ($F_{1,23} = 168.95$, $p < 0.05$) between the menu selection techniques. The continuous marking menu was significantly faster than the continuous loop menu.
(4) **Trial Duration.** There was no significant difference ($p = 0.32$) in the mean trial duration between the two menus (6.85 s, sd = 3.48 s and 6.06 s, sd = 2.19 s for the continuous marking and loop menu respectively).
(5) **Subjective Ratings.** There was no significant difference in the median Likert scale ratings of the perceived accuracy ($p = 0.11$), selection duration ($p = 0.83$) and comfort ($p = 0.25$) between the continuous marking and loop menu. The median ratings for accuracy, selection duration and comfort for the continuous marking and loop menu were 4.0, 3.5, 4.0 and 4.0, 3.0, 4.0 respectively.

4.5.2 *DiscreteMenu Comparison (Figure 4).*

(1) **Accuracy.** We found no significant difference between the menu selection techniques for the mean number of correct selections ($p = 1.00$). We found a significant difference for the mean number of additional gesture attempts ($t = -3.0051$, $p < 0.05$, $R^2 = 0.4984$) between the menu selection techniques. It took users more attempts on average to complete the 16 menu selections for the discrete marking menu. There were no significant differences for repeated incorrect selections ($p = 0.35$) or exit boundary count ($p = 0.051$).

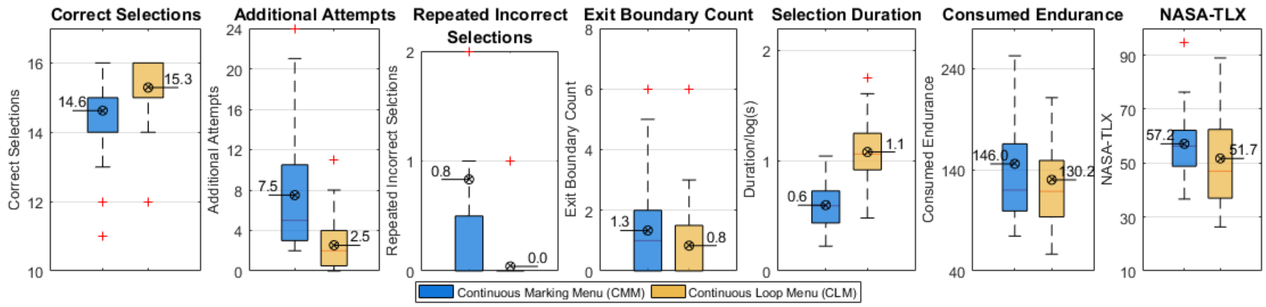

Figure 3: Box-and-whisker plots for continuous marking menu vs. continuous loop menu in Phase 1.

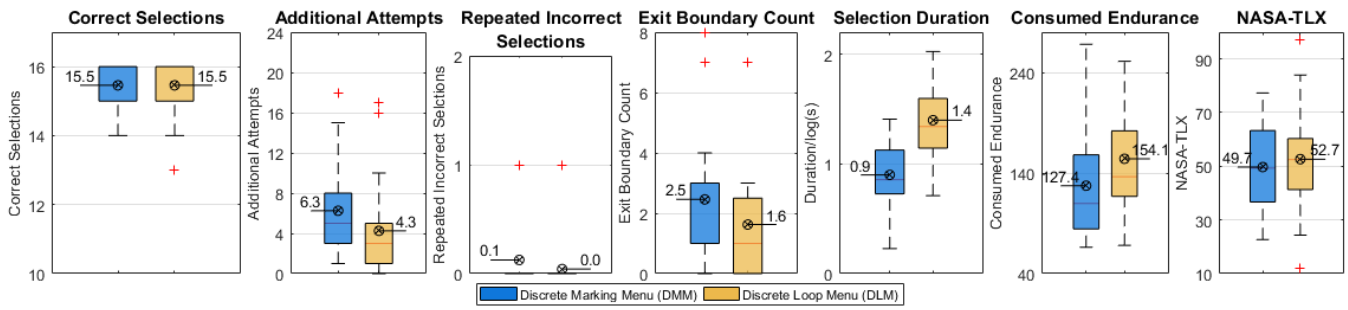

Figure 4: Box-and-whisker plots for discrete marking menu vs. discrete loop menu in Phase 1.

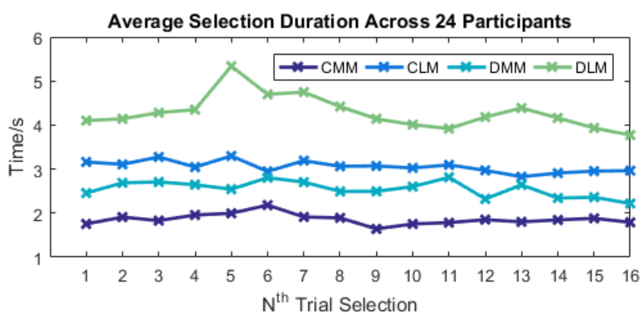

Figure 5: Average selection duration across all participants for trial selection N in each test condition in Phase 1.

(2) **Workload/Comfort.** There were no significant differences for the overall workload score for the NASA-TLX questionnaire ($p = 0.68$) or the mean Consumed Endurance score ($p = 0.07$) between the menu selection techniques.

(3) **Selection Duration.** There was a significant difference for the mean duration per gesture ($F_{1,23} = 90.89$, $p < 0.05$) between the menu selection techniques. The discrete marking menu was faster than the discrete loop menu.

(4) **Trial Duration.** There was a significant difference ($F_{1,23} = 6.03$, $p < 0.05$) in the mean trial duration between the two menus (6.11 s , sd = 1.99 s and 8.16 s, sd = 4.18 s for the discrete marking and loop menu respectively). Participants required less time to complete a trial selection with the discrete marking menu than the discrete loop menu.

(5) **Subjective Ratings.** There was no significant difference in the median Likert scale ratings of the perceived accuracy ($p = 0.62$), selection duration ($p = 0.11$) and comfort ($p = 0.09$) between the discrete marking and loop menu. The median ratings for accuracy, selection duration and comfort for the discrete marking and loop menu are 4.0, 3.5, 4.0 and 4.0, 3.0, 3.0 respectively.

4.6 Phase 2: Hybrid Loop/Marking Menu

In Phase 1 participants were prevented from using the full flexibility of the loop menu as the objective was to assess the merits of loops and linear strokes.

In Phase 2, we first introduced the combined loop and marking menu design which provides full functionality of selecting with both loops and linear marking strokes, in an 8×2 menu layout with compass labels. To avoid confusion with the restricted loop menu configurations in Phase 1, we will refer to this loop menu with full functionality as the *hybrid loop/marking menu.* As there was no learning effect in Phase 1 (see Figure 5), we will compare the performance of this hybrid loop/marking menu against the marking menu results obtained in Phase 1 using the same menu layout.

Thereafter we investigated the effects of altering the menu configuration. We tested two conditions. In the first condition (hybrid loop/marking menu with arbitrary labels) we used the same hybrid menu with an 8×2 menu layout, but instead of compass labels the menu items were labeled in colours for the parent hierarchy and shapes for the child. For example, a menu item could be labeled as red-square or pink-triangle. This simulated a menu layout where the participant would not know where the intended menu items were beforehand, and thus required them to explore the menu first. The second condition (hybrid menu 8×4) used the same hybrid menu but with a 8×4 menu layout and compass labels. This investigated the effect of a more complex, but known, menu configuration.

All participants would proceed with the default hybrid loop and marking menu first, and the following two menu layout conditions were counterbalanced across participants. Each participant went through Phase 2 in one sitting, in about 45 minutes.

The dependent variables were the same as in Phase 1, with the exception that we also investigated the frequency of each selection method (loops vs. linear strokes).

4.7 Phase 2 Results

4.7.1 Overall Comparison. We took the performance data of the hybrid loop/marking menu (HM) and compared with the earlier menu variants in Phase 1: continuous marking and loop menu, discrete marking and loop menu (CMM, CLM, DMM, DLM respectively).

(1) **Accuracy.** We used the same GLME analysis as in Phase 1. We found a significant difference between the menu selection techniques for the mean number of correct selections ($t = -2.3063$, $p < 0.05$, $R^2 = 0.0091$). The hybrid menu was significantly more accurate than the continuous marking menu. We found a significant difference for the mean number of additional gesture attempts between the menu selection techniques. It took users fewer attempts ($R^2 = 0.4048$) on average to complete the 16 menu selections for the hybrid menu than continuous marking menu ($t = 7.4354$, $p < 0.05$), discrete marking menu ($t = 6.0719$, $p < 0.05$) and discrete loop menu ($t = 3.3469$, $p < 0.05$) respectively. Out of the total additional attempts, the hybrid menu also has lesser attempts than the continuous marking menu ($t = 3.0641$, $p < 0.05$) due to repeated incorrect selections ($R^2 = 0.1777$), and lesser attempts than the discrete marking menu ($t = 4.181$, $p < 0.05$) and discrete loop menu ($t = 2.4282$, $p < 0.05$) due to participants exiting the camera boundary ($R^2 = 0.2770$).
(2) **Workload/Comfort.** We found significant differences for the mean workload score for the NASA-TLX questionnaire ($\chi^2 = 11.8326$, $p < 0.05$), and the mean Consumed Endurance score ($F_{4,92} = 4.77$, $p < 0.05$) between the menu selection techniques. The hybrid menu had a significantly lower NASA-TLX workload than the continuous marking menu ($p = 0.0099$). The hybrid loop menu also had a significantly lower Consumed Endurance score than the continuous marking menu ($p = 0.0101$) and the discrete loop menu ($p = 0.0012$).
(3) **Selection Duration.** There was a significant difference for the mean duration per gesture ($F_{4,92} = 83.12$, $p < 0.05$) between the menu selection techniques. The hybrid menu was faster than the continuous loop menu ($p < 0.001$), discrete marking menu ($p < 0.001$) and discrete loop menu ($p < 0.001$).

4.7.2 Importance Rating of Factors. The participants were also asked to rate the importance of accuracy, articulation time and comfort as factors in a freehand gestural menu system. We found a significant difference for ratings between the three factors ($\chi^2 = 11.91$, $p < 0.05$). The participants on average rated accuracy ($p = 0.0337$) and comfort ($p = 0.0029$) as more important than articulation time. There was no significant difference between accuracy and comfort.

4.7.3 Overall Preference. Participants were asked to choose their 1) overall preferred menu between the hybrid menu (HM) and the earlier menu variants in Phase 1: continuous marking and loop menu, discrete marking and loop menu (CMM, CLM, DMM, DLM respectively), and 2) overall preferred menu when the hybrid menu is not an option. 19/24 people preferred the hybrid menu. When the hybrid loop menu was excluded as an option, the majority of the participants preferred the continuous loop menu (11/24) and the discrete marking menu (9/24), both of which are selection methods widely used in the hybrid menu throughout Phase 2.

4.7.4 Selection Pattern. We recorded the corresponding selection pattern for all gesture selections made by the participants. For each loop, straight line and 'click' release (discrete selection mode), we labeled it as 'L', 'S' and 'M' accordingly. For example, if a participant started with a loop, followed by a release of the 'click' gesture, and ended with a straight line, it would be recorded as 'LMS'. We used a GLME model with a Binomial kernel and a log link function to assess the effect of menu layout on the number of gestures performed in discrete selection mode.

We found a significant difference for the number of selections made using the discrete selection mode between the menu layouts. There were more gestures performed with discrete selection mode when using the hybrid menu with arbitrary menu labels than the hybrid menu with compass labels for the 8×2 menu ($t = 9.3116$, $p < 0.05$, $R^2 = 0.8591$). There were also more gestures performed with discrete selection mode when using the hybrid loop menu with the 8×4 menu compared to the hybrid loop menu with the 8×2 menu ($t = 3.0349$, $p < 0.05$, $R^2 = 0.7743$).

5 DISCUSSION

We here discuss the findings in relation to the five hypotheses:

Hypothesis 1: Menu selection with loops is more accurate than linear strokes.

The loop menu resulted in 4.38% more correct selections from a given set of selection stimuli (for continuous selection mode only), and required 21.3% and 8.9% fewer total attempts than the marking menu in both the continuous and the discrete selection mode respectively to complete the selection tasks (see Figure 3 and Figure 4).

While the tendency to exit the camera boundary with the marking menu is higher, the main contributor to the increase in the attempt count was due to participants cancelling incorrectly drawn gestures midway (additional attempts - (repeated incorrect selections + exit boundary count)) before selection confirmation due to the difficulty to produce straight movements with freehand gestures (with 72% and 58.7% of total reattempts for continuous and discrete selection mode respectively; see Figure 3 and 4).

Hypothesis 2: Menu selection with loops causes lower workload than linear marking strokes.

The results do not indicate that the loop menu reduces workload or increases comfort compared to the marking menu. This means

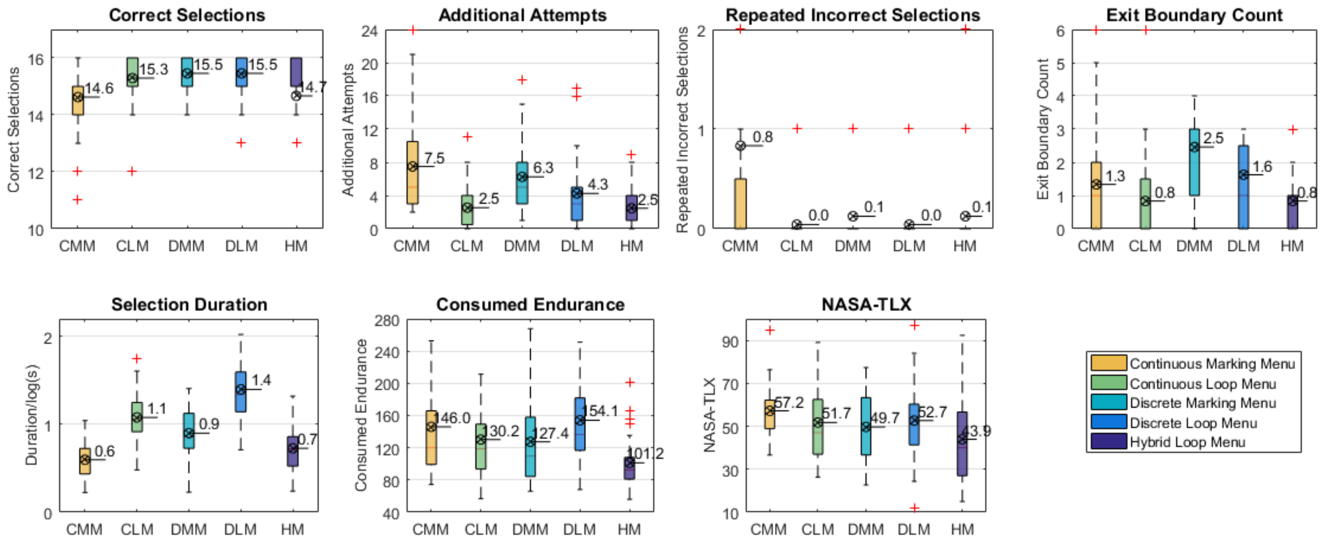

Figure 6: Box-and-whisker plots for for all menu variants in Phase 1 and Phase 2 (using the same menu layout).

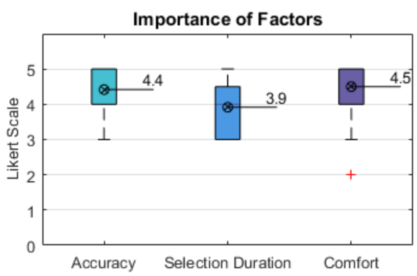

Figure 7: Participants' rating of factors in Phase 2.

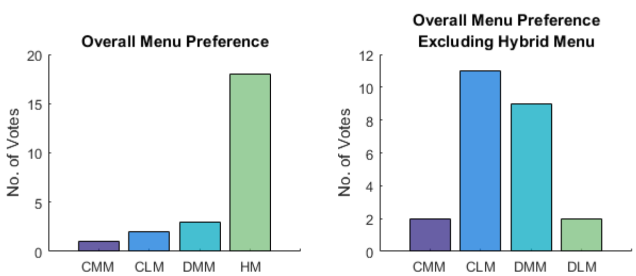

Figure 8: Participants' preferred menu in Phase 2.

using loops instead of straight lines improves accuracy without adding much, if any, additional workload or discomfort.

Hypothesis 3: Menu selection with loops is faster than linear marking strokes.

The average gesture distance per second (i.e., gesture speed) of the marking menu in the continuous and discrete selection mode (0.16m/s and 0.13m/s respectively) was lower than the loop menu (0.21m/s and 0.15m/s for continuous and discrete selections respectively). This indicates that participants were able to move faster with the loop menu compared to when using the marking menu. However, the additional gesture distance required in the loop menu (see Figure 10) outweighed its superior movement speed, resulting in a shorter selection duration for the marking menu than the loop menu in both selection modes (see Figure 3 and 4). Nonetheless, the equal selection duration ratings of the two menus by the participants show that using loops instead of straight lines did not incur any user perceivable penalty in selection duration.

Hypothesis 4: Using both loops and linear marking strokes together will improve selection performance.

The hybrid loop/marking menu outperformed in all measurements (see Figure 6) and was also preferred by most participants. It retained all positive characteristics from each individual menu variant.

In Figure 3 and 4, we can observe a classic speed-accuracy tradeoff between performing straight lines in the marking menu and loops in the loop menu during Phase 1. However, the hybrid menu has both accuracy and speed. It is on par, or better, in terms of the number of correct selections with all the menu variants. It has, together with the continuous loop menu, the lowest additional attempts, the lowest repeated incorrect selection count and the lowest exit boundary count while simultaneously demonstrating approximately the same selection duration as the fastest menu, the continuous marking menu. The hybrid menu is also either on par, or better, than all menu variants in terms of workload/comfort (NASA-TLX and Consumed Endurance; see Figure 6).

To ensure the improved performance with the hybrid menu was not due to learning effects, we compared the average selection duration across all participants for each trial selection for all test conditions in Phase 1. Figure 5 shows that the learning effect is minimal as the average selection duration did not improve significantly from the first trial selection to the last over all test conditions.

Hypothesis 5: There will be a difference in preferred selection methods when using the hybrid menu for different selection tasks.

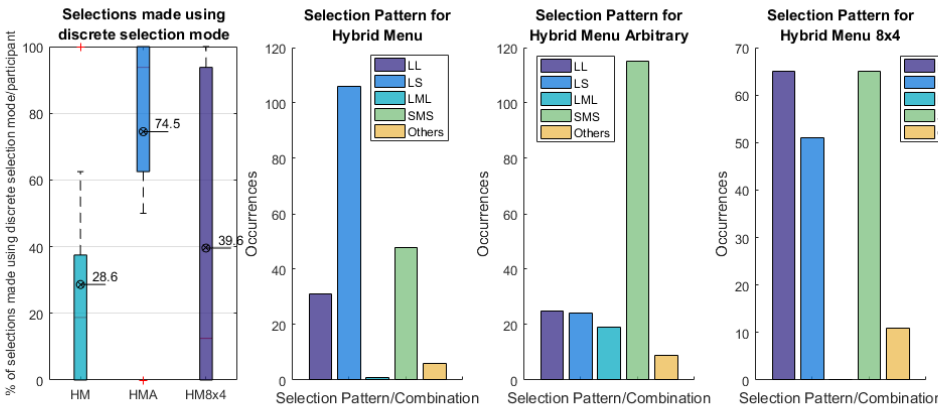

Figure 9: Selection patterns in Phase 2.

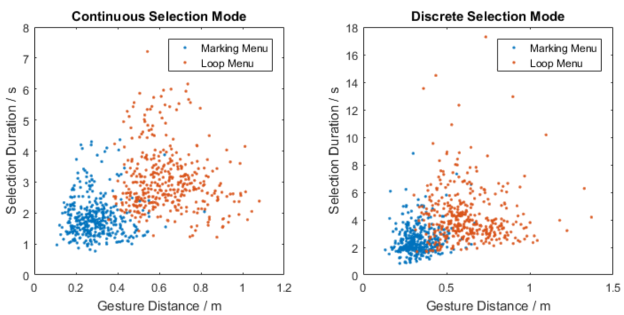

Figure 10: Gesture distance vs. selection duration of every gesture performed by the participants in Phase 1.

The main gesture combinations used throughout Phase 2 were 'SMS'/'SMSMSMS' (analogous to the discrete marking menu) and 'LL'/'LLLL'/ 'LS'/'LLLS (analogous to the continuous loop menu ending with a linear stroke).

There were approximately twice as many gestures performed with discrete selection mode when users were using the hybrid menu with arbitrary menu items as compared to the compass menu items (see Figure 9). We believe the need to search for items in each of the two menu hierarchies encouraged the participants to perceive the task as accomplishing two discrete objectives. This resulted in a pause realized through the release of the 'click' gesture after selecting the menu item from the parent hierarchy.

Participants tended to chunk their gestures more during command elicitation when they already knew the location of the menu items. In both the 8×2 and the 8×4 compass layouts of the hybrid loop menu, the ratio of continuous selection gestures to discrete selection gestures were approximately 3:1 and 2:1 respectively, while the ratio was approximately 1:2 for the arbitrary menu layout (see Figure 9).

With the deeper menu hierarchy in the 8×4 menu layout, there was an increase in variance of the preferred method (see Figure 9). While there were participants who felt that holding the 'click' gesture is less energy consuming, others preferred to release the 'click' gesture more often to rest their arm.

6 CONCLUSIONS

This paper has demonstrated the benefits of using loops over straight lines for making selections on scale-independent menu systems in OST HMDs. Despite the additional distance for gesture travel and the longer gesture articulation duration, the nonrestrictive looping movements improves selection accuracy without compromising users' perceived gesture articulation duration or effort.

Users prioritized accuracy and comfort over speed. Participants' ratings on the importance of the factors in Figure 7 (accuracy, articulation time and comfort) and their preference for menu variants that are more accurate but slower (see Figure 6 and 8), demonstrated that users indeed prioritized accuracy and comfort over articulation time. We conjecture this indicates the importance of providing users with a sense of internal locus of control or *agency* [36], especially for mid-air gestural systems, which involve generally more taxing actions than, for example, capacitive touchscreens.

Users exploited the flexibility of the hybrid menu to enhance their selection performance. Users could make selections with the hybrid menu as quickly as with the marking menu while retaining an accuracy level as high as the original loop menu. Thereby we demonstrated the additional noise resilience provided by the loop menu compared to the marking menu, a key design factor for OST HMDs, which rely on inherently noisy sensing of mid-air gestures.

Finally, users were also able to exploit the flexibility of the hybrid menu to cater for different selection tasks. We observed that participants switched between chunking (continuous gestures) and discrete gestures depending on the selection task as well as on the level of familiarity with the menu layout.

The principle of using loops in gestural interfaces is not new (e.g. [17, 31]). However, this paper demonstrates empirically that looping is a useful technique for designing a practical usable OST HMD menu interface given today's hardware limitations, in particular high lag, lack of precision, and a limited field-of-view, which necessitate a robust and flexible design solution.

REFERENCES

[1] Johnny Accot and Shumin Zhai. 2002. More Than Dotting the I's — Foundations for Crossing-based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 73–80. https://doi.org/10.1145/503376.503390

[2] Ryosuke Aoki, Yutaka Karatsu, Masayuki Ihara, Atsuhiko Maeda, Minoru Kobayashi, and Shingo Kagami. 2011. Gesture identification based on zone entry and axis crossing. *Human-Computer Interaction. Interaction Techniques and Environments* (2011), 194–203.

[3] Gilles Bailly, Eric Lecolinet, and Laurence Nigay. 2007. Wave menus: improving the novice mode of hierarchical marking menus. In *IFIP Conference on Human-Computer Interaction*. Springer, 475–488.

[4] Gilles Bailly, Eric Lecolinet, and Laurence Nigay. 2008. Flower Menus: A New Type of Marking Menu with Large Menu Breadth, Within Groups and Efficient Expert Mode Memorization. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '08)*. ACM, New York, NY, USA, 15–22. https://doi.org/10.1145/1385569.1385575

[5] Gilles Bailly, Eric Lecolinet, and Laurence Nigay. 2017. Visual menu techniques. *ACM Computing Surveys (CSUR)* 49, 4 (2017), 60.

[6] Gilles Bailly, Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2011. Gesture-aware remote controls: guidelines and interaction technique. In *Proceedings of the 13th international conference on multimodal interfaces*. ACM, 263–270.

[7] Olivier Bau and Wendy E. Mackay. 2008. OctoPocus: a dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*. ACM, 37–46.

[8] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Q Ly. 2001. Using pinch gloves for both natural and abstract interaction techniques in virtual environments. (01 2001).

[9] William A. S. Buxton. 1995. Human-computer Interaction. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, Chapter Chunking and Phrasing and the Design of Human-computer Dialogues, 494–499. http://dl.acm.org/citation.cfm?id=212925.212970

[10] J. Callahan, D. Hopkins, M. Weiser, and B. Shneiderman. 1988. An Empirical Comparison of Pie vs. Linear Menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '88)*. ACM, New York, NY, USA, 95–100. https://doi.org/10.1145/57167.57182

[11] H. Cheng, L. Yang, and Z. Liu. 2016. Survey on 3D Hand Gesture Recognition. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 9 (Sept 2016), 1659–1673. https://doi.org/10.1109/TCSVT.2015.2469551

[12] Raimund Dachselt and Anett HÃijbner. 2007. Three-dimensional menus: A survey and taxonomy. *Computers & Graphics* 31, 1 (2007), 53 – 65. https://doi.org/10.1016/j.cag.2006.09.006

[13] William Delamare, Céline Coutrix, and Laurence Nigay. 2015. Designing Guiding Systems for Gesture-based Interaction. In *Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '15)*. ACM, New York, NY, USA, 44–53. https://doi.org/10.1145/2774225.2774847

[14] Assaf Y. Dvorkin, Robert V. Kenyon, and Emily A. Keshner. 2006. Reaching within a dynamic virtual environment. *Journal of NeuroEngineering and Rehabilitation* 4 (2006), 23 – 23.

[15] D Elliott, R Chua, and W.F. Helsen. 2001. A century later: Woodworth's (1899) two-component model of goal-directed aiming. 127 (05 2001), 342–357.

[16] V.H. Franz, K.R. Gegenfurtner, Heinrich BÃijlthoff, and Manfred Fahle. 2000. Grasping Visual Illusions: No Evidence for a Dissociation Between Perception and Action. 11 (02 2000), 20–5.

[17] François Guimbretiére and Terry Winograd. 2000. FlowMenu: combining command, text, and data entry. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*. ACM, 213–216.

[18] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[19] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[20] Clare-Marie Karat, Christine Halverson, Daniel Horn, and John Karat. 1999. Patterns of entry and correction in large vocabulary continuous speech recognition systems. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 568–575.

[21] Kenrick Kin, Björn Hartmann, and Maneesh Agrawala. 2011. Two-handed Marking Menus for Multitouch Devices. *ACM Trans. Comput.-Hum. Interact.* 18, 3, Article 16 (Aug. 2011), 23 pages. https://doi.org/10.1145/1993060.1993066

[22] Per-Ola Kristensson and Shumin Zhai. 2004. SHARK^2: a large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, 43–52.

[23] Arun Kulshreshth and Joseph J. LaViola, Jr. 2014. Exploring the Usefulness of Finger-based 3D Gesture Menu Selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1093–1102. https://doi.org/10.1145/2556288.2557122

[24] Gordon Kurtenbach and William Buxton. 1993. The limits of expert performance using hierarchic marking menus. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. ACM, 482–487.

[25] Gordon Paul Kurtenbach. 1993. *The design and evaluation of marking menus*. University of Toronto Toronto.

[26] Geniva Liu, Romeo Chua, and James T. Enns. 2008. Attention for perception and action: task interference for action planning, but not for online control. *Experimental Brain Research* 185, 4 (01 Mar 2008), 709–717. https://doi.org/10.1007/s00221-007-1196-5

[27] Jennifer Mankoff and Gregory D. Abowd. 1998. Cirrin: A Word-Level Unistroke Keyboard for Pen Input. In *ACM Symposium on User Interface Software and Technology*.

[28] David Milner and Mel Goodale. 2006. *The Visual Brain in Action*. Oxford University Press.

[29] Takashi Nakamura, Shin Takahashi, and Jiro Tanaka. 2008. Double-crossing: A new interaction technique for hand gesture interfaces. In *Asia-Pacific Conference on Computer Human Interaction*. Springer, 292–300.

[30] Tao Ni, Doug A. Bowman, Chris North, and Ryan P. McMahan. 2011. Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. *International Journal of Human-Computer Studies* 69, 9 (Aug. 2011), 551–562. https://doi.org/10.1016/j.ijhcs.2011.05.001

[31] Ken Perlin. 1998. Quikwriting: continuous stylus-based text entry. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*. ACM, 215–216.

[32] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + Pinch Interaction in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. ACM, New York, NY, USA, 99–108. https://doi.org/10.1145/3131277.3132180

[33] Stuart Pook, Eric Lecolinet, Guy Vaysseix, and Emmanuel Barillot. 2000. Control Menus: Excecution and Control in a Single Interactor. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00)*. ACM, New York, NY, USA, 263–264. https://doi.org/10.1145/633292.633446

[34] Gang Ren and Eamonn O'Neill. 2013. 3D selection with freehand gesture. *Computers & Graphics* 37, 3 (May 2013), 101–120. https://doi.org/10.1016/j.cag.2012.12.006

[35] A.W. Salmoni, R.A. Schmidt, and C.B. Walter. 1984. Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin* 95, 3 (1984), 355–386. https://doi.org/10.1037/0033-2909.95.3.355 cited By 717.

[36] Ben Shneiderman. 1997. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[37] Garth Shoemaker, Leah Findlater, Jessica Q. Dawson, and Kellogg S. Booth. 2009. Mid-air Text Input Techniques for Very Large Wall Displays. In *Proceedings of Graphics Interface 2009 (GI '09)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 231–238. http://dl.acm.org/citation.cfm?id=1555880.1555931

[38] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. 2012. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 179–188. https://doi.org/10.1145/2207676.2207702

[39] Dan Venolia and Forrest Neiberg. 1994. T-Cube: A Fast, Self-disclosing Pen-based Alphabet. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. ACM, New York, NY, USA, 265–270. https://doi.org/10.1145/191666.191761

[40] Daniel Vogel and Ravin Balakrishnan. 2005. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST '05)*. ACM, New York, NY, USA, 33–42. https://doi.org/10.1145/1095034.1095041

[41] David A. Whitney, David A. Westwood, and Melvyn A. Goodale. 2003. The influence of visual motion on fast reaching movements to a stationary object. *Nature* 423 (2003), 869–873.

[42] R.S. Woodworth. 1970. The Accuracy of Voluntary Movement. 3 (01 1970), i–114.

[43] Shumin Zhai, Per Kristensson, Caroline Appert, Tue Andersen, and Xiang Cao. 2012. Foundational Issues in Touch-Surface Stroke Gesture Design: An Integrative Review. *Found. Trends Hum.-Comput. Interact.* 5, 2 (Feb. 2012), 97–205. https://doi.org/10.1561/1100000012

[44] Shumin Zhai and Per-Ola Kristensson. 2003. Shorthand writing on stylus keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 97–104.

[45] Shengdong Zhao, Maneesh Agrawala, and Ken Hinckley. 2006. Zone and polygon menus: using relative position to increase the breadth of multi-stroke marking menus. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 1077–1086.

[46] Shengdong Zhao and Ravin Balakrishnan. 2004. Simple vs. compound mark hierarchical marking menus. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, 33–42.

[47] Jingjie Zheng, Xiaojun Bi, Kun Li, Yang Li, and Shumin Zhai. 2018. M3 gesture menu: Design and experimental analyses of marking menus for touchscreen mobile interaction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 249.