Simple vs. Compound Mark Hierarchical Marking Menus

Shengdong Zhao, Ravin Balakrishnan

Department of Computer Science

University of Toronto

sszhao | ravin @dgp.toronto.edu

www.dgp.toronto.edu

ABSTRACT

We present a variant of hierarchical marking menus where items are selected using a series of inflection-free simple marks, rather than the single "zig-zag" compound mark used in the traditional design. Theoretical analysis indicates that this simple mark approach has the potential to significantly increase the number of items in a marking menu that can be selected efficiently and accurately. A user experiment is presented that compares the simple and compound mark techniques. Results show that the simple mark technique allows for significantly more accurate and faster menu selections overall, but most importantly also in menus with a large number of items where performance of the compound mark technique is particularly poor. The simple mark technique also requires significantly less physical input space to perform the selections, making it particularly suitable for small footprint pen-based input devices. Visual design alternatives are also discussed.

Keywords: Marking menus, pie menus, simple mark

INTRODUCTION

Marking menus [6] are a variant of pie or radial menus [4, 12] that allow a user to perform a menu selection by either selecting an item from a popup radial menu, or by making a straight mark in the direction of the desired menu item without popping-up the menu. As with linear menus, marking menus can also support menu hierarchies, where users make “zig-zag” compound marks to select from multiple levels of submenus (Figure 1a). Extensive research [6, 7, 11] has shown that marking menus have significant advantages over regular pie/radial menus and linear menus, including support for seamless transition from novice to expert usage, and selection speeds that are up to 3.5 times faster. While these advantages make marking menus a very promising alternative to status-quo linear menus, they nonetheless have several limitations which can hinder their use in several situations.

In this paper, we first discuss the key limitations of existing compound mark hierarchical marking menus. We propose an alternative design that could overcome these limitations by using a series of simple inflection-free marks instead of a single compound zig-zag mark. We then present an experiment that compares menu selection performance using our new simple mark technique to the traditional compound mark technique.

LIMITATIONS of the COMPOUND MARK TECHNIQUE

Breadth vs. Depth, and Speed vs. Accuracy Trade-off

Kurtenbach and Buxton [6] have shown that as the number of levels (menu depth) and items per level (menu breadth) in the menu hierarchy increases, error rates increase significantly, even for experts. Their results indicate that in order to maintain high selection speed and an acceptable error rate of under 10%, a menu with breadth of four-items per level can be at most four levels deep, while increasing breadth to eight-items per level limits depth to at most two levels. While these breadth vs. depth and speed vs. accuracy tradeoffs may be acceptable for some applications with relatively small numbers of menu items, other applications can require much larger numbers of menu items organized in a complex hierarchy that is both broad and deep. For example, Kurtenbach et al. [8] describe the challenges of deploying marking menus in a commercial graphics application having approximately 1200 menu items, necessitating a hybrid marking and linear menu solution that relinquishes some of the advantages of a pure marking menu approach.

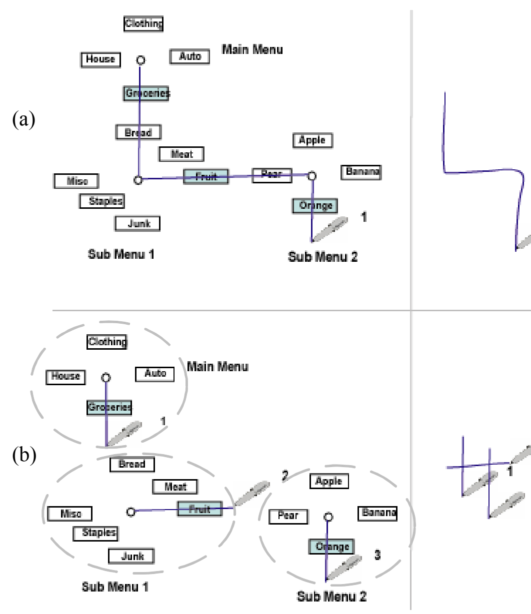

Figure 1. (a) Compound mark technique (b) Simple mark technique. Images on the left show selection from the popup radial menu. Images on right show the same corresponding selection made using the marks alone without popping-up the menu. With simple marks, the marks can overlap.

Ambiguous Marks due to Scale Invariance

When used without popping up the menu, marking menus treat all marks as being scale invariant in that only the changes in directions of the marks are considered when determining which submenu to select when traversing the menu hierarchy. Scale invariance is a key factor contributing to the speed advantages of marking menus, since users only need to make marks in the correct directions without worrying about the size of the marks. However, a compound mark with no inflections is treated as a single mark intended for the leaf node in the given direction, regardless of the depth of the menu hierarchy, leading to possible ambiguities. For example, in the two-level menu layout shown in Figure 2a,b, the system does not distinguish between the mark required for selecting S versus that for selecting S-S. This is not a problem for menus up to two levels deep, since the first level menu simply invokes the second, final, submenu whose leaf nodes are the only selectable items. However, this ambiguity is problematic for menus that are three or more levels deep. For example, in a three level menu, the inability to distinguish between the marks for S-S-N and S-N-N forces a significant number of leaf nodes to be left unassigned since they are not selectable with unique marks (Figure 2c,d). Note that this problem only occurs when the marks are made without the menu being displayed. Appendix A describes how to calculate the number of unambiguous leaf nodes.

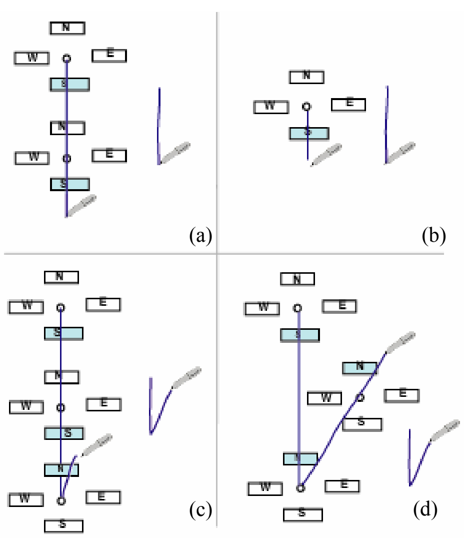

Figure 2. Ambiguity in compound mark technique. The S-S mark in the two-level menu in (a) cannot be distinguished from the S mark in the one level menu in (b). The S-S-N mark in the three-level menu in (c) cannot be distinguished from the S-N-N mark in (d).

Physical Space Requirement

Since marking menus use one continuous compound mark to select items from a menu hierarchy, the amount of physical space required to make the mark grows quadratically as the depth of the hierarchy increases. For example, a four level deep menu requires roughly 16 times more space than a one level menu. (Figure 3) This may not be an issue for desktop computer users with relatively large working areas, but could be a major limitation for users of smaller form factor computing devices. For example, it would be difficult to operate a four level marking menu using a laptop computer's touchpad or a pen on a small screen PDA. While these marks are scale invariant, and the user can ostensibly make smaller scale marks on small footprint devices as the menu depth increases, in practice there's a limit as to how small one can make the marks.

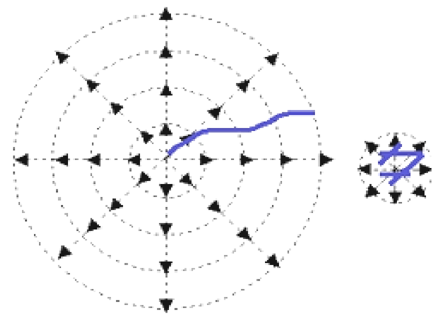

Figure 3. Space requirement. Four-level compound mark hierarchical menu (left) requires roughly sixteen times more space than a four-level simple mark hierarchical menu (right) where all marks are made overlapping one another.

SIMPLE MARK HIERARCHICAL MARKING MENUS

In an effort to alleviate the aforementioned limitations, we have developed a variant of hierarchical marking menus. Instead of using "zig-zag" compound marks, we use a series of inflection-free simple marks performed in quick succession to select from the menu hierarchy (Figure 1b). A simple mark is completed with a pen or mouse-button up event, and the successor mark is initiated with a pen down event. The menu item invoked, whether it's an internal node or a leaf node, solely depends on the menu structure, and is unaffected by the drawing technique. For example, the third simple mark or third segment of a compound mark will similarly invoke a leaf node for a 3-level hierarchic menu, or an internal node for a 4-level hierarchic menu).

A theoretical analysis of this simple mark technique indicates that it retains many of the benefits of the compound mark technique, while having several advantages that could resolve the previously identified three main limitations of the compound mark technique:

Increased depth in menu hierarchy

In terms of the physical actions required, selecting an item in an *n*-level hierarchy using the simple mark technique is effectively equivalent to selecting from *n* single-level marking menus in quick succession. As such, theoretically there should be no limit to how deep the menu hierarchy can go without incurring a performance penalty beyond the simple linear additive cost of selecting from multiple single-level menus. The limits on menu breadth for a single-level menu would still remain, but should be constant as the menu depth increased. As such, the breadth vs. depth tradeoff observed in the compound mark technique should be eliminated.

Unambiguous Marks

As discussed earlier, scale invariance in the compound mark technique results in some non-unique marks and consequently requires that some leaf nodes in menus that are three or more levels deep be left unassigned. As with the compound mark technique, the simple marks are also scale invariant, thus retaining the nice property that users only need to worry about making marks in the correct direction with no regard to scale. However, unlike compound marks, since each simple mark required for selection from each level in the hierarchy is independent of the previous mark, there is no ambiguity as to what items a series of marks is intended to select. As such, all leaf nodes in all levels in the hierarchy can be effectively assigned.

Physical Space Efficiency

Since the simple marks can overlap one another in space, it is possible that the amount of physical space required for menu selection can be significantly reduced. For example, in a 4 level menu hierarchy, the simple mark technique could use only 1/16 of the space required by the compound mark technique (Figure 3). This space efficiency could be advantageous particularly in smaller footprint devices.

EXPERIMENT

Goals

To validate our analysis that using simple marks could be a promising technique, we conducted an experiment that compares the simple and compound mark techniques, with particular emphasis on the following questions:

Q1: How do simple marks compare to compound marks in terms of speed and accuracy?

Previous research [6] has conclusively shown that compound mark hierarchical marking menus are an effective technique, and significantly outperform regular radial and linear menus, albeit with the limitations discussed earlier. Since the simple mark technique can eliminate these limitations, in order to be useful they do not necessarily have to outperform the compound mark technique, but should be at least equivalent in speed and accuracy. If they indeed do outperform the compound mark technique, that would be an added advantage, but arguably not a strictly necessary one.

In terms of the physical motions required, making a series of simple marks is quite different from making a continuous compound mark. For simple marks, no inflections are involved, but multiple pen (or mouse-button) up/down actions are required to delineate the marks. The number of pen up/down actions are linearly proportional to the menu depth. For compound marks, only a single pen up/down action is required regardless of menu depth, but the mark requires several inflections. The trade-off is in the complexity of the mark vs. multiple pen up/down actions. It is unclear how this difference between the two techniques will affect performance. Also, since the simple mark technique allows marks to overlap, the user could make the marks using finger movements alone, while resting the wrist, potentially allowing for better performance [1, 13].

Q2: Do simple marks permit deeper hierarchies?

Our theoretical analysis indicates that using simple marks should allow us to select from an infinitely deep menu hierarchy without penalty beyond the simple linear additive cost of selecting from an additional single-level menu, assuming no change in menu breadth. However, it is entirely possible that this may not be achieved in practice. Of particular interest is whether we can beat the limitation of the compound mark technique where the error rates significantly increases with menus that go beyond two levels with eight items per level (i.e., 8*8=64 items). If using simple marks allows us to transcend this limitation, to say three levels of eight items each (8*8*8=512 items), that would be a very significant advantage.

Q3: Is there a performance difference for certain combinations of mark directions?

Previous research by Kurtenbach and Buxton [6] on compound marks has shown that performance with on-axis marks are significantly better than with off-axis marks. On-axis marks are those in the primary compass directions (N, S, E, W or up, down, right, left), while off-axis marks are those at 45° angles (NE, SE, SW, NW). Further, they showed that two and three level deep combinations consisting of all off-axis directions resulted in additional performance penalties. We explore if similar effects exist with simple marks. Since each level is somewhat independent from the others, we expect that there will not be additional penalties incurred due to the combination of mark directions. In other words, the total performance cost should be the sum of the cost of each individual mark.

Q4: How does device footprint relatively affect performance of compound and simple marks?

One possible advantage of simple marks is that they theoretically require less physical space to execute than compound marks. Indeed, as pen based computing devices get more popular, and often with much smaller physical footprints than traditional mousepads and digitizing tablets, there could be significant value in a menuing technique that performs well in such small scales. We manipulate the size of the input area to explore possible differences between the two techniques in this aspect.

Q5: Are simple marks typically performed by overlapping previous marks in space?

Our reasoning for possible physical space savings relies on users actually making a series of overlapping marks. It is possible, however, that given enough space users will not overlap their simple marks but create the equivalent of a "dashed" compound mark (Figure 1b, image on left)

Q6: What is the timeout threshold required to distinguish inter-command marks from intra-command marks.

With simple marks, since each mark requires a pen up/down action, there are no obvious boundaries between inter-command marks (i.e., making a set of marks to select different items from completely different menus) and intra-command marks (i.e., making a set of marks to select one item from a single hierarchical menu). The number of levels in the menu hierarchy is one way to group the marks, but this approach will fail if the user is interrupted during the marking process. One reliable approach is to introduce a timeout threshold that determines whether a mark belongs to one set or another. We determine an appropriate threshold using our experimental data.

Participants

Twelve right-handed participants, 4 women and 8 men ranging in age from 20 to 35 years, recruited from within the university community, volunteered for the experiment. None had previous experience with marking menus.

Apparatus

The experiment was conducted on a Pentium4 2Ghz workstation running MS WindowsXP, with a 19" display. A pen on a 12" x 18" Wacom Intuous2 tablet was used for input. The pen operated in absolute mode on the tablet. The useable space on the tablet was defined using cardboard cutouts whose size was manipulated as an experimental variable to determine the effect of device footprint. All software was implemented in Java 1.4. While we have attempted to ensure that our implementation of compound mark marking menus is as close to that of Kurtenbach and Buxton [6], it is possible that small differences exist that could potentially bias our comparison in favour of the simple mark technique. However, this issue is moot given the results we obtained, as will be discussed later.

Task and Stimuli

Most of the possible advantages of simple marks occur when users make selections without waiting for the menu to be displayed. This typically occurs once the user is completely familiar with the layout of the menu and knows the exact mark(s) required to select a particular item. Research on compound mark marking menus [5, 6] has shown that users eventually reach this level of expertise, where 90% of selections are made without waiting for the menu to popup. However, it is impractical to expect participants in our experiment to devote sufficient practice to achieve such expert behavior using menus with arbitrarily or realistically labeled items. As such, following Kurtenbach and Buxton [6], we assist in achieving expert behavior by using menus with 4 and 8 items oriented in the standard compass directions and with corresponding labels. The 4 item *compass4* menu has all "on-axis" items labeled "N", "S", "E", and "W", while the 8 item *compass8* menu has an additional four "off-axis" items labeled "NE", "SE", "SW", and "NW". Since participants are typically already familiar with such a compass layout, the overhead in learning the menu layout is significantly reduced, allowing them to make the marks without popping up the menu.

We note that Balakrishnan and Patel [2] in an experiment evaluating a marking menu variant used an alternate approach for eliciting expert behavior. Rather than using the familiar compass layout as in [6], they simply displayed the required mark for the user to emulate, thus completely eliminating any need for familiarity with the menu layout. This approach, however, is unsuitable for our experiment since our comparison of simple and compound marks will necessitate the display of different stimulus (i.e., a compound mark or a series of simple marks) for the two techniques, thus introducing a possible confound.

A pictorial representation of the *compass4* or *compass8* menu layout, as appropriate for given trial, was displayed on the right to assist the user in recalling these layouts. A rectangle was drawn in the screen centre to represent the input area's size, and mapped directly to the constrained area of the digitizing tablet. A small circle was drawn in the middle of the rectangle, and denoted the start position. Instructions to the participant appeared on the top of the screen. Details of the current experimental manipulation were displayed on the left. Figure 4 illustrates the experiment setup. A trial occurred as follows: The participant was first instructed to move their pen (and the corresponding cursor) into the circle. Once the pen cursor dwelled in the circle for half a second, the required menu selection instruction was displayed in red (e.g., "Select N-W"). The participant then responded by making the required marks (compound or simple depending on current experimental condition) to select that menu item. Once the marks were completed, then the resulting menu selection was displayed, in blue if it was a correct selection and in grey if incorrect. The ink-trail of the marks the participant made was also displayed. This served to reinforce learning and aided in correcting errors for subsequent trials. Finally, the participant was instructed to tap on the tablet with the pen to clear the screen and begin the next trial.

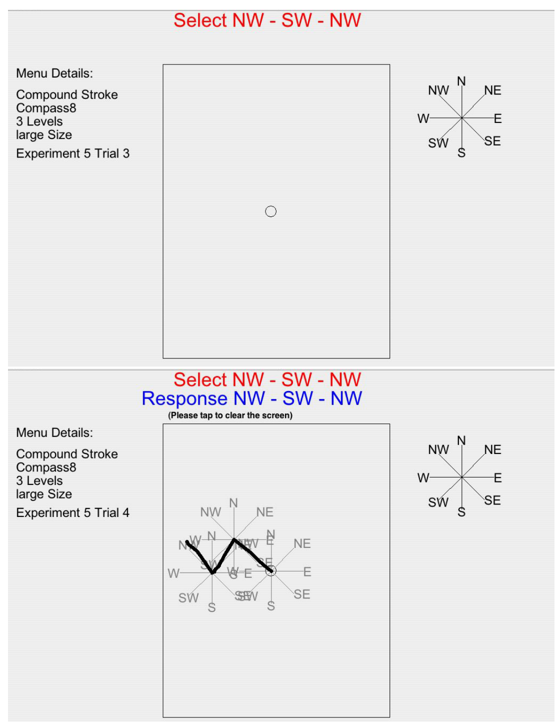

Figure 4. Screenshot of experiment setup. (top) Participant is instructed to select a particular item from the menu. (bottom) After the selection is made, the ink trail is displayed as feedback to the participant.

Design

A within-participants design was used. Participants were randomly assigned to two groups of six participants each. The first group performed the experiment with the compound mark technique first, followed by the simple mark technique. The second group did it in reverse order.

For each technique, participants made selections using three different sizes of input space: *large*, *medium*, and *small*. *Large* was 7.8" x 8.8", approximately the size of a typical mouse pad. *Medium* was 3.5" x 4.25", approximately the size of a typical PDA screen. *Small* was 1.25" x 1.25", resembling the size of a watch or mobile phone screen. Presentation order of the three display sizes was completely counterbalanced across participants and groups (i.e., the first participants in each group did one of the six possible orderings, the second set of participants did the next ordering, … the sixth set of participants did the last possible ordering).

For each technique and size combination, participants made selections from four menus layouts: *compass4-2*, *compass4-3*, *compass8-2*, and *compass8-3*. The first number in these acronyms refers to the number of items in each menu in the hierarchy, and the second number refers to the number of levels in the hierarchy. For example, the *compass4-2* menu has two levels in the hierarchy with each level having four items in a *compass4* layout.

If we were to require participants to make selections from all possible items in these four menu layouts, the experiment would be too large. Thus, following Kurtenbach and Buxton [6], we chose a subset of menu items, equally distributed between on-axis (N, S, E, W) and off-axis (NE, SE, SW, NW) marks as follows:

compass4-2: the only possible marks at both levels are on-axis, and we randomly chose 16 items with replacement (Note since we eliminate menu combinations containing sequential strokes for same directions, the total possible items for compass4-2 is 4*3 = 12 items).

compass4-3: as with *compass4-2,* all possible marks are on-axis, and we randomly chose 16 items.

compass8-2: there are four possible combinations of on- and off- axis marks (two choices at each of the two menu levels): on-on, on-off, off-on, off-off. We randomly chose 8 items from each combination, resulting in 8x4 = 32 items.

compass8-3: there are eight possible combinations of on and off- axis marks (two choices at each of the three menu levels): on-on-on, on-on-off, on-off-on, on-off-off, off-on-on, off-on-off, off-off-on, off-off-off. As with *compass8-2* we randomly chose 8 items from each possible combination, resulting in 8x8 = 64 items.

The order of presentation of the four menu layouts were from easy to hard (i.e., *compass4-2, compass4-3, compass8-2, compass8-3*), to allow for participants to ease gradually into the more complex layouts. Participants could take voluntary breaks between trials, and breaks were enforced between different size conditions. In addition, at the start of each technique, a set of 80 warm-up trials were given at the beginning of the experiment to familiarize the participants with the experiment and the relevant technique. Each participant performed the entire experiment at one sitting, including breaks, in approximately 2 hours. In summary, the design was as follows (excluding warm-ups):

12 participants x
2 techniques (*compound* and *simple*) x
3 sizes (*large, medium, small*) per technique x
(16+16+32+64) items for the four menu layouts
= 9216 menu selections in total.

Results

Accuracy

Accuracy is measured as the percentage of menu selections made that matched the given stimulus.

Analysis of variance showed a significant main effect for technique ($F_{1,11}$ = 131.38, p < .0001), with the simple mark technique having an accuracy of 93%, while the compound mark technique was 80%. Figure 5 illustrates the various effects discussed in this section.

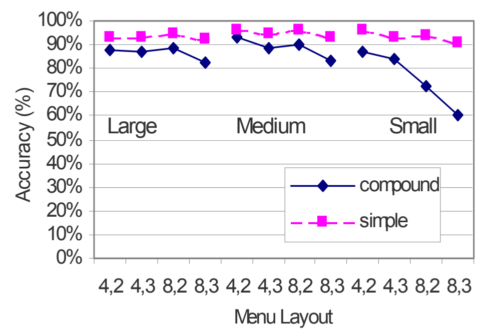

Figure 5. Accuracy by technique, input size, and menu layout.

There was also a significant main effect for input size ($F_{2,22}$ = 30.79, p < .0001). There was also a significant size x technique interaction ($F_{2,22}$ = 21.02, p < .0001), indicating that changes in input size affected the two techniques differently. Pairwise means comparisons (t Tests with 5% alpha-level) indicate that size had no significant effect on accuracy (p > .05) for the simple mark technique. For the compound mark technique, the *medium* and *large* sizes did not significantly differ in their effect on accuracy (p > .05), however, the *small* size resulted in significantly less accuracy than both *medium* and *large* sizes (p < .01).

Both menu layout ($F_{1,11}$ = 42.91, p < .0001) and hierarchy level ($F_{1,11}$ = 39.54, p < .0001) had a significant effect on accuracy. There was also a significant level x technique interaction ($F_{2,22}$ = 4.85, p < .05), with pairwise means comparisons indicating that the compound mark technique with 3 level menus had significantly worse accuracy compared to the other pairs.

There was no significant effect for trial number on accuracy, for all partitions of the data by technique, size, and menu layout (all p > .05). This lack of a significant learning effect indicates that participants were performing at close to expert levels right after the warm-up trials, and that our experiment did likely measure expert performance.

Comparing on-axis and off-axis selections (Figure 6), we found that for compound marks, accuracy was significantly ($F_{2,22}$ = 4.57, p < .05) higher (86.05%) for on-axis selections than for off-axis (75.87%) or mixed on-off axis selections (77.20%). For simple marks, there was no significant difference ($F_{2,22}$ = 2.28, p > .05) between on and off axis selections.

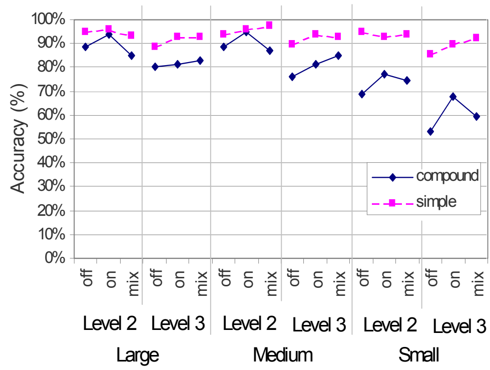

Figure 6. Accuracy by on-axis, off-axis, and mix-axis selections for both techniques, two and three level menus, and different input sizes.

Menu Selection Time

We have three timing measures*:* *Reaction time* for a trial is measured as the time from when the stimulus first appears to the time when the participant begins drawing the mark(s) (i.e., the pen begins to move). This represents the time the participant takes to process the stimulus and figure out what mark to draw. *Drawing time* is the time from when the participant begins drawing the mark(s) to when the menu selection is completed. This can be thought of as the task's motor component when performed by experts. *Total time* is the sum of reaction and drawing times, and represents the complete process of selection.

Analysis of variance showed a significant main effect for technique on total time ($F_{1,11}$ = 5.11, p < 0.05), with average total times of 2.92 seconds for the simple mark technique and 3.09 seconds for the compound mark technique. A similar main effect was seen for technique on drawing time ($F_{1,11}$ = 17.73, p < .0001), with average drawing times of 1.79 and 1.97 seconds for the simple and compound mark techniques respectively. However, there was no significant effect for technique on reaction time ($F_{1,11}$ = 1.36, p > .05).

There was a significant effect for menu layout on total time ($F_{1,11}$ = 107.45, p < .0001), drawing time ($F_{1,11}$ = 138.44, p < .0001), and reaction time ($F_{1,11}$ = 8.42, p < .01). Figures 7-9 illustrate this effect, and those discussed in the rest of this section. Pairwise means comparisons showed that all four layouts resulted in significantly different total and drawing times (p < .01), with the two deeper menus (*compass4-3* and *compass8-3*) requiring more time than the shallow menus (*compass4-2* and *compass8-2*). This is not surprising since the deeper menus require one additional mark segment. With regards to reaction time, the *compass8* layouts required more time than the *compass4* layouts.

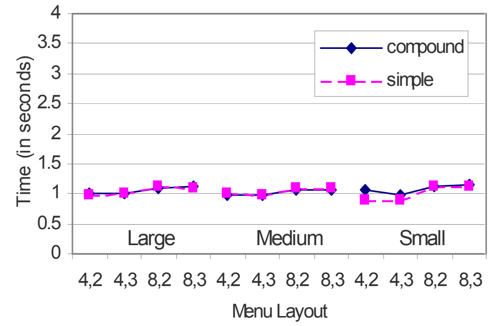

Figure 7. Reaction time by technique, input size, and menu layout.

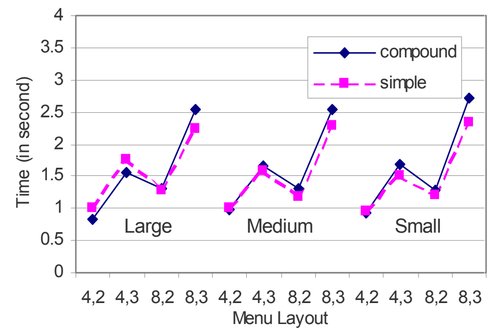

Figure 8. Drawing time by technique, input size, and menu layout.

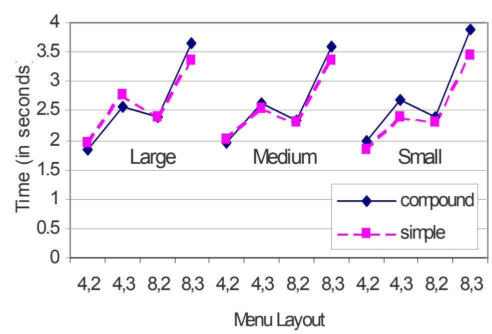

Figure 9. Total time by technique, input size, and menu layout.

There was no significant effect for input size on total time ($F_{2,22}$ = 2.86, p > .05) nor drawing time ($F_{2,22}$ = 0.22, p > .05). However, there was a significant effect for input size on reaction time ($F_{2,22}$ = 6.60, p < .01), which was somewhat surprising since we might expect size to effect motor performance more than the task's pre-planning stage.

There was a technique x size interaction for total time ($F_{2,22}$ = 5.57, p < .01) and drawing time ($F_{2,22}$ = 9.27, p < .0001), with pairwise means comparisons indicating that the compound mark technique with small size being significantly different from other pairs. There was also a menu layout x technique interaction on drawing time ($F_{2,22}$ = 6.12, $p < .05$), with the *compass8-3* layout for compound marks being the worst performer.

There was a significant size x menu layout interaction for total time ($F_{2,22}$ = 5.25, p = .0052), but not for reaction or drawing times. Pairwise means comparisons showed that this effect was due to the *compass4-3* and *compass8-3* layouts in *small* size resulting in significantly different times from the same layouts in the other sizes for the compound mark technique ($p < .01$). None of the other pairs were significantly different ($p > .05$).

Comparing on-axis and off-axis selections (Figures 10-12), we found that drawing time was significantly different for compound marks ($F_{2,22}$ = 247.54, $p < .0001$) and simple marks ($F_{2,22}$ = 100.48, $p < .0001$) for on-axis, off-axis, and mixed-axes selections. Similar effects were seen for reaction time with compound marks ($F_{2,22}$ = 19.24, $p < .0001$) and simple marks ($F_{2,22}$ = 6.41, p = .0017); and for total time with compound marks ($F_{2,22}$ = 207.20, $p < .0001$) and simple marks ($F_{2,22}$ = 63.27, $p < .0001$). Figures 10-12 illustrate these effects.

There was no significant effect for trial number on either total or drawing time, for all partitions of the data by technique, size, and menu layout (all $p > .05$). This reinforces the similar finding for the accuracy measure, and provides further evidence that the experiment measured expert performance.

Physical Space Usage

In order to compare the amount of physical space used for making the marks in the different techniques and conditions, we computed the areas of the bounding box surrounding the mark(s) for each selection. Figure 13 illustrates the space usage for all participants, with the data for participant 1 shown in greater detail simply as an example case. There was a significant main effect for technique on space usage with the compound mark technique taking significantly more space than the simple mark technique ($F_{1,11}$ = 14.90, $p < .001$). Input size significantly affected space usage for compound marks ($F_{2,22}$ = 220.95, $p < .0001$). As expected from our theoretical analysis, participants took advantage of the larger input sizes to make larger marks. However, this effect was far less pronounced with simple marks, indicating that participants were making more overlapping marks and using space more efficiently. In other words, the simple mark technique was not regressing to a simple "dashed" mark version of the compound mark technique. Menu layout also had a significant effect on space usage ($F_{1,11}$ = 23.29, $p < .0001$) for both techniques.

Subjective Preference

Participants completed a post-experiment questionnaire in which they were asked to rate their preferences with respect to technique and input size. Of the twelve participants, three preferred the compound mark technique and nine preferred the simple mark technique. Note that none of the participants had any prior experience with either technique, and thus had no idea beforehand which technique was the status-quo and which was the new technique. Therefore it is unlikely that this preference is due to the "good participant effect" often observed in experiments where participants will tell the experimenter that they prefer the technique which they know the experiment is trying to evaluate.

For those who preferred the simple mark technique, one preferred the large size, three preferred medium, one preferred small, and four had no preference. For those who preferred the compound mark technique, two preferred the large size and one preferred medium.

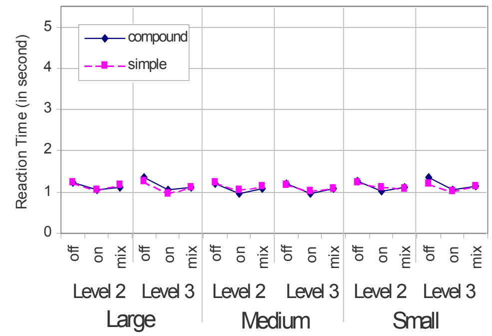

Figure 10. Reaction time by on-axis, off-axis, and mix-axes selections for both techniques, two and three level menus, and different input sizes.

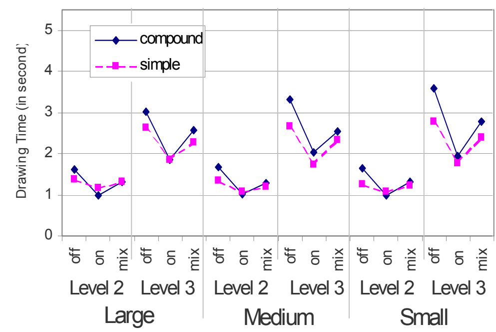

Figure 11. Drawing time by on-axis, off-axis, and mix-axes selections for both techniques, two and three level menus, and different input sizes.

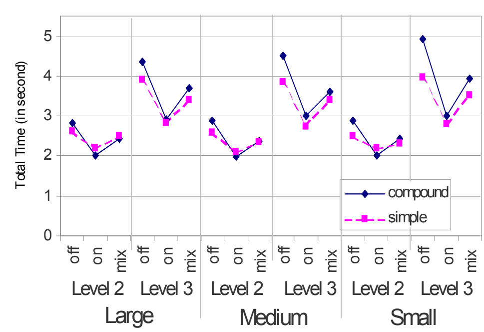

Figure 12. Total time by on-axis, off-axis, and mix-axes selections for both techniques, two and three level menus, and different input sizes.

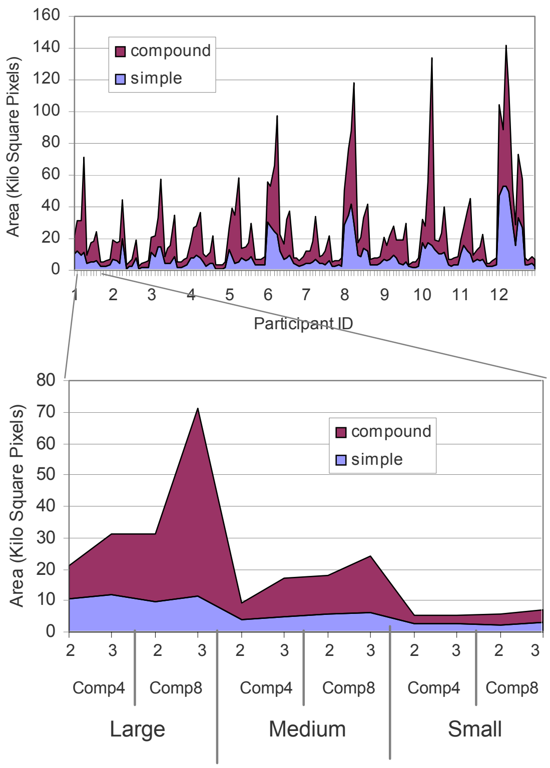

Figure 13. Physical space usage for both techniques, broken down by participants. The bottom graph shows the data for participant #1, and illustrates the coding of the top graph

Discussion

When describing the apparatus used in our experiment, we noted that possible slight differences in our implementation of compound mark marking menus could bias the results in favor of the simple mark technique. However, our results for the compound mark technique are comparable to that of [6], and the results for the simple mark technique clearly outperform the results in [6], so any issue of bias is moot. In light of the results of our experiment, we can now revisit and attempt to answer the questions posed earlier:

Q1: How do simple marks compare to compound marks in terms of speed and accuracy?

Our results have clearly shown that the simple mark technique enables slightly faster menu selections (~6%) than the compound mark technique. More importantly, however, is the significant difference in accuracy: 93% for simple marks overall versus 80% for compound marks.

Q2: Do simple marks permit deeper hierarchies?

From Kurtenbach and Buxton [6] we know that error rates significantly increase when compound mark menus go beyond two levels with eight items per level. Our experimental data also showed a similar trend, with the *compass8-3* layout resulting in significantly higher error rates compared to the narrower and shallower layouts. With the simple marks technique, however, accuracy remained approximately constant at about 93% for the four layouts tested: *compass4-2, compass4-3, compass8-2, compass8-3.* While our experiment did not explore hierarchies deeper than three levels, the excellent performance with regards to the *compass8-3* layout indicates that the simple mark technique can at least overcome the *compass8-2* limit of the compound mark technique by one additional level. This alone can increase the number of viable menu items from the previously established [6] maximum of 4*(3^3) + 4 = 112 for a *compass4-4* layout with compound marks to 8*8*8 = 512 for a *compass8-3* layout with simple marks, which is a significant improvement. Note that the number of items for the *compass4-4* layout with compound marks is not 4^4 = 256 due to the ambiguity problem discussed earlier (also see Appendix A). In terms of selection time, the simple mark technique was also faster than the compound mark technique for deeper hierarchies, with differences ranging approximately from 5% for the *compass4-3* layout to 14% for the *compass8-3* layout. Overall, the results show that the simple mark technique can significantly reduce, if not eliminate, the breadth vs. depth and speed vs. accuracy tradeoffs observed in the compound mark technique [6].

Q3: Is there a performance difference for certain combinations of mark directions?

Our results showed that all combinations of mark directions resulted in similar selection accuracy with the simple mark technique. In contrast, consistent with the previous research [6], the compound mark technique showed significant differences, with on-axis marks performed more accurately than off-axis marks. This is likely due to the interdependence between levels in the compound mark technique. However, the effect of mark direction on selection times were similar for both techniques.

Q4: How does device footprint relatively affect performance of compound and simple marks?

Our results showed that changes in device or input size ranging from large 7.8’ x 8.8’ to small 1.25’ x 1.25’ had no effect on both accuracy and selection time for the simple mark technique. However, for the compound mark technique, both accuracy and selection time significantly degraded at the smallest size tested. These results clearly indicate that moving from compound to simple marks can make marking menus a viable technique for very small footprint devices like watches and small PDAs.

Q5: Are simple marks typically performed by overlapping previous marks in space?

From our observations during the experiment, participants tended not make simple marks in an optimally overlapping fashion. Rather, the marks were slightly offset, but not to the extent of making a “dashed” compound mark. Our bounding box analysis confirmed that the space usage in the simple mark technique was significantly more efficient than the compound mark technique. On average, the true space usage with simple marks was roughly halfway between the theoretical lower bound where all marks are made exactly overlapping, and the upper bound space usage of the compound mark technique

Q6: What is the timeout threshold required to distinguish inter-command marks from intra-command marks.

In our experiment, we deliberately did not impose a timeout threshold between marks in the simple mark technique. This allowed us to observe what participants would do if given no predetermined limits on how quickly they had to make the successive marks. This experimental design allowed us to empirically determine an appropriate threshold to use in discriminating between inter-command and intra-command marks for future menu designs.

Across all the trials for the simple mark technique, we found that participants on average took 0.481 seconds between marks. The standard deviation was 0.377 seconds, and the median was 0.375 seconds. We also found that 99% of the trials took less than 1.875 seconds between marks, and 95% of the trials tool less than 1.156 seconds. In real use, we can expect that users will get even more familiar with the technique than they did in our experiment. Thus, it seems reasonable to treat these results as a worst case bound on the desired threshold. Accordingly, we estimate that a real simple mark design should have a threshold somewhere between 1-2 seconds. In other words, if a user pauses for longer than this threshold between successive marks, the next mark is considered the first mark in a new menu selection.

DESIGN ALTERNATIVES for SIMPLE MARK HIERARCHICAL MARKING MENUS

Now that our experiment has shown the simple mark technique to be a viable contender to traditional compound mark hierarchical marking menus, it is worth exploring a few details with regards to the visual design of a simple mark hierarchical marking menu.

When the menu is popped up in compound mark hierarchical marking menus, each sublevel is displayed in a new location, in the standard stacked fashion shown in Figure 1a. The previous level submenu can either remain on screen or fade away. Either design does not change the amount of screen space used. For simple marks, however, if we are to try to optimize screen space usage, it is worthwhile considering alternative layouts. As Figure 14 (top) shows, one option is to simply fade away the previous level menu from which a selection has just been made, and display the next level menu in its place, or just slightly offset. This design has the advantage of being very space efficient, but the disadvantage of perhaps not retaining the previous context for the user's reference. Figure 14 (bottom) shows an alternate design, where previous level is displayed in a faded hue, but is moved away from its original location, while the next level menu appears in its place. This allows the user to minimize the amount of movement required to make successive marks, while retaining the context of the previous menu selection. If only the immediate previous level menu is displayed faded, screen space usage is still kept fairly minimal.

In all cases, just like with regular compound mark marking menus, the menu only pops up if the user dwells in place, at any level of the menu hierarchy, for a short time,. If the user makes a mark right away, then the menu does not pop up, and the user makes the selection using marks alone. If successive marks are made within the defined time threshold, they are treated as making selections from a single multi-level menu. If the time between marks exceeds the threshold, then it is assumed that the previous selection is to be aborted and a new one begun. Besides exceeding the time threshold, one can also abort the menu by drawing a special gesture (such as a circle or a pigtail).

Another design question is how the simple mark technique should support backing up a level in the hierarchy. In the compound mark technique, if the menu displayed one just has to roughly retrace the compound mark to the desired menu level. With simple marks, however, once a level has been traversed with a mark, one cannot retrace that mark since the pen has already been lifted and the next mark likely begun. One solution is to simply support backing up levels by clicking on the middle circle of the menu. Another, perhaps less desirable, option is to reserve one menu item in each level as the "backup" command.

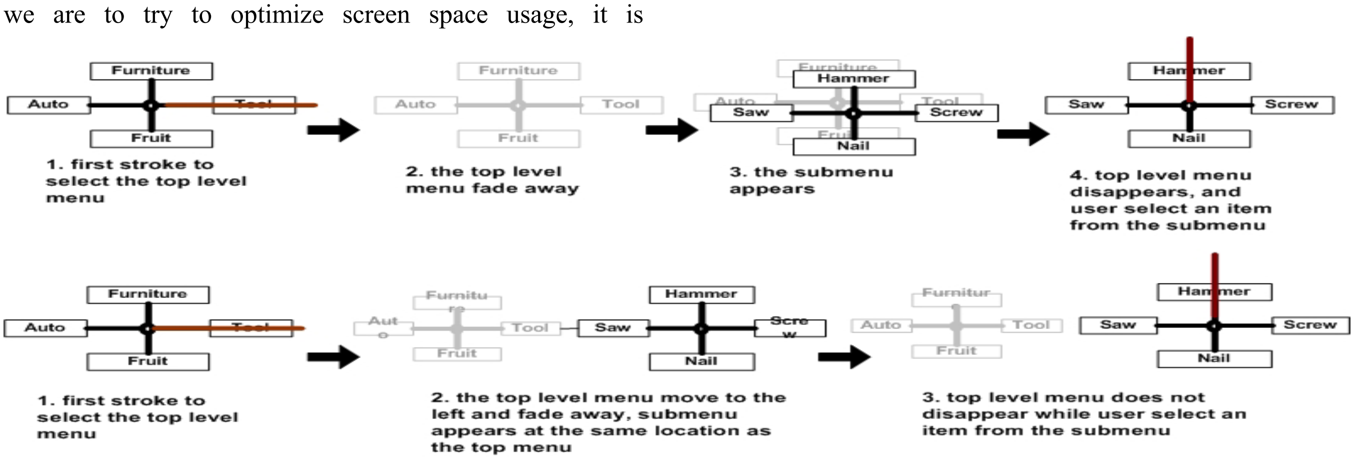

Figure 14. Alternate designs for displaying menus for the simple mark technique. (top) After a selection is made from one level of the menu, the menu for that level fades away, and is replaced with the next level menu. (bottom) The previous level menu is pushed aside and faded out, while the next level is displayed in its place.

CONCLUSIONS and FUTURE WORK

We have presented a variant of hierarchical marking menus that use a series of simple inflection-free marks to make selections, instead of the traditional single compound mark. An experiment comparing the two techniques showed that using simple marks outperformed compound marks in both accuracy and selection speed for four different menu layouts and three difference input sizes. In particular, we showed that the simple mark technique can effectively select from a 512 item menu (*compass8-3*) without degradation in selection accuracy. This significantly increases the viable menu size for marking menus by a factor of five, from the previously established effective limit of 112 items (*compass4-4*) [6]. In addition, the data shows that performance with the simple mark technique remained fairly consistent as input or device size decreases, making it particularly suitable for small footprint devices.

While our work has explored many of the issues surrounding the simple mark technique, one issue that remains unanswered is whether users will find it easier or harder to remember a series of simple marks compared to a single compound mark. In order to make selections without displaying the menu, users have to memorize the marks required for each level of the menu. With compound marks, this task may be chunked [3, 10] as a single whole, whereas it is possible that simple marks will result in users thinking of the task as several subtasks, thus hindering their transition from novice (i.e., selecting with menu displayed) to expert (i.e., making marks without menu displayed) performance. Also, although our results do not indicate it, there is a small possibility that the simple mark technique's serial pen up/down events could lead to more mode errors [9] in practice than the compound mark technique where the single pen/up down event per mark essentially acts like a kinesthetically held mode while selecting from multiple levels. These issues along with a real world implementation of the simple mark technique should be explored further.

ACKNOWLEDGEMENTS

We thank Mark Chignell, Jingnan Yang, Xiao Wu, Michael McGuffin, our experiment participants, and members of the Dynamic Graphics Project lab (www.dgp.toronto.edu) at the University of Toronto.

Appendix A: Calculation of Unambiguous Leaf Nodes in Compound Mark Marking Menus

Let B be the branching factor of the menu (e.g., 4, 8)

Let D be the depth of the menu (i.e., number of levels)

Then, the total number of leaf nodes = B^D

Number of leaf nodes with unambiguous marks =
(number of marks with maximal number D-1 inflections) +
(number of marks with no inflections at all) =

B*(B-1)^(D-1) + B

Example calculations:
compass8-2 layout = 8*(7^1) + 8 = 64 (i.e., all leaves)
compass4-4 layout = 4*(3^3) + 4 = 112 (43% of all leaves)
compass8-3 layout = 8*(7^2) + 8 = 400 (78% of all leaves)

REFERENCES

1. Balakrishnan, R., & MacKenzie, I.S. (1997). Performance differences in the fingers, wrist, and forearm in computer input control. *ACM CHI Conference on Human Factors in Computing Systems*. p. 303-310.
2. Balakrishnan, R., & Patel, P. (1998). The PadMouse: Facilitating selection and spatial positioning for the non-dominant hand. *ACM CHI Conference on Human Factors in Computing Systems*. p. 9-16.
3. Buxton, W., Chunking and phrasing and the design of human-computer dialogues, in *Readings in human-computer interaction: Towards the year 2000*, R. Baecker, J. *Grudin, W. Buxton, S. Greenberg.*, Editors. 1986, Morgan Kaufmann. p. 494-499.
4. Callahan, J., Hopkins, D., Weiser, M., & Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. *ACM CHI Conference on Human Factors in Computing Systems*. p. 95-100.
5. Kurtenbach, G. (1993). The design and evaluation of marking menus. PhD Thesis thesis. University of Toronto.
6. Kurtenbach, G., & Buxton, W. (1993). The limits of expert performance using hierarchical marking menus. *ACM CHI Conference on Human Factors in Computing Systems*. p. 35-42.
7. Kurtenbach, G., & Buxton, W. (1994). User learning and performance with marking menus. *ACM CHI Conference on Human Factors in Computing Systems*. p. 258-264.
8. Kurtenbach, G., Fitzmaurice, G.W., Owen, R.N., & Baudel, T. (1999). The Hotbox: efficient access to a large number of menu-items. *ACM CHI Conference on Human Factors in Computing Systems*. p. 231-237.
9. Sellen, A., Kurtenbach, G., & Buxton, W. (1992). The prevention of mode errors through sensory feedback. *Human Computer Interaction*. 7(2). p. 141-164.
10. Simon, H. (1974). How big is a chunk? *Science*. 183. p. 482-488.
11. Tapia, M., & Kurtenbach, G. (1995). Some design refinements and principles on the appearance and behavior of marking menus. *ACM UIST Symposium on User Interface Software and Technology*. p. 189-195.
12. Wiseman, N., Lemke, H., & Hiles, J. (1969). PIXIE: A new approach to graphical man-machine communication. *Proceedings of the 1969 CAD Conference*.
13. Zhai, S., Milgram, P., & Buxton, W. (1996). The influence of muscle groups on performance of multiple degree-of-freedom input. *ACM CHI Conference on Human Factors in Computing Systems*. p. 308-315.