Mélange: Space Folding for Multi-Focus Interaction

Niklas Elmqvist[1] **Nathalie Henry**[1,2,3] **Yann Riche**[1,2,4] **Jean-Daniel Fekete**[1]
{elm, nhenry, yann, fekete}@lri.fr

[1]INRIA
Saclay, France

[2]LRI, Univ. Paris-Sud
Orsay, France

[3]University of Sydney
Sydney, Australia

[4]University of Queensland
Brisbane, Australia

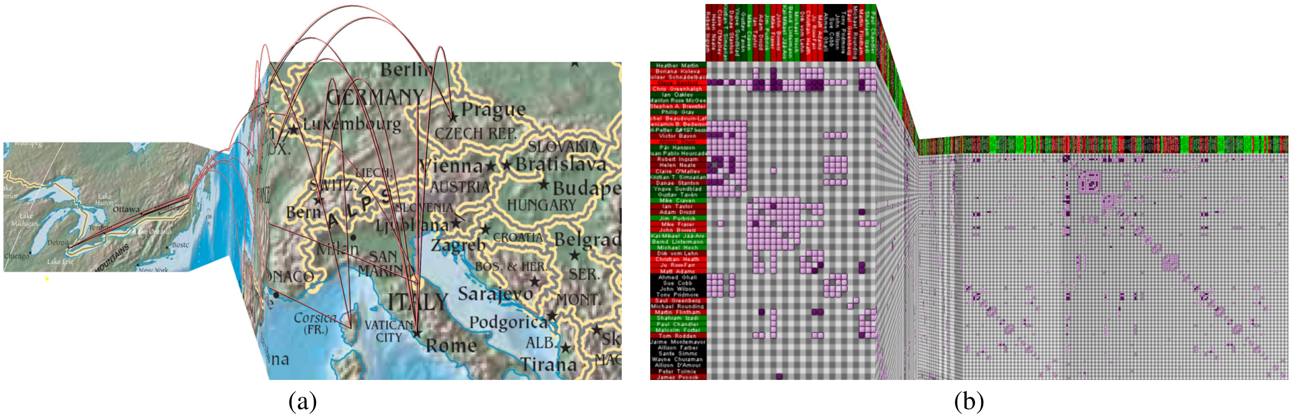

Figure 1. Examples of the Mélange technique: (a) Browsing flight routes on a world map. (b) Displaying a large matrix visualization of a network.

ABSTRACT

Interaction and navigation in large geometric spaces typically require a sequence of pan and zoom actions. This strategy is often ineffective and cumbersome, especially when trying to study several distant objects. We propose a new distortion technique that folds the intervening space to guarantee visibility of multiple focus regions. The folds themselves show contextual information and support unfolding and paging interactions. Compared to previous work, our method provides more context and distance awareness. We conducted a study comparing the space-folding technique to existing approaches, and found that participants performed significantly better with the new technique.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—*Interaction styles*; I.3.6 Computer Graphics: Methodology and Techniques— *Interaction techniques*

Author Keywords

interaction, visualization, navigation, exploration, folding, split-screen, space distortion, focus+context

Submitted to CHI 2008. Please do not distribute.

INTRODUCTION

Current visualization applications often involve navigation in large visual spaces — many times the size of the screen — using a sequence of zoom and pan operations. The tasks that are performed on these spaces typically require multiple objects to be displayed at sufficient scale for precise manipulation, yet these objects may be separated by long distances. Zooming and panning is tedious, potentially disorienting, and often ineffective [6, 7]. In order to retain view of multiple objects, the standard practice is to split the screen into several subwindows, but this means that the context between the objects is lost.

Consider a researcher planning a conference trip from eastern Canada to Florence, Italy. Beyond constraints such as cutting costs and minimizing the number of stops and the flight time, the researcher may be interested in combining such a long trip with visits to other labs and professional acquaintances in Europe. Thus, our traveler wants to study maps of both the source and destination areas at sufficiently high detail to make informed decisions about departure and arrival airports as well as appropriate ground transportation and lodging, yet is also interested in seeing the context between these areas to get an idea of opportunities for potential detours and research visits. Panning and zooming the map to solve this task is burdensome and ineffective. Similarly, splitting the screen to show several regions of the map simultaneously causes loss of context of the intervening space. Figure 1(a) illustrates how the Mélange technique presented in this paper solves this problem.

The situation is very similar when exploring social networks. These can be represented as matrices to avoid node overlap or edge crossings, which is particularly useful for dense and large networks [9]. Here, nodes are placed on the row and column axes, and a filled cell in the matrix indicates an edge between nodes. Often, several different parts of the same matrix are interesting for a particular task, such as collaborating actors, as well as the intermediate context between them (the communities they belong to). However, no efficient interaction technique exists for exploring these matrices, leaving the user no option but to pan and zoom to navigate. Figure 1(b) shows our approach.

These are two examples of multi-point interaction tasks [26] that require several concurrently visible focus points. In our generalized *multi-focus interaction* model, we also stipulate that each focus must be independently zoomed so that the user can adapt the magnification to the task. Furthermore, as much display space as possible should be dedicated to each focus to show its surrounding context. Finally, our intended user tasks often require an awareness of the content and quantity of space that lies between the foci. For the world map example, context and distance helps the user quickly estimate flight time and stopovers on the way. For the social network, they give an indication of the global communities and collaboration patterns.

Based on these requirements, we formulate a number of design goals for our approach to exploring large visual spaces:

G1 *guaranteed focus visibility:* multiple foci at the desired zoom level should be visible simultaneously, regardless of their location on the space;

G2 *surrounding context visibility:* as much as possible of the area surrounding each focus region should be visible;

G3 *intervening context awareness:* the space between focus regions should be shown to give a frame of reference; and

G4 *distance awareness:* some notion of the distance between the focus regions should be available.

No existing interaction technique is known to fulfill all of the above design goals. Therefore, we present the *Mélange* technique that automatically folds intervening space between focus regions to guarantee their visibility.

The rest of this paper is structured as follows: We begin with a review of the existing work on space deformation and similar techniques for exploring visual spaces. We then present the Mélange interaction technique. We describe our controlled experiment and present the results and a discussion of our findings.

RELATED WORK

There are a number of existing techniques (or combinations of techniques) that partially fulfill the design goals outlined above. This section reviews the main approaches:

- *General navigation:* interaction techniques for navigating in large visual spaces;
- *Split-screen:* dividing the viewport into smaller subwindows, each showing a small region of the space;
- *Space distortion:* deforming geometric space; and
- *Semantic distortion:* deforming semantic space.

Table 1 gives a summary of these strategies and how they fulfill our design goals.

Solution strategy	G1	G2	G3	G4	Techniques
General navigation	–	–	–	–	[1, 13]
Split-screen	Y	Y	–	–	[26]
Fisheye views	Y	P	Y	–	[6, 26]
Rubber sheet	P	P	Y	–	[25, 17, 27]
Semantic distortion	Y	Y	Y	–	[4, 21]

Table 1. Design goals fulfilled by existing strategies (P = partially).

General Navigation

Zooming and panning are the standard actions for interacting with large visual spaces that exceed the size of the viewport. Furnas and Bederson present the space-scale diagram [8] as a comprehensive model for describing these actions as paths through scale-space. In general, using both zoom and pan in combination is both more efficient and more informative than using just panning [3, 8, 28]. However, zooming and panning do not directly support any of our design goals.

A number of approaches have been developed to better support navigation in zoomable spaces. Speed-dependent automatic zooming [13] (SDAZ) seamlessly zooms out to maintain a fixed visual flow depending on the speed of scrolling governed by the user's cursor. Bourgeois and Guiard [3] show that bimanual multi-scale navigation outperforms standard navigation. OrthoZoom [1] allows for controlling both zoom and pan using the orthogonal axes of the mouse in a 1D scrolling task, and was recently shown to be the fastest one-dimensional scrolling technique.

For larger visual spaces, standard navigational aid include an overview window showing the position and general context of the viewport on the canvas [20]. A recent trend integrates the overview in the detail view to provide off-screen target awareness; examples include Halo [2], where circles emanating from off-screen targets indicate their approximate distance and location, City Lights [29] that show the "shadows" of off-screen targets on window borders, and the EdgeRadar [10] that provides a rectangular context region on window edges. Hopping [14] extends the idea by also allowing for direct teleportation to any of the off-screen targets indicated on the viewport edge. However, again, these techniques do not provide multiple foci, and provide poor context awareness.

Split-Screen

Splitting the screen into several windows showing different parts of the visual space is a standard method employed by commercial applications such as Microsoft Excel and Adobe PhotoShop. However, there exists no evaluation on the performance of navigation in such split-screen setups.

Shoemaker and Gutwin [26] present an interaction technique called split-scrolling that automatically divides the screen into two viewports when two interaction points move apart, but they do not empirically evaluate this technique.

For time-series data, it is useful to be able to summarize or condense periods of times into aggregated representations. An example is LifeLines [22], where the time navigation scrollbar can be split into several regions with multiple foci.

By definition, split-screen setups support the guaranteed visibility (G1) and surrounding context (G2) goals, but intervening context (G3) and distance (G4) is lost. Adding an overview helps to show the context, but overviews are typically small and placed in the periphery of the viewport, splitting the user's attention and consuming screen real estate.

Space Distortion

Instead of having the user travel through the visual space, space-distortion techniques deform the space non-linearly to optimize browsing. Fisheye views [6, 7] describe ways of doing this, both in geometric as well as information space. The Table Lens [23] is an example of applying fisheye distortion to a tabular visualization. The Document Lens [24] visualizes a large document as a rectangular array of pages with a focused region in 3D. This use of 3D perspective foreshortening as a distortion technique is also used in the Perspective Wall [16]. However, most of these approaches have no direct support for our design goals, although they can be used as starting points for fulfilling them.

The rubber sheet stretching metaphor [25] is one model for distorting 2D space. Accordion Drawing [17] (AD) is an extension of the rubber sheet with support for guaranteed visibility. Slack et al. [27] present a general application framework for accordion drawing. The AD method supports all of our design goals, but some of them only partially. Focus regions cannot be zoomed independently (G1), the model is not view-dependent so surrounding context is not automatically allocated a maximum amount of space (G2), and the compressed space gives no direct distance awareness (G4).

Instead of distorting the whole space, Shoemaker and Gutwin [26] describe a multi-point interaction technique based on automatic creation of fisheye lenses for each interaction point. As for the AD method, this approach supports design goals G1 and and G3, but there is no automatic space allocation given the available space (G2), and distance awareness (G4) is difficult to attain when the space is non-linearly deformed. For our exploration task, it would make more sense to deform the context regions and leave the focus unchanged and of maximum size, whereas fisheye lenses allocate space for the foci and leave the context unchanged.

Semantic Distortion

As stated earlier, fisheye views [6] also allow for semantic zooming [19]. In other words, we can distort semantic space instead of geometric space. DOITree [4] and SpaceTree [21] are examples of such techniques for hierarchical structures. However, while this approach can support design goals G1 through G3, it is again distance awareness (G4) that is lacking due to the scale-independent graphical representation.

MÉLANGE: FOLDING 2D SPACE INTO 3D

Mélange is a space deformation technique that folds 2D space into 3D in order to bring several focus regions of interest into view at the same time. Figure 1 shows a large world map being folded using Mélange to bring both northern Italy and eastern Canada into view at high magnification, as well as a matrix visualization of a social network being folded to simultaneously view different parts of the network.

Multiple Foci: Guaranteed Focus and Context Visibility

Given a set of focus points and the location and extents of the current viewport on the canvas, the objective of the Mélange technique is to combine different parts of the visual space so that the focus points and as much as possible of their surrounding context are visible on the user's screen. This fulfills the *guaranteed focus visibility* (G1) and *surrounding context visibility* (G2) design goals.

Focus points are specified as 2D positions on the visual space, and also have an associated depth factor that allows each point to be zoomed independently of the others. This supports interactions where different parts of the visual space must be viewed at different scales, such as a social scientist studying a particular actor in relation to a larger clique of actors on a matrix representation of a social network.

Folding Space: Intervening Context Awareness

A split-screen approach to multiple foci would remove space outside of the focus regions and show each region as small subwindows in the main viewport. Mélange instead *folds* the space into the negative depth dimension (i.e. into the screen, see Figure 1). If there is no extraneous space to fold away, the space is instead stretched, similar to the rubber sheet [25] but with support for independent depths for each focus point.

The folds themselves are shown in 3D perspective as they stretch away into the depths of screen, and they also indicate the relative positioning of the focus points. Thus, this fulfills the *intervening context awareness* (G3) design goal. Furthermore, this mechanism gives a tangible and compelling metaphor for the user that is close to how real paper or fabric is folded. We believe that this metaphor is easier to understand than merely compressing the space, as in rubber sheet-inspired models.

Figure 2 shows a schematic overview of the folding process. The user's viewport (denoted by the smaller rectangle in the left part of the figure) is centered on the focus point A — the main focus — but the user has also designated a second focus point, B. Given the available space in the viewport, the Mélange technique folds away some of the intervening space below and to the left of A to also bring B onto the screen. All folds are rectilinear to simplify understanding of the deformed space. A certain amount of screen real estate (*foldSize*) is used to show the contents of the folded space in 3D perspective as it stretches away into the depths of screen. These regions serve as context between the focus regions.

The above method generalizes to any number of additional focus points. One of the foci is always designated as the main one and is used as a baseline for computing the size allocations for the others.

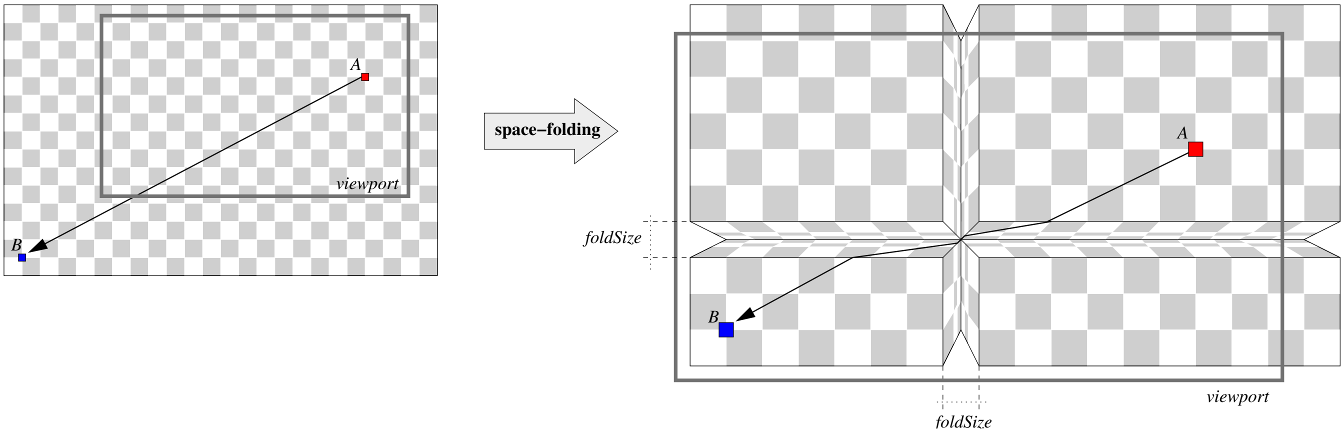

Figure 2. Folding a 2D space with two focus points *A* (main) and *B*. The space is folded to make best use of the available area in the viewport. Focus points can be independently zoomed by changing their 3D depths.

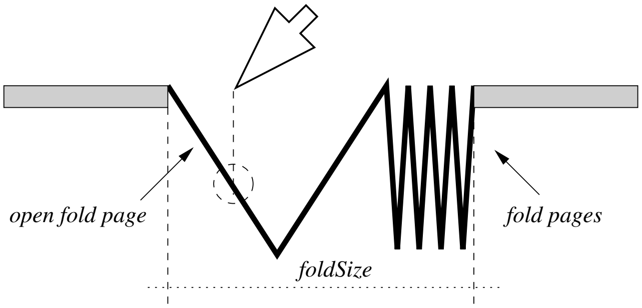

Figure 3. Fold pages for conveying a sense of distance between focus regions. Supports flipping and defining new focus points.

Interacting with Folds: Context and Distance Awareness

Deforming the space to bring several foci onto the screen may give the user a false sense of the size of the visual space. For example, folding a world map to bring London and New York into focus at high detail level will certainly convey a false sense of the distances between the two cities.

Mélange supports better distance awareness than compression-based techniques like the rubber sheet method [25] since the 3D perspective of the folds gives an indication of the distance between the regions.

To further improve distance awareness, we introduce fold pages and interaction techniques for flipping between them. The folded space is split by a suitable and tangible unit, such as the size of the screen. Only one such unit is shown at full detail, and the rest are shown as thin fold pages (Figure 3). Each fold page represents one screen of compressed space. This helps fulfill the *distance awareness* (G4) design goal by allowing the user to quickly estimate the number of fold pages to find the distance between the focus points (like estimating a book's length from its thickness).

Another benefit is that context awareness is improved by allocating more screen estate to each individual fold page (although some overview is lost). Pages could potentially also show condensed context information on its one-pixel representation, akin to the compact contextual views of the City Lights [29] technique.

Hovering with the mouse over the pages flips through them like leafing through a book. Furthermore, clicking on a fold adds a focus point on the designated location, and double-clicking removes all of the other focus points and creates a new primary focus point at the position. The effect is that the user stops folding space and travels to the new location.

Design Decisions

In this section we deal with some of the specific design decisions underlying Mélange. Note that the method does not stipulate how the user interacts with the focus points, allowing it to be combined with advanced multi-scale navigation techniques like OrthoZoom [1] or SDAZ [13].

Fold Geometry

The Mélange space-folding mechanism is different to most focus+context techniques in that it compresses uninteresting space as opposed to expanding the focused space. The geometry of the actual folds is an interesting design issue; to fully support the metaphor of folding paper or fabric, the space should probably be folded in a smooth curve. However, this would cause most screen estate to be afforded to the middle region of the compressed space.

Most often, the space closer to a focus region is more important than the space halfway between regions. Therefore, in our realization, the folds are sharp and angular (more like paper origami than fabric folding), similar to the Perspective Wall [16]. 3D perspective foreshortening gives a form of fisheye effect on the contents of the folds.

Perspective Correction

When rendering the visual canvas and the folds in 3D, we must correct for perspective to get a correct visual appearance for the folds. Otherwise, the perspective projection of the 2D space deformed into 3D causes uneven distribution of screen space. Carpendale [5] calls this *folding* a region over other regions, unrelated to our use of the term. We solve this by performing all layout in the 2D screen space, and then unprojecting to 3D world space.

USER STUDY

We performed a controlled experiment to evaluate whether the Mélange technique assists users in exploring large visual spaces, by comparing it to single and split-screen viewports. We designed the experiment to test our design goals in the context of a matrix visualization of a large graph with MatLink [12] arcs connecting relevant nodes in the graph.

Participants

We recruited 12 unpaid subjects (1 female, 11 male) for our study. The participants were from 20 to 35 years of age, had normal or corrected-to-normal vision, and were screened to not be color-blind. No specific skills were required other than basic computer experience.

Apparatus

The experimental apparatus consisted of an Apple iMac Core 2 Duo 2.33 GHz workstation with 2 GBs of memory and equipped with a standard two-button mouse (with wheel) and keyboard. The 21-inch display was fixed at 1680 × 1050 resolution and powered by an ATI Radeon X1600 with 256 MB of video memory.

Tasks

Participants were given a source node and its neighborhood on an adjacency matrix representation of a social network, and were then asked to perform three tasks in sequence:

T1 Find one destination node connected to the source node with the same neighborhood [G1 and G2]

T2 Estimate the distance between the source and destination nodes (in 1:1 screen units) [G4]

T3 Estimate the number of contextual targets between the source and destination nodes [G3]

This scenario was inspired by social network analysis, where a common task is to compare the local neighborhood of two actors to find similar patterns of collaboration.

Potential targets in our study were blue squares measuring 20 pixels (at 1:1 zoom level), surrounded by a neighborhood of four half-size (10 pixel) squares of different colors (Figure 4). We chose five colors for these neighborhood squares: white, magenta, orange, green, and blue (a selection that is preattentively perceptible [11]). Neighborhood nodes were placed in a 5 × 5 grid around the blue rectangle, and whole targets were placed in one line on the visual space, like columns in a matrix visualization.

Targets were identical if both the position and color of their neighborhood nodes are identical. Only one other target neighborhood matched the source target, other were distractors. Connections between the source node and the potential targets were visualized using MatLink arcs. Not all nodes on the visual space had a MatLink arc from the source node; those without were background nodes that also served as distractors, and participants were instructed to disregard them when looking for the destination target.

Contextual targets (T3) were red squares six times the size of primary targets (i.e. 120 pixels) and below the line of primary targets. The motivation for this was that being aware of intervening context is only applicable for large-scale features such as mountain ranges or large bodies of water on a map, or communities of actors in a social network.

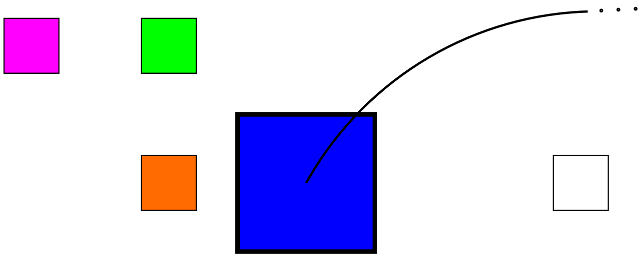

Figure 4. Example of a source target with its four-node neighborhood.

All targets on the visual space — i.e. target nodes, neighborhood nodes, and contextual targets — were guaranteed to be rendered with at least a single pixel, forcing them to be visible even if the view was zoomed out or distorted.

The visual space itself was represented by a checkered gray rectangle that was 30 screens wide and one screen high. Each scenario had randomly-generated distractors. The source node was always located on the left edge of the rectangle, so the participant would always have to pan right to find the target. The view was initialized to center on the source node at 1:1 zoom level for every new scenario (started by T1), and was then left in its previous position for each consecutive task (T2 and T3).

Finally, to give users a frame of reference for distance, screen units were indicated on the visual space by black lines drawn on the checkered gray rectangle. Figure 5 shows a screenshot of our experiment application.

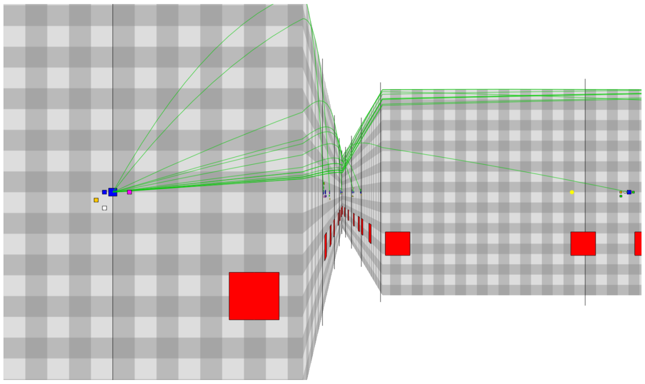

Figure 5. Screenshot from the user study application.

Experimental Conditions

The factors were presentation technique, off-screen distance, distractor density, and contextual target density.

Presentation Technique

The primary objective of our experiment was to study the performance of different presentations of the visual space for supporting our design goals. In addition to the Mélange technique, we included single and split-screen viewport conditions for comparison. While none of these two fulfill our design goals, they are commonly used in practice, suggesting that they are suitable competitors.

We considered comparing our technique against Accordion Drawing [17]. However, AD does not seem to support independently zoomed foci. Furthermore, Nekrasovski et al. [18] have shown that pan and zoom for a large hierarchical dataset is more efficient than navigation in AD spaces, hence our choice of competing techniques.

- *Single viewport (SV).* The standard baseline consisting of a single window showing a view of the visual space. Has no direct support for any of our stated design goals, these must be achieved through interaction.
- *Split-screen viewport (SSV).* The main viewport is split vertically into two equal-sized subwindows, each showing a different view of the visual space. In our setup, the left subwindow was fixed to always show the source node at 1:1 zoom, while the user could interact with the view of the right subwindow.
- *Mélange (M).* Our space-folding technique with the primary focus point on the source node and the secondary point controlled by the user. Moving the focus point (in the horizontal and depth dimensions) will thus cause the visual space to be folded to accommodate both focus points in the viewport. Fold pages were disabled to not unfairly give a direct distance measure to the participants (i.e. only the 3D perspective foreshortening of the folds indicated distance).

All three techniques were controlled using standard zoom and pan operations. Dragging the mouse while clicking the left mouse button caused horizontal movement of the focus point (the camera for single viewport, the right subwindow for split-screen, and the folding focus point for Mélange). The focus could be zoomed in and out by dragging with the right mouse button, or by spinning the mouse wheel.

Off-Screen Distance

We wanted to see whether performance varied with the distance to traverse on the visual space, so we tested three different distances: 4, 8, and 16 screen widths of distance (in our experimental setup, the screen width was 1680 pixels). In a matrix representation, this corresponds approximatively to networks containing 400, 800, and 1600 actors.

Distractor Density

The number of false targets (i.e. distractors) between the source and destination nodes will clearly affect the time spent finding the destination node (T1). Thus, we included two different densities: *low* or *high*. This corresponded to one or two potential targets per screen (half of them background nodes with no MatLink arcs to them).

Contextual Target Density

We studied two levels of density for the contextual targets between the source and destination nodes: *few* (less than or equal to five) or *many* (more than five).

Experimental Design

We used a $3 \times 3 \times 2 \times 2$ within-subjects factorial design. The factors (described above) were:

- Presentation technique: single (SV), split (SSV), and Mélange (M)
- Off-screen distance: 4, 8, or 16 screens
- Distractor density: 1 or 2 per screen (average)
- Contextual target density: few (≤ 5) or many (> 5)

The order of the techniques was counterbalanced: two participants were assigned to each order. Participants were asked to complete 3 blocks of 24 trials (3 distances $\times$ 2 distractor densities $\times$ 2 contextual target densities $\times$ 2 trials) in randomized order. With 12 participants, the study software collected 864 trials in total.

Procedure

Participants were introduced to the study and randomly assigned to one of the six order groups for the presentation technique. They then performed three blocks of trials, one per technique, in succession. Before each block, the test administrator explained how to use the technique and then let the participant practice on six training trials. Participants were not allowed to proceed past each training trial without answering correctly to all three tasks.

Each trial consisted of performing the three tasks T1 to T3 in sequence. A screen with instructions were given prior to each task, and the participant proceeded to the task by clicking a button or pressing the space bar. Task T1 ended when the participant clicked the right target (which then turned from blue to yellow); for the other tasks, the participant pressed the space bar to end the task. After task T2 and T3, participants were presented with a multiple-choice question asking about their answer to the task.

Participants were instructed to work as quickly as possible. For every trial, the software silently collected the time and correctness measures for the three tasks (only time for T1). Participants were instructed to pause between each block to avoid fatigue affecting the results. At the end of the test, they were given a preference questionnaire to complete.

Predictions

P1: Mélange is as fast as single or split-screen viewport

We believe that the space-folding technique will not introduce significantly slower completion times for standard visual search (task T1). In other words, we think that the added visual complexity and space allocations of the fold region and the additional focus point will not cause slow-downs for a user trying to locate a specific target on the canvas.

P2: Mélange provides more efficient context awareness

None of the two techniques we compare Mélange to support contextual views explicitly, but participants are nonetheless exposed to this context when navigating over the visual space. We submit that the intervening context shown in the fold regions of the technique will cause significantly lower completion times for tasks T2 and T3.

Task	Factors	F	p
T1	Distance	38.740	**
	Distractors	55.155	**
T2	Technique	8.695	*
	Distance	6.560	*
	Technique*Distance	6.658	**
	Distance*Distractors*Context	4.216	*
T3	Distance*Context	5.335	*
	Technique*Distance*Context	2.660	*

* = $p \leq 0.05$, ** = $p \leq 0.001$.

Table 2. Significant effects of completion time on the factors.

P3: Mélange provides more accurate context awareness

Analogously to P2, we also believe that participants will be more accurate when answering contextual tasks (T2 and T3) with Mélange than the other two presentation techniques. Mélange provides an integrated overview of the context, whereas the other two require the user to manually pan and zoom around in the space to discover this information.

RESULTS

We analyzed the measurements collected from the study for efficiency (completion time) and correctness (error rate).

Completion Time

Table 2 summarizes the main effects for time. Figure 6 shows mean time to completion for all tasks.

For task T1, the average completion time was 18.05 (s.d. 1.42) seconds for SV, 16.98 (s.d. 0.85) seconds for SSV, and 19.18 (s.d. 0.99) seconds for M (SSV < M < SV). A repeated-measures analysis of variance (ANOVA) showed no significant main effect of Presentation technique.

For task T2, the average time was 4.13 (s.d. 0.64) seconds for SV, 4.02 (s.d. 0.43) seconds for SSV, and 2.74 (s.d. 0.35) seconds for M (M < SSV < SV). ANOVA yielded a significant main effect for Presentation technique ($F_{2,22} = 9.203, p = .001$).

For T3, the average time was 1.72 (s.d. 0.57) seconds for SV, 1.90 (s.d. 0.50) seconds for SSV, and 1.64 (s.d. 0.19) seconds for M (SV < M < SSV). ANOVA yielded no significant main effect for Presentation technique.

For task T2, the average correctness was 0.986 (s.d. 0.007) for SV, 0.948 (s.d. 0.013) for SSV, and 0.983 (s.d. 0.008) for M (SV > M > SSV). This is a significant difference (Friedman test, $p = .008$). A Wilcoxon test for paired comparison shows that M and SV have higher correctness than SSV (M vs SSV: $p < .025$, SV vs SSV: $p < .012$). Figure 7 shows the mean correctness for T2.

For task T3, the average correctness was 0.983 (s.d. 0.008) for single viewport, 0.965 (s.d. 0.011) for split-screen, and 0.983 (s.d. 0.008) for Mélange. This is a non-significant difference (Friedman test, $p = .189$).

Subjective Preference

When asked about their preference on the presentation technique, 5 out of 12 participants ranked the Mélange technique first (5 for split-screen and 2 for single viewport). Comments from the participants were favorable for our new technique, particularly for contextual tasks.

DISCUSSION

Summarizing the previous section, our user study yields the following results:

- Our experiment shows no significant differences between the three techniques for visual search (T1) so we cannot conclude about our prediction P1. With 12 participants, the techniques seemed comparable in performance.
- Mélange is significantly faster for the contextual task T2 than both single and split-screen viewport, confirming prediction P2. The difference is almost one-third of the completion time for the competing techniques.
- Mélange promoted significantly better correctness than split-screen viewport. This partially confirms prediction P3. There was no difference for Mélange in comparison to single viewport, but this may be due to single viewport simply not supporting quick contextual assessment.

In the following sections, we try to explain and generalize these results, and see how our work can be used in practice.

Explaining the Results

These results confirm that the Mélange space-folding technique provides extra benefit beyond the standard split-screen method. More specifically, the results show that providing an awareness of intervening context and distance between focus points helps for contextual tasks, while clearly not consuming too much screen space or cognitive effort to cause poorer performance than split-screen viewports.

Looking at the completion times for task T1, we note that there is no large difference between single-focus (single viewport) and the two double-focus (split-screen and Mélange) presentation techniques. The reason for this is that T1 is a relatively simple visual search task where the target appearance can be memorized, so two foci are not strictly necessary. We designed the study this way to avoid punishing the participants with very long completion times — instead, the objective of task T1 (rather than strictly confirming G1 and G2) is to show that space-folding does not introduce slow-downs in navigation compared to single or split-screen viewports (prediction P1).

It is also worth noting that Mélange is a novel and relatively complex presentation technique, whereas our participants had all encountered single and split-screen viewport presentations before. This may account for the long completion times for task T1 using Mélange, but we believe that this measure will decrease as a user becomes more comfortable using the technique.

We found no significant difference in completion time for the T3 task, so our prediction P2 only holds for contextual task T2. However, we observed that participants in the user study tended to solve both T2 and T3 simultaneously during the T2 time. This was possible because distance indicators and contextual targets were visible for both tasks. If we combine

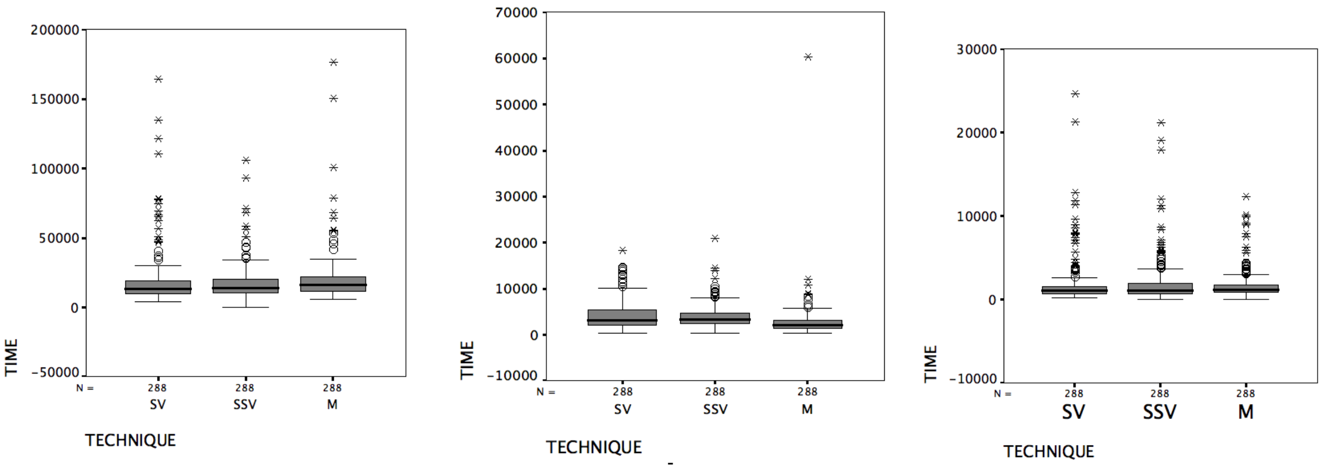

Figure 6. Average completion times for presentation technique across T1, T2, and T3.

the completion times for both tasks, the average time was 5.74 seconds for SV, 5.79 seconds for SSV, and 4.17 (s.d.) seconds for M (M < SV < SSV). Removing outliers, this is a significant difference ($F_{2,22} = 4.289, p = .027$).

While Mélange was significantly more correct than split-screen, there was no difference in comparison to single viewport. We believe this is due to single viewport simply not supporting quick assessment of context. With Mélange, users can easily retrieve the contextual information, whereas split-screen and single viewport require users to invest considerable time to reach the same accuracy.

Generalizing the Results

Our results show that the Mélange technique fulfills most of our predictions for the chosen scenario and tasks. The question is naturally whether these results generalize to the whole class of large visual spaces discussed in the introduction.

The answer to this question is two-fold: We believe that the tasks and the scenario used in the study are realistic enough to be ecologically valid, yet general enough to allow us to extend the results to other domains. For the first point, the tasks selected are based on typical user tasks for network analysis [15]. For the second, the study scenario is sufficiently abstract so that there is nothing in the tasks or the scenario that limits the results. In fact, for the world map example (depicted in Figure 1(b)), contextual tasks may become even easier due to the inherent multi-scale properties of a map (i.e. large-scale features like ocean, land, and mountains are visible even from long distances and under great distortion).

One specific threat to generalizing the results is that we only tested one-dimensional navigation (horizontal) in one direction (left to right). Two-dimensional tasks may exhibit differences depending on the relative positions of the foci.

For larger distances (more than the 16 screens tested in our study), the performance may degrade since the folds become very small and dense. This would happen when navigating a DNA sequence, for example. Supporting this situation is left for future work.

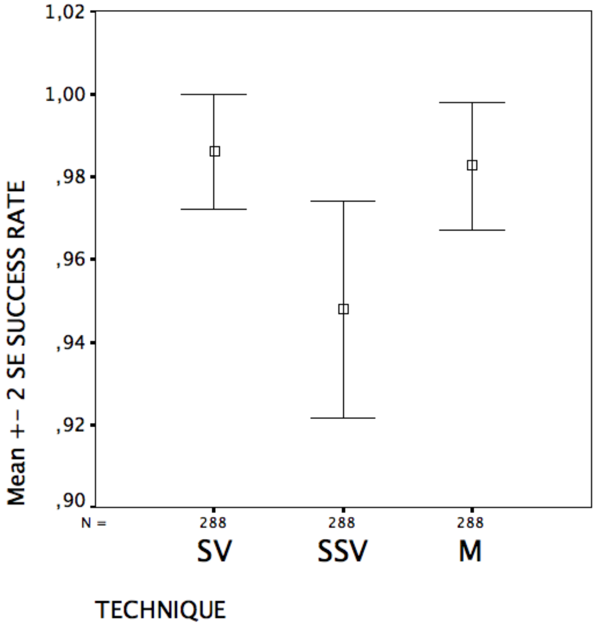

Figure 7. Correctness for presentation technique for T2.

Multi-Focus Interaction in Practice

One important issue with all multiple-foci techniques, including split-screen and space-folding as well as overview windows, is that they divide the user's attention between several different viewports and consume valuable screen estate. Even for a focus+context technique like Mélange, there is a non-trivial cognitive effort associated with comparing the different focus regions. As for screen space, users typically interact with only one area of the visual space at a time, so multiple-foci techniques reduce the amount of screen space available for this interaction. Mélange is slightly worse than split-screen due to the fold regions also consuming screen space. Having just a single viewport sidesteps both of these concerns. However, this loss of screen space is balanced by

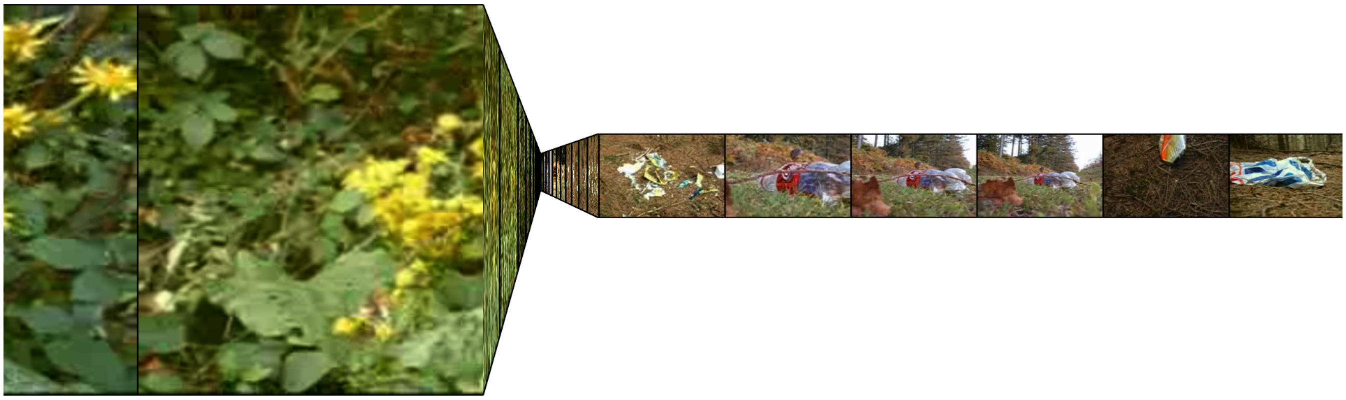

Figure 8. Folding a 1D video editing timeline using the Mélange technique.

improved context awareness.

As has been shown in this paper, split-screen is perhaps the primary competitor to space-folding. One of its major advantages is its simplicity, both for interaction and implementation. Mélange is unquestionably more complex in both aspects, but we believe that its advantages outweigh this fact. Not only does space-folding better show contextual information, as has been proven in this paper, but it also integrates several foci into the same continuous view, and directly gives the relative positioning of the foci. By the same token, split-screen viewports are fully independent of each other, so they give no intrinsic indication of what part of the space they are showing in relation to the others. In fact, both subviewports may be showing the same target, causing the user to mistake the source node for the destination node, as happened to one of our study participants.

We can anticipate many additional applications for Mélange beyond those discussed in this paper. Figure 8 shows an example of a video editing timeline — essentially a 1D visual structure — being folded using our technique. This may be useful for an editor who is synchronizing shots in different parts of a video, or looking to perform color correction between different clips on the timeline. Other potential applications include maps, blueprints, large images, documents, and even user interface components.

CONCLUSION AND FUTURE WORK

We have introduced Mélange, a space-distortion technique that folds 2D space into 3D to guarantee visibility of multiple focus points. The technique supports our definition of large visual space exploration by showing the distance and the intervening context between the foci. This is also what distinguishes it from existing distortion-based techniques. We have presented results from a controlled experiment that confirm the technique's usefulness in this regard.

Space-folding is our first step to supporting exploration of large visual spaces. We have so far focused mostly on the visualization aspects for the Mélange technique and less on the interaction. In the future, we anticipate designing more sophisticated interaction techniques to directly support the exploration design goals. We are also interested in continuing studying its use for visualization of social networks.

ACKNOWLEDGMENTS

Anonymized for double-blind review.

The inspiration for the technique in this paper comes from Frank Herbert's classic science-fiction novel *Dune* from 1965, where interstellar space travel is performed through a process known as "folding space". Here, specially trained navigators with prescient abilities utilize the influences of a special spice known as *Mélange* to traverse fold-space and jump from one end of the galaxy to another.

> *The most precious substance in the universe is the spice Mélange. The spice extends life. The spice expands consciousness. The spice is vital to space travel. The Spacing Guild and its navigators, who the spice has mutated over 4000 years, use the orange spice gas, which gives them the ability to fold space. That is, travel to any part of the universe without moving.*
>
> — Princess Irulan's introduction to the movie adaptation of *Dune* (David Lynch, 1984).

REFERENCES

1. Caroline Appert and Jean-Daniel Fekete. OrthoZoom scroller: 1D multi-scale navigation. In *Proceedings of the ACM CHI 2006 Conference on Human Factors in Computing Systems*, pages 21–30, 2006.

2. Patrick Baudisch and Ruth Rosenholtz. Halo: a technique for visualizing off-screen objects. In *Proceedings of the ACM CHI 2003 Conference on Human Factors in Computing Systems*, pages 481–488, 2003.

3. Frédéric Bourgeois and Yves Guiard. Multiscale pointing: facilitating pan-zoom coordination. In *Proceedings of the ACM CHI 2002 Conference on Human Factors in Computing Systems*, pages 758–759, 2002.

4. Stuart K. Card and David Nation. Degree-of-interest trees: A component of an attention-reactive user interface. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*, 2002.

5. M. S. T Carpendale and Catherine Montagnese. A

framework for unifying presentation space. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 61–70, 2001.

6. George W. Furnas. Generalized fisheye views. In *Proceedings of the ACM CHI'86 Conference on Human Factors in Computer Systems*, pages 16–23, 1986.
7. George W. Furnas. A fisheye follow-up: further reflections on focus + context. In *Proceedings of the ACM CHI 2006 Conference on Human Factors in Computing Systems*, pages 999–1008, 2006.
8. George W. Furnas and Benjamin B. Bederson. Space-scale diagrams: Understanding multiscale interfaces. In *Proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems*, pages 234–241, 1995.
9. Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4(2):114–135, 2005.
10. Sean G. Gustafson and Pourang P. Irani. Comparing visualizations for tracking off-screen moving targets. In *Extended Abstracts of the ACM CHI 2007 Conference on Human Factors in Computing Systems*, pages 2399–2404, 2007.
11. Christopher G. Healey. Choosing effective colours for data visualization. In *Proceedings of the IEEE Conference on Visualization*, pages 263–270, 1996.
12. Nathalie Henry and Jean-Daniel Fekete. MatLink: Enhanced matrix visualization for analyzing social networks. In *Proceedings of INTERACT*, 2007. to appear.
13. Takeo Igarashi and Ken Hinckley. Speed-dependent automatic zooming for browsing large documents. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 139–148, 2000.
14. Pourang Irani, Carl Gutwin, and Xing Dong Yang. Improving selection of off-screen targets with hopping. In *Proceedings of the ACM CHI 2006 Conference on Human Factors in Computing Systems*, pages 299–308, 2006.
15. Bongshin Lee, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. Task taxonomy for graph visualization. In *Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)*, pages 82–86, 2006.
16. Jock D. Mackinlay, George G. Robertson, and Stuart K. Card. The Perspective Wall: Detail and context smoothly integrated. In *Proceedings of the ACM CHI'91 Conference on Human Factors in Computing Systems*, pages 173–179, 1991.
17. Tamara Munzner, François Guimbretière, Serdar Tasiran, Li Zhang, and Yunhong Zhou. TreeJuxtaposer: scalable tree comparison using focus+context with guaranteed visibility. In *Proceedings of ACM SIGGRAPH 2003*, pages 453–462, 2003.
18. Dmitry Nekrasovski, Adam Bodnar, Joanna McGrenere, Francois Guimbretiere, and Tamara Munzner. An evaluation of pan & zoom and rubber sheet navigation with and without an overview. In *Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems*, pages 11–20, 2006.
19. Ken Perlin and David Fox. Pad: An alternative approach to the computer interface. In *Proceedings of Computer Graphics (SIGGRAPH 93)*, pages 57–64, 1993.
20. Catherine Plaisant, David Carr, and Ben Shneiderman. Image browsers: Taxonomy and guidelines for developers. *IEEE Software*, 12(2):21–32, March 1995.
21. Catherine Plaisant, Jesse Grosjean, and Benjamin B. Bederson. SpaceTree: Supporting exploration in large node link tree, design evolution and empirical evaluation. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 57–64, 2002.
22. Catherine Plaisant, Brett Milash, Anne Rose, Seth Widoff, and Ben Shneiderman. LifeLines: Visualizing personal histories. In *Proceedings of the ACM CHI'96 Conference on Human Factors in Computing Systems*, pages 221–227, 1996.
23. Ramana Rao and Stuart K. Card. The Table Lens: Merging graphical and symbolic representations in an interactive focus+context visualization for tabular information. In *Proceedings of the ACM CHI'94 Conference on Human Factors in Computing Systems*, pages 318–322, 1994.
24. George G. Robertson and Jock D. Mackinlay. The Document Lens. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 101–108, 1993.
25. Manojit Sarkar, Scott S. Snibbe, Oren J. Tversky, and Steven P. Reiss. Stretching the rubber sheet: A metophor for visualizing large layouts on small screens. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 81–91, 1993.
26. Garth Shoemaker and Carl Gutwin. Supporting multi-point interaction in visual workspaces. In *Proceedings of the ACM CHI 2007 Conference on Human Factors in Computing Systems*, pages 999–1008, 2007.
27. James Slack, Kristian Hildebrand, and Tamara Munzner. PRISAD: A partitioned rendering infrastructure for scalable accordion drawing (extended version). *Information Visualization*, 5(2):137–151, 2006.
28. Jarke J. van Wijk and Wim A. A. Nuij. Smooth and efficient zooming and panning. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 15–22, 2003.
29. Polle T. Zellweger, Jock D. Mackinlay, Lance Good, Mark Stefik, and Patrick Baudisch. City lights: contextual views in minimal space. In *Proceedings of ACM CHI 2003 Conference on Human Factors in Computing Systems*, pages 838–839, 2003.