A Multiscale Interaction Technique for Large, High-Resolution Displays

Sarah M. Peck* Chris North† Doug Bowman‡
Virginia Tech

ABSTRACT

This paper explores the link between users' physical navigation, specifically their distance from their current object(s) of focus, and their interaction scale. We define a new 3D interaction technique, called multiscale interaction, which links users' scale of perception and their scale of interaction. The technique exploits users' physical navigation in the 3D space in front of a large high-resolution display, using it to explicitly control scale of interaction, in addition to scale of perception. Other interaction techniques for large displays have not previously considered physical navigation to this degree. We identify the design space of the technique, which other researchers can continue to explore and build on, and evaluate one implementation of multiscale interaction to begin to quantify the benefits of the technique. We show evidence of a natural psychological link between scale of perception and scale of interaction and that exploiting it as an explicit control in the user interface can be beneficial to users in problem solving tasks. In addition, we show that designing against this philosophy can be detrimental.

KEYWORDS: interaction technique, large tiled display, physical navigation

INDEX TERMS: H.5.2 User Interfaces - Interaction styles

1 INTRODUCTION

The decreasing price of displays has enabled the exploration of ever-larger and higher-resolution displays. Previous research has quantified benefits from both the increased size and the increased resolution. Several studies have shown that with large datasets, such as those found in geospatial analysis, the larger viewport size improves users' performance time and decreases frustration [1,14]. A key benefit of large high-resolution displays is that they afford greater opportunity for physical navigation (moving one's body to navigate the displayed information). These studies found a correlation between faster user performance time, a decrease in virtual navigation and an increase in physical navigation. Evidence indicated that physical navigation was more efficient, effective, and preferred than virtual navigation.

While increasing display size has significant benefits for user performance, it also creates a new difficulty – how do users interact with it? In particular, given that users physically navigate when using large display information spaces, how should that affect the design of interaction techniques? Stationary interaction devices, such as the traditional keyboard and mouse, can tether users and discourage physical navigation. Considering 3D input (such as tracking head, hand or body movement) for interaction with large 2D displays offers new possibilities for interface design that supports physical navigation [12].

However, when physically navigating, users move in and out from the display to zoom into details or out for an overview, essentially *changing the scale* at which they perceive the information. Untethered interaction techniques can enable users to interact with the display from anywhere as they move around in the space in front of the display, but this interaction is static.

During physical navigation, *users' visual scale changes, but their scale of interaction does not change.* This is a problem because users often do different types of tasks at different levels of visual scale. As they step out to see an overview, they are still interacting with it on the detail level, even though they are no longer able to see any of the details. Introducing additional controls that enable users to change the scale of interaction can cause extra confusion, visual and interactive clutter, and difficulties with accuracy, analogous to virtual navigation.

There is evidence that people naturally interact on different levels of scale in the literature on pointing. People can indicate a single nearby object by pointing alone. However, people either refer to faraway areas only, or compensate for the ambiguity brought on by distance by adding verbal descriptions [8,11]. This change in how people point to a target based on their distance from it can be modeled approximately by a cone extending out from the finger, representing the degree of pointing accuracy with which people perceive they are pointing [8].

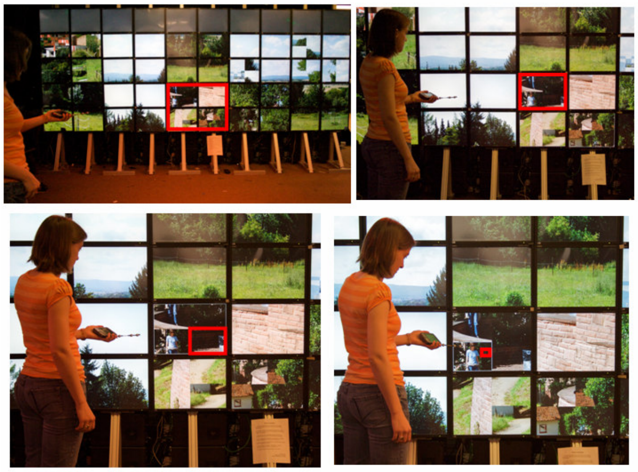

Figure 1. Selecting on different scales with multiscale interaction.

Thus, two related behaviors have been demonstrated: people move to perceive varying levels of scale, and people point at varying degrees of detail according to distance. *Is it possible to combine these behaviors to provide an efficient and explicit multiscale pointing capability on large displays?* Can physical navigation be exploited, using it as an explicit operator to control the scale level of interaction, in addition to perception?

We suggest a class of interaction techniques, called multiscale interaction, which links the two behaviors by *changing the user's scale of interaction depending on their distance from the current object(s) of interaction.* To demonstrate this, we implemented a 3D interaction technique [2] that automatically changes the scale of a 2D cursor according to the user's distance from the display, using a 3D input device to interact directly with a large 2D information space on the display, as seen in Figure 1. This paper first describes the design space for multiscale interaction, and the benefits we believe can be derived from it. Second, it discusses a

---

{*speck, †north, ‡dbowman}@vt.edu

IEEE Symposium on 3D User Interfaces 2009
14-15 March, Lafayette, Louisiana, USA
978-1-4244-3965-2/09/$25.00 ©2009 IEEE

user study to evaluate an implementation of multiscale interaction, and the consequent results.

2 RELATED WORK

Much work has already been done developing pointing and selection techniques for large display environments, both virtual and not. These techniques can be classified by the degree to which they incorporate physical navigation: those that ignore it, those that support it, and those that exploit it. Multiscale interaction falls into the last category.

2.1 Ignoring Physical Navigation

Direct pointing is a cognitively direct interaction technique which uses ray-casting to point to and select objects. Jiang, et al's Direct Pointer was designed for situations when only remote interaction with large displays is possible [7]. This technique seeks to enable interaction from afar while completely discounting close interaction. It uses simple handheld digital cameras, such as those found in cellphones to do direct, laser pointer-like interaction. The Direct Pointer indicates cursor position with a red circle of diameter 48 pixels. Another technique seeks to refine typical seated mouse interaction to better fit the scale of large displays used for personal workspaces [4].

2.2 Supporting Physical Navigation

Other direct pointing techniques, however, seek to enable interaction from multiple locations in the display space. The VisionWand technique passively tracks a wand in 3D space to interact with the display [3]. The technique focuses less on direct pointing and more on enabling coarser interactions from anywhere in the camera-tracked space. These coarser interactions include a set of wand gestures for specific commands. Several of the gestures incorporate movement in relation to the display, but only small arm movements as part of the gesture. The Vision-Wand technique provides visual feedback: two differently colored circles indicating the orthogonal position of the wand markers on the screen, as well as a set of black crosshairs indicating the intersection point of the ray.

Vogel and Balakrishnan describe a set of pointing and clicking techniques for large, high-resolution displays that use the human hand as the implement [15]. The authors emphasize the need for mobile interaction techniques that can transition smoothly from close to far interaction. The techniques involve tracking hand gestures for both direct and relative pointing, and several clicking techniques. The cursor is indicated by the normal arrow, and is augmented with additional visualizations to compensate for lack of tangible feedback from interactions such as clicking.

Liang and Green first describe a 3D selection technique that extends the ray-casting concept, called the "spotlight" technique [9,10]. This technique shoots a cone out of a pointing implement, with the apex of the cone at the end of the pointer. Objects that fall inside the cone are candidates for selection. Only one object may be selected at a time; object selection is disambiguated by selecting the one closest to the origin of the cone. The selection area is visualized by a transparent cone, with the intensity of light indicating closeness to origin.

Aperture based selection is based on the "spotlight" technique, and is designed to alleviate problems associated with aiming and selection of an object from a distance [7]. Much like the original "spotlight" technique, things inside the cone are candidates for selection; however, the user is additionally able to control the size of the selection cone by moving the pointer away from/towards a fixed point called the eye. Again, the technique is intended only to select one object at a time. Selection area is shown using the aperture cursor, reducing the clutter from visual feedback about what is selected.

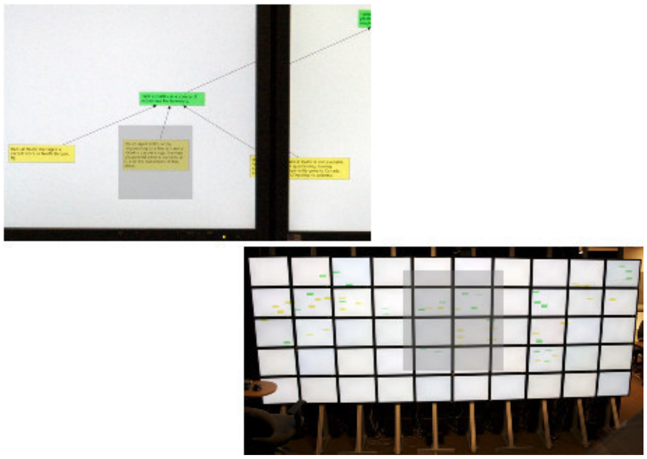

Figure 2. Selecting on the (a) detail and (b) overview scales.

2.3 Exploiting Physical Navigation

There is an opportunity space here to exploit users' physical navigation for more meaningful interactive purposes, by recognizing the semantics of the physical navigation space. Multiscale interaction is singular in that it exploits users' physical navigation to control scale of interaction with large 2D data displays using 3D interaction techniques.

3 DESIGN SPACE FOR MULTISCALE INTERACTION

The design space focuses on applying multiscale interaction concept to selection and navigation; to the former by defining *what* gets selected and to the latter by defining navigation speed.

3.1 Selection

Different amounts of detail and different portions of the overall visualization are visible to users depending on their distance from the display. Multiscale selection enables users to select and manipulate information at levels of detail that correspond to what is visible to the user at their current position. From an information space perspective, there are two cases for specifying scale:

3.1.1 Continuous Selection Size

This case is very similar to the "spotlight" technique in that the 'size' of the on-screen cursor changes smoothly based on the user's distance from the display. However, there is a distinct difference in semantic use – unlike the "spotlight" technique, which focuses on making a single target easier to select from a distance, the change in the multiscale cursor size changes the scale at which the user is interacting with the data on the display. A larger cursor size allows the user to select and/or manipulate everything that falls inside its area. This approach varies the size of the cursor in a continuous (or smooth) manner, according to a predefined mapping, such as intersecting a cone (extending from a handheld wand) with the display, or an exponential function of the orthogonal distance of the user to the display.

This case is most appropriate for continuous data spaces or discrete data that does not have natural hierarchical groupings. The scale can be interpreted in two ways. The cursor can grow or shrink to select either (a) larger or smaller groups of objects or (b) larger or smaller objects themselves.

An example of (a) is the visual knowledge synthesis application, called Storyboard, that we prototyped for intelligence analysis (Figure 2). It enables them to construct scenarios out of a large collection of unorganized intelligence information. It is based on Wigmore charts, a structured approach for evidence marshalling that builds, in essence, a graph that illustrates the analyst's chain of reasoning from evidence to hypotheses. Details, such as individual evidential facts gathered in the field, can be manipulated while standing close to the display (Figure 2a), including selecting notes and then editing them using a handheld PDA. Evidence can be meaningfully positioned and linked to hypotheses to form graphs. Stepping back, larger graphs are built by selecting and combining several smaller graphs together to arrive at some more encompassing scenario (Figure 2b), which ties together the suppositions represented by the smaller graphs. Whole graphs are selected by moving further from the display.

Continuous variation of selection size can also be used to select differently sized objects. An example of this is selecting windows versus selecting icons in a desktop metaphor environment.

3.1.2 Hierarchical Level Selection

Instead of smoothly varying the effective selection area of the cursor, another case of multiscale interaction would be to vary the hierarchical level of interaction of the cursor according to hierarchical structure in the data space. Essentially, this would still change the size of the cursor, but by discrete intervals, according to a predefined hierarchy within the data. An example application of this is geographic information visualization, where a frequent problem is specifying the scale of selection in a hierarchy of political boundaries. Multiscale interaction can automatically adjust selection scale to different levels of boundary hierarchy (e.g. district, county, state, country).

For instance, in a visualization of US demographics, stepping back from the display would enable a user to select state level aggregations (Figure 3b), while stepping closer to the display would enable selection of county level aggregations (Figure 3c).

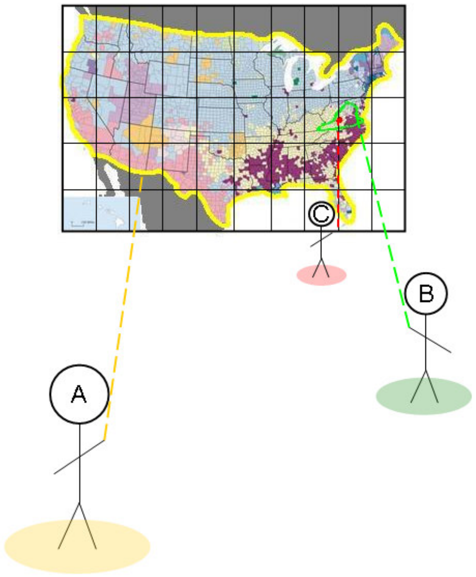

Figure 3. Mulitscale interaction can discretely vary selection area according to hierarchical levels in the data.

3.2 Navigation

Multiscale interaction may also be applied to preexisting visualization navigation strategies. Virtual navigation is typically needed even in the presence of physical navigation because total data size often exceeds total display size. One example application is rate control in 2D and 3D navigation. In large-scale 2D and 3D navigable information spaces, rate of navigation motion could be tailored to the user's visual perspective. For example, consider a ZUI (zoomable user interface) such as a space-scale zoom+pan navigation interface for large 2D satellite imagery analysis. Multiscale interaction can provide automatic rate adjustment for navigation controls, such as panning and zooming, based on users' physical position relative to the display. For example, when users are up close, they would have fine movement control and when they are further back, they would make more coarse movements. Hence, objects move across the visual field at a constant visual-angular rate regardless of the distance between user and object.

3.3 Design Considerations

Though many of these issues are not unique to multiscale interaction, we explored several design options through a case study, using the Storyboard knowledge synthesis scenario previously discussed in section 3.1.1.

Pointing Technique -- Pointing technique used can vary with domain. Many things can be tracked for pointing, including wand or the human hand. Instead of using buttons on a mobile device like a PDA, gestures can be developed for selection or other actions. A combination of PDA and tracked wand worked well, as it allowed for simple mobile text input and menu interface.

Cursor Shape -- Rather than casting a ray from the wand with a single intersection point with the display, we cast a square pyramid from the tip of the wand. The intersection with the display is no longer a point, but a square slice of the pyramid. The square shape of the cursor better fit the form factor of our data. Other selection techniques like spotlight use conic sections to form circular shaped cursors. Shape could also vary with distance if one shape is more appropriate for detail tasks than for overview tasks; or it could vary with a secondary user input, such as wand distance from eye as used in the aperture technique. Another important consideration of shape is how to handle oblique pointing angles. We chose to preserve the rectilinear square shape of the selection cursor regardless of pointing angle (essentially casting a skewed orthographic pyramid) to fit the data form factor, and avoid long narrow cursors when obliquely pointing up close.

Cursor Size -- As the user moves away from the display, the surface area of the intersection slice increases. We used an exponential function to grow the cursor size, as we found that it better matched the transition between tasks at different distances from the screen. The cursor size function should be calibrated according to the size of the space available in front of the display, the range of scale needed between detail and overview tasks, and general usability. Another choice is how to handle distance for oblique pointing angles, either ray cast distance or orthogonal distance. Users tended to prefer orthogonal distance for stability.

Feedback -- Storyboard's cursor was represented by a filled medium gray translucent square, which stood out well against the white background and could be distinguished easily from the colored evidence tidbits while also not obscuring them. Because the organization of the evidence tidbits, suppositions and links were not predefined, the filled cursor worked well in showing the selection area and whether items fell inside it. However, if color is crucial, as with the puzzle solving example in section 4.2.3, a cursor with no fill color that outlines the selection area should be considered.

3.4 Potential Benefits of Multiscale Interaction

We hypothesize that users will benefit from multiscale interaction due to improvements in several areas. We hypothesize that these benefits will manifest themselves in measurable effects on users' performance, which will be discussed in more detail in section 4.

3.4.1 Physical Navigation

Multiscale interaction exploits users' physical navigation to enable efficient interaction at multiple levels of scale. Prior experiments have shown that physical navigation is advantageous over virtual navigation for many reasons, including increasing users' efficiency and decreasing their frustration [1, 14]. By exploiting physical navigation, multiscale interaction could gain similar time performance benefits. As users move away from the display, their perception of the data changes. Information visually aggregates, and users see patterns and groups instead of individual details [16]. With this change in perception comes a change in the tasks that are appropriate to perform on the data. By designing the interaction to change as a function of the distance from the display, users' changing perceptions are supported with equivalently changing interactive controls.

3.4.2 Visual and Interaction Design

Multiscale interaction simplifies the visual design and interaction design, and therefore simplifies the users' interaction cycle. Based on Norman's *gulf of evaluation* [13], there is a perceptual advantage to linking physical navigation with a change in interaction scale – it semantically overloads the selection cursor. In addition to indicating the items for selection, the multiscale cursor also informs the user of the current interaction scale level, and gives information about the scale levels in general, reducing the need to clutter the display with other visual cues about scale levels. Based on Norman's *gulf of execution*, another benefit is less "interactive clutter." There is less need for separate user interface controls for scale level, such as mode menus or scale selectors. Thus users can focus more on their task and less on operating an interaction scale control. For example, Geospatial or Treemap visualizations would not need special 'handles' for selecting nodes at different levels of the hierarchy. Together these benefits reduce the number of steps in the interaction cycle.

3.4.3 Embodied Interaction

Multiscale interaction applies the philosophy of embodied interaction, by linking physical navigation with interaction to "give meaning to space" in front of the display. Embodied interaction theory refers to an individual's feeling of presence within the surrounding world and his or her participation with what is going on in the world. It claims that people interact with the world around them to increase their understanding of it [5]. Movement closer to and further away from an object to change the scale of visual information is nothing new; this phenomenon is already present in the world around us. The link between physical navigation and interaction scale should therefore be an already learned and natural interaction technique. More importantly, there is likely a fundamental link between perceptual scale and interactive scale wired deeply into human cognitive processes [8,11], perhaps enabling people to problem-solve more effectively in such environments.

For example, consider a patron at an art gallery. While he is up close to a particular painting, he can see one aspect of the painting or perhaps individual brush strokes. When he backs up, he can no longer see the brush strokes because they have faded together to form the entire painting. This is a change in visual scale. Interaction also changes. One art student might say to another while pointing very close at details in a painting, "Look at the type of brush strokes the artist uses here." Once the two students step back, the other might gesture at entire paintings and say, "I prefer that scene to this one." They can do this automatically without thinking about the operations. Multiscale interaction uses distance from the focus object(s) to change the scale of interaction, which corresponds to the task contexts at each location, thus leveraging the knowledge and assumptions users already have about the world around them.

4 EXPERIMENT

4.1 Motivation

We conducted a user study to evaluate the effectiveness of the multiscale interaction technique and verify whether it showed improvements over other techniques that do not exploit users' physical navigation. To this end, we wished to explore several issues. Multiscale interaction enforces a link between users' physical navigation and their interaction scale. Is this natural? Do users in fact benefit from this linkage? We hypothesized that they would, for several reasons:

1. Multiscale interaction exploits the fact that users are already physically navigating, perhaps even encourages it. We hypothesize that similar time performance benefits seen from physical navigation with perception tasks will hold for interaction tasks as well.
2. Users prefer physical navigation over virtual because it is more natural. We believe this "naturalness" will manifest itself in reduced mental workload and tendency towards similar interaction in the other, non-multiscale interaction techniques, measured by the correlation between their distance from the screen and their current interaction level.
3. Multiscale interaction reinforces the user's sense of embodiment, leveraging preexisting knowledge, while also reducing both the visual and interactive clutter. This frees up cognitive resources users may then apply to better problem solve, manifested in greater accuracy and performance.

To evaluate these hypotheses, we compared the multiscale interaction implementation to two other techniques, one representing a standard multi-selection tool and another which allowed interaction on different scales, but did not link it to physical navigation. We especially wanted to explore whether users of the non-multiscale techniques tended to physically navigate similar to multiscale users, because of hypothesized natural human tendencies to link perceptual and interactive scale.

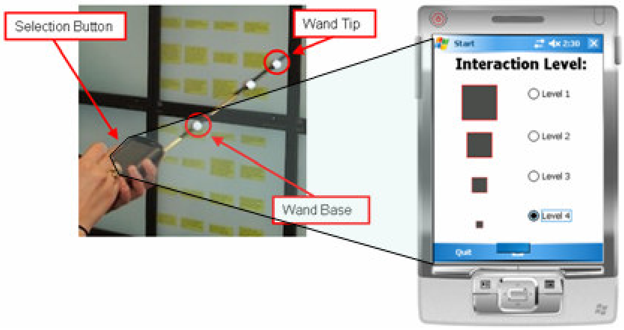

Figure 4. Wand (left) and PDA (right) were used in combination as interaction device.

4.2 Method

4.2.1 Hardware/software

Participants performed the tasks on a large, high resolution display, called the "Gigapixel" display, consisting of 50 tiled 1600x1200 pixel LCD monitors, arranged in ten columns of five monitors each, for a total of 96 million pixels (96 megapixels).

The Gigapixel display is located so that the area in front of the display is open, allowing users of the display to move around freely. This space, which is roughly 3.5 meters wide (parallel to the display) by 3 meters deep, is tracked by a Vicon MX motion-capture system, consisting of eight cameras. The Vicon system is a six degree-of-freedom, near-infrared vision-based tracking system, which uses retro-reflective balls as positional markers.

For the purpose of this experiment, we used a simple metal rod instrumented with three reflective markers as a pointing device, or wand (Figure 4 left). We similarly marked the display, allowing us to track its plane. Using this plane and the line formed by the base and tip markers of the wand, we used a simple ray-casting technique to project a line out of the tip of the wand to the plane of the display, calculating the intersection point on the display. In addition to the wand (used for pointing), participants also carried an HP iPAQ personal digital assistant (PDA), which was used for selection. Depending on the interaction type condition, participants saw different on-screen controls (Figure 4 right); however, all participants used the side button to select items during the experiment tasks.

4.2.2 Experimental Design

The only independent variable was interaction technique. There were three interaction techniques: physical navigation, explicit, and lasso, each described in more detail below. This was a between-subject variable to avoid any learning effects that might occur with task repetition, and because participants had to complete a long puzzle-solving task, with an average completion time of around 20 minutes.

Each participant was asked to complete two short and one long puzzle solving tasks. The order of the two short puzzle-solving tasks was alternated so that half of the participants performed one short task first and the other half performed the second short task first. The dependent variables for these tasks were time, accuracy (measured in total number of swaps), and total movement, a combined measure of in/out and sideways movement (taken from Vicon tracking data).

For each interaction technique we had eight participants, for a total of 24 participants. All participants were either undergraduate or graduate students. All participants were engineering majors. The average age of the participants was 28, with a minimum age of 21 and a maximum age of 37. Nineteen of the participants were male and five were female. Participants had a wide range of prior experience working with large display environments, from no prior experience to having conducted their own experiments with large display environments.

4.2.3 Tasks

The task was designed to represent other typical scenarios that could utilize multiscale interaction, specifically the case in which the dataset would be pre-organized hierarchically, and where the size of the cursor area would vary according to those predefined hierarchical levels (Section 3.1.2).

Participants completed a task that was multiscale in both the visual sense and the task sense, in order to properly evaluate the multiscale interaction technique. To be multiscale in the visual sense, at different levels of zoom, different amounts of information must be available. For example, details present in the visualization fade or become indistinct when far away, but are clear when close up. To be multiscale in the task sense, there must be different types of applicable tasks at each hierarchal level or information scale of the visualization.

Participants were asked to solve a hierarchical puzzle. In order to solve the puzzle, the participant must swap various pieces, placing them so they form a final image. The pieces in the puzzle are arranged hierarchically (pieces within pieces), so that pieces may only be swapped within the confines of their "parent piece." There are 4 levels of scale (Figure 1), with scale level 4 being the smallest at 1/16 of a monitor. Scale levels 3 through 1 are ¼ of a monitor, 1 monitor, and 4 monitors respectively. Users swap two pieces by simply selecting them.

Participants were asked to complete a series of these puzzle-solving tasks. The first two were short tasks, requiring the participant to only make two swaps on the largest scale level and two swaps on the smallest scale level. The last task was longer, requiring the user to make many swaps on all scale levels in order to solve a puzzle. Because of the difficulty and time commitment required to completely solve a puzzle, participants always started with the puzzle partially solved. Each of the 3 tasks used a different puzzle picture. The puzzle pictures and initial shuffle states remained constant across all participants.

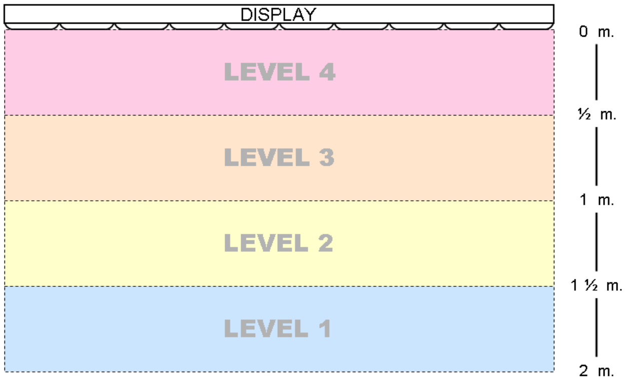

Figure 5. Interaction scale levels and distance from display for physical navigation technique. (Note: figure is not to scale)

The short puzzle tasks were designed specifically to evaluate users' efficiency at switching among the various interaction scale levels. The long task also looked at these issues (over the more long-term), but was also designed to capture richer movement data, to help evaluate the "naturalness" of the technique.

After participants completed all tasks, they then completed the NASA Task Load Index (NASA-TLX) to rate workload.

4.2.4 Interaction Techniques

There were three techniques: physical navigation, explicit, and lasso.

The *physical navigation* technique refers to the implementation of multiscale interaction being evaluated in the study. The users' cursor automatically changes size (and therefore the puzzle scale level at which the users can swap pieces) depending on their distance from the display (Figure 5). The four levels are kept well within the bounds of the tracked area described in section 4.2.1.

Much like the physical navigation technique, the cursor for the *explicit* technique can discretely change size, allowing the user to select pieces at different levels in the puzzle hierarchy. However, changing between the various scale levels must be done explicitly by the user via a menu on the PDA (Figure 4 right). This allows users to immediately switch to any level, without necessarily visiting intervening levels. This technique differentiates between the effects of the multiscale nature of the cursor and its coupling with user physical navigation.

The *lasso* technique keeps the cursor size constant at the detail level, and requires the user to manually select larger scale items by drawing a box around them. The box must enclose at least half of the intended item in order for it to be selected. This technique is reminiscent of conventional multi-selection tools found in many current systems, and is designed to examine the case of a typical detail-level interaction technique that is not multiscale.

4.3 Results

There were interesting findings related to completion times, number of piece swaps, and various movement measures, including total movement, correlation with interaction scale and user strategies. Because this study evaluates a new type of interaction technique, there was little else to build on and compare it to, and it was thus exploratory in nature. In addition, the effects of embodiment and naturalness are very difficult to measure. For these reasons, we feel it important to report quantitative and qualitative trends. Even though some of the results are not strong statistically, they raise some interesting potential issues.

4.3.1 Completion Times

Task completion time was measured for both the short tasks and the long task. The short tasks were not designed to be difficult, having only two swaps on both the largest and smallest scale levels, but instead were designed to test the performance difference between interaction types with respect to switching among scale levels. However, despite the fact that the participants were informed of the location of the out-of-place pieces prior to the task beginning, some participants still had difficulty finding them. We believe that this skewed the short task completion times, causing high variance.

For the long task, we performed a one-way ANOVA on interaction technique. There are no significant differences in completion time between any of the interaction types. However, in closer inspection of the individual participants, we can look in detail at the spread of completion times. Unlike the explicit and lasso types, the physical navigation type appears to have escaped having outlier participants with unusually long completion times. The outliers in the explicit and lasso types are participants who have lesser experience with large displays. This may indicate that the physical navigation technique allows users to perform a task more consistently, and allows less experienced users to perform on par with those having more experience.

4.3.2 Piece Swaps

Total number of swaps and swaps per scale level were counted for both the long task and short tasks. Swaps were intended to be a measurement of accuracy; however, post-experiment, we noted that this might not always be the case. We considered more swaps to indicate less accuracy (due to mistakes); however, more swaps could be a positive effect, indicating that users were comfortable enough with the technique to make many trial-and-error swaps.

For the long task, a one-way ANOVA on interaction technique showed that overall piece swaps was weakly significantly different (F(2,21)=3.21, p=0.061). Further analysis using the student's t-test showed a significant difference between the explicit and lasso techniques (p=0.022), as well as a weakly significant difference between physical navigation and lasso (p=0.096). Participants using the lasso technique made statistically significantly fewer piece swaps than either explicit or physical navigation participants.

Looking more closely at data for individual participants, as in Figure 6, there are several noteworthy findings. First, the physical navigation participants seem to be very consistent in their swap performance, while the explicit and lasso participants have greater variance. Within both the lasso and explicit types, there are outliers, indicated by red circles. The explicit outlier participants made a distinctly larger number of swaps than any of the other participants, apparently due to "mistake" swaps and inaccurate pointing. The lasso outlier participants made a very small number of swaps. This is perhaps because the lasso technique does not scale the cursor according to preset hierarchical levels, so it is more difficult to select on different scale levels; therefore, lasso participants avoid swapping as much as possible.

Coloring the markers according to each participant's prior experience with large displays shows that the outliers in the explicit type both have no prior experience. We can also notice that both of the most experienced users are in the lasso type's low-swapping sub-group. The physical navigation technique again appears to equalize participants across experience level.

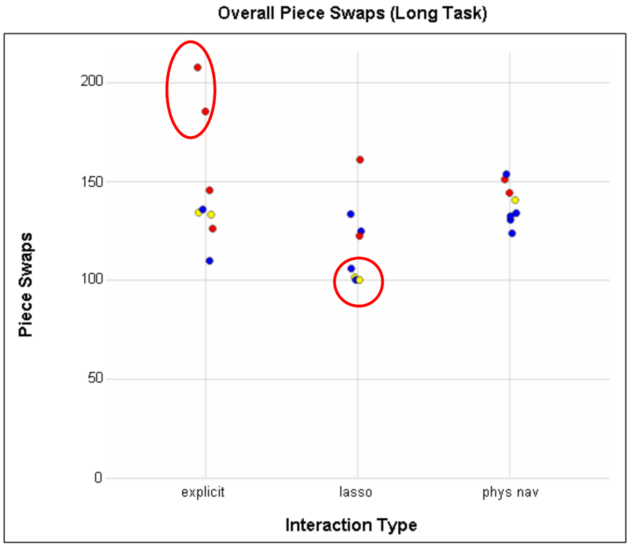

Figure 6. Total number of piece swaps by interaction technique, colored by experience (red=none; blue=some; yellow=expert) .

4.3.3 Movement

We recorded data about users' movement in the space in front of the display during the long task, using the Vicon motion tracking cameras. For three participants, one from each interaction technique, the position data did not correctly record and were removed from movement analysis.

Correlation

As a measure of the naturalness of multiscale interaction, we examined the correlation of the user's position and interaction scale level data. If the explicit or lasso techniques exhibit a correlation, it would indicate the people naturally tend to link their visual scale and interaction scale, e.g. people naturally step back when selecting large-scale objects and vice versa. This would indicate that the physical navigation technique appropriately models user's behavior and would probably have helped them.

We matched each user's position data with their current interaction scale (based on their most recent selection made), for every second in task time. We then calculated the correlation between the position and interaction scale data, using the Pearson product-moment correlation. Coefficients for the physical navigation type were unsurprisingly all high since that technique enforces a link between the two (but not exactly 1.0 because of differences between continuous motion and discrete scale data).

All correlation coefficients were consistent in sign (Figure 7), except for participant number 19, a lasso participant who had a correlation coefficient of +0.4857. All other correlation coefficients relating measured scale and distance were negative, indicating an inverse relationship: if interaction scale number increased (meaning a smaller scale, see Figure 5), then distance between the participant and the display decreased. This makes sense, as scale level 4 pieces were smaller than scale level 1 pieces. Despite the wide range of correlation coefficients, this indicates that there is a natural link between users' distance from the display and their interaction scale. This begs the question: will this link grow stronger as tasks push the multiscale aspect further and pack even more detail into the information space?

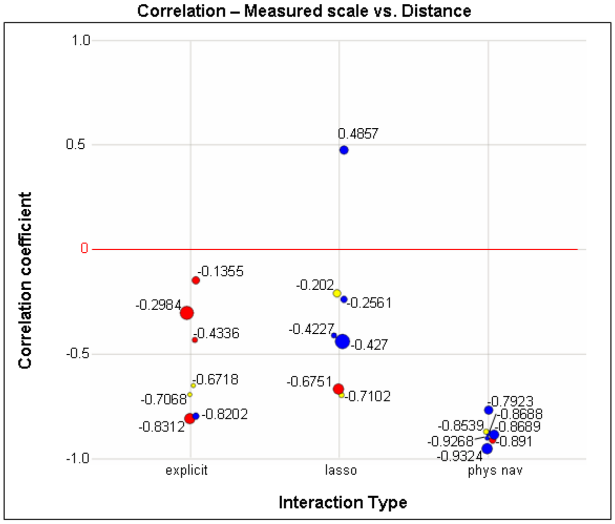

Figure 7. Correlation between user-selected interaction scale and user distance to the display, by interaction type.

Total Distance Moved

We used the position information from each user to calculate their total distance moved throughout the course of the long task. User position was noted every second, after removing some jitter from the tracking data. The distance between each consecutive pair of position samples was calculated, and these distances were summed to calculate the total distance moved (Figure 8). An ANOVA on interaction technique showed that total distance moved was weakly significant (F(2,18)=2.6739, p=0.096). Further analysis using a t-test showed a significant difference between the explicit and lasso techniques (p=0.038). Explicit participants moved significantly less than lasso participants. Though not significant, the difference in means between explicit and physical navigation is also fairly high (p=0.1246).

Though there is no difference in distance moved between the lasso and physical navigation techniques, if we look at the individual participant data, we see that the averages are the same for very different reasons. While the total distance moved is fairly consistent across all participants for the physical navigation technique, it varies greatly for the participants in the lasso technique. There are two distinct sub-groups: those who tended not to move, and those who moved a lot. For those lasso participants in the latter group, there was an interesting behavioral trend. These participants tended to move up very close to the display to make all selections, regardless of piece size. We think this is because participants in this group had great difficulty pointing with accuracy using this technique, since they had to precisely interact at level 4 scale (detail), even to specify a level 1 scale (overview) selection. In order to plan their next swap, these users would move away from the display to gain some amount of overview, and then move back in close to the display to actually make the swap. This resulted in a large amount of "thrashing".

In addition, if we look at how distance moved correlates with the absolute value of the above correlation between physical navigation and interaction scale, we see two very interesting trends. For the lasso participants, the correlation coefficient is -0.6971, meaning there is a fairly strong correlation that as the total distance moved increases, the physical navigation / interaction scale correlation decreases. This makes sense after examining how the lasso participants have changed their movement strategy to compensate for accuracy issues. On the other hand, for the explicit participants, the correlation coefficient is 0.6265, meaning there is a fairly strong correlation that as the total distance moved increases, the physical navigation/ interaction scale correlation becomes stronger. In other words, explicit users who move more tend to behave more like physical navigation users.

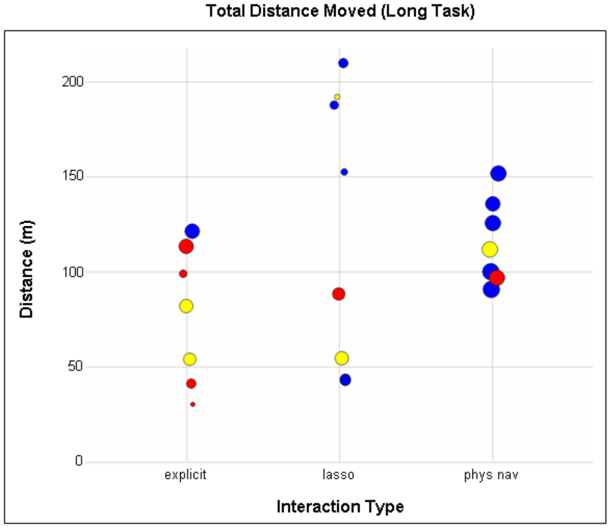

Figure 8. Total Distance Moved by Interaction Type. Markers are colored by experience and sized by |correlation|.

Strategies

We found that participants using the physical navigation technique used a fairly consistent strategy to solve the puzzle. The majority of these participants started further out from the display and arranged large scale (Level 1) pieces first, by selecting them and swapping them by swiveling their bodies to "pan." As they continued on with the task, they progressively moved closer to the display and worked with smaller scale pieces, and also tended to dynamically change their panning strategy from swiveling to shifting left and right across the length of the display. They also tended to leave the smallest scale level pieces (Level 4) for last.

There was a similar trend with the explicit technique; however, only 3 of these users seemed to follow the same patterns as physical navigation users. Another 3 users seemed stubborn to move, sometimes lunging close to peer at details but then stepping back out to select, and only inching closer to the display when they had problems pointing accurately. Many of these participants said they had problems with pointing accurately, especially with the smallest scale pieces.

For the lasso participants, the only strategy they all had in common was that they all worked within one focus area (in one large scale piece), solving it completely before moving on to another. Four of these participants had moderate to severe accuracy problems when selecting pieces; many of these participants coped with this by moving close to the display, some especially for detail work and some for all selections.

5 DISCUSSION

Since we did not expect to find strong performance differences, the study was exploratory in nature and looked for evidence of 'naturalness' in long tasks. There is evidence that there is a natural link between users' distance from the display and their selected scale of interaction, even when using techniques that do not enforce nor exploit this link.

Across all performance measures, the physical navigation technique remained very consistent in its results, unlike the explicit and lasso techniques which caused outliers. The physical navigation technique appeared to act as an equalizer, since the poor-performing outliers in the other techniques were participants who had little or no experience with large displays. Physical navigation participants also used a very consistent puzzle solving strategy, in which they progressively moved in towards the display while solving pieces on progressively smaller scale levels.

The link between interaction scale and distance enforced in the physical navigation technique appears to occur naturally across the other interaction techniques. There was a strong trend that as participants changed interaction scale to work with smaller scale pieces they also moved forward to be closer to the display. Explicit participants moved significantly less than lasso participants and though not significant, also appeared to move less than physical navigation participants. There is a possibility that this lack of movement on the part of the explicit participants could be detrimental to their performance. The key example of this is found in explicit users' piece swaps, where there are several participants who swapped pieces much more frequently than either of the other techniques. These additional piece swaps appear to be "mistake" swaps. While the mean distances moved for the lasso and physical navigation techniques were similar, this extra movement as compared to the explicit type seemed to only be beneficial to the physical navigation participants. The lasso participants had distinct trouble selecting pieces accurately and compensated for this by moving up close to the display to select any piece, regardless of scale level. They would then immediately move back out to survey their work. This resulted in a large amount of "thrashing" movement in and out from the display that did not benefit the participants' performance in any other way.

6 CONCLUSIONS AND FUTURE WORK

What do these results indicate about multiscale interaction? The multiscale interaction technique produced consistent performance with its participants, across all long task measures. Is this consistency of performance a good thing – does it mean that this technique successfully taps into the users' sense of embodiment – or is it merely the result of forcing the users to over conform? The other results do show that changing interaction scale level according to physical navigation movement is natural to people and that, while requiring users to move more, this extra movement with the physical navigation technique was beneficial. Overall, these results provide evidence that multiscale interaction is indeed a natural behavior and can be exploited for useful value in interaction design for large high-resolution displays.

The results also point out negative impacts when interaction techniques are designed against this natural behavior. The lasso technique forces users to interact on the detail level even when selecting large scale objects. They must compensate for this by selecting up close even while wanting to view at a distance. Users' physical navigation is used against them; they must move to perform interactions, but this movement has no benefit for them and is instead disruptive to their strategy. Similarly some users of the explicit technique chose to stand still at a medium distance, but had difficulty selecting accurately.

How else did people benefit from multiscale interaction other than quantitative performance measures? We looked at puzzle solving strategy as a way to gauge how interaction helped people think or reason about the data and solve problems. It was interesting to note that the physical navigation participants were consistent also in their puzzle solving strategy. This might indicate that the technique helped users better understand the structure of the data and have a more complete mental model of the puzzle, which in turn allowed them to plan a more organized solution to the problem. It would be interesting to test this on a more complex dataset.

All of this strengthens the argument for physical navigation and 3D interaction. Ball *et al.* showed that physical navigation improved user performance in the perception of data from visualizations [1]. Visualization involves two activities: perception of information, and interaction on the information. This work now shows that physical navigation also benefits visualization interaction, in addition to perception.

This work has discovered some benefits of multiscale interaction, but also raises many questions. How will these benefits hold as the dataset is changed? The multiscale aspect could be pushed to the extreme, and many more levels of hierarchy could be packed into the information space. Would the benefits of multiscale interaction hold or become even more apparent? What new kinds of multiscale interactions can be invented? We look forward to the further exploration of the multiscale interaction design space.

ACKNOWLEDGMENTS

The authors wish to thank Christopher Andrews for his initial collaboration and work on the Storyboard project.

REFERENCES

[1] Ball, R., C. North, and D.A. Bowman, Move to improve: promoting physical navigation to increase user performance with large displays, in *ACM CHI 2007*. p. 191-200.

[2] Bowman, D.A., Kruijff, E., LaViola, J.J., and Poupyrev, I. 3D User Interfaces: Theory and Practice. Addison-Wesley, 2004.

[3] Cao, X. and R. Balakrishnan, *VisionWand: interaction techniques for large displays using a passive wand tracked in 3D*, in *Proceedings ACM UIST,* 2003, p. 173 – 182.

[4] Dasiyici, M., *Multi-Scale Mouse: Optimizing Mouse Interaction for Large Personal Workspaces*, M.S. Thesis, *Computer Science*, 2008, Virginia Polytechnic Institute and State University: Blacksburg, VA.

[5] Dourish, P., *Where the Action Is: The Foundations of Embodied Interaction*. 2004: MIT Press.

[6] Forsberg, A., K. Herndon, and R. Zeleznik, *Aperture based selection for immersive virtual environments*, in *ACM UIST* 1996, p. 95-96.

[7] Jiang, H., E. Ofek, N. Moraveji, and Y. Shi, *Direct pointer: direct manipulation for large-display interaction using handheld cameras*, in *ACM CHI* 2006, p. 1107-1110.

[8] Kranstedt, A. and I. Wachsmuth. *Incremental Generation of Multimodal Deixis Referring to Objects*. in *proc. European Workshop on Natural Language Generation (ENLG2005)*. 2005.

[9] Liang, J. and M. Green, *Geometric modeling using six degrees of freedom input devices*, in *proc. CAD & Computer Graphics,* 1993.

[10] Liang, J. and M. Green, *JDCAD: A highly interactive 3D modeling system.* Computers & Graphics, 1994. **18**(4): p. 499-506.

[11] Lucking, A., H. Rieser, and J. Stegmann. *Statistical Support for the Study of Structures in Multimodal Dialogue: Inter-rater Agreement and Synchronisation*. in *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue*. 2004.

[12] Ni, T., G. Schmidt, O. Staadt, M. Livingston, R. Ball, and R. May, A survey of large high-resolution display technologies, techniques, and applications, in *proc. of IEEE Virtual Reality 2006*, p. 223-236.

[13] Norman, D.A., *The Design of Everyday Things*. 1988, New York: Basic Books.

[14] Shupp, L., R. Ball, B. Yost, J. Booker, and C. North, *Evaluation of viewport size and curvature of large, high-resolution displays*, in *proc. Graphics Interface (GI'06)*. 2006, p. 123-130.

[15] Vogel, D. and R. Balakrishnan, *Distant freehand pointing and clicking on very large, high resolution displays*, in *proc. ACM UIST* 2005, p. 33 – 42.

[16] Yost, B., Y. Haciahmetoglu, and C. North, Beyond visual acuity: the perceptual scalability of information visualizations for large displays, in *ACM CHI 2007*, p. 101 -- 110.