

Faster Document Navigation with Space-Filling Thumbnails

Andy Cockburn[†], Carl Gutwin[‡], Jason Alexander[†]

[†]Department of Computer Science
University of Canterbury
Christchurch, New Zealand
{andy, jason}@cosc.canterbury.ac.nz

[‡]Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
gutwin@cs.usask.ca

ABSTRACT

Scrolling is the standard way to navigate through many types of digital documents. However, moving more than a few pages can be slow because all scrolling techniques constrain visual search to only a small document region. To improve document navigation, we developed Space-Filling Thumbnails (SFT), an overview display that eliminates most scrolling. SFT provides two views: a standard page view for reading, and a thumbnail view that shows all pages. We tested SFT in three experiments that involved finding pages in documents. The first study (n=13) compared seven current scrolling techniques, and showed that SFT is significantly faster than the other methods. The second and third studies (n=32 and n=14) were detailed comparisons of SFT with thumbnail-enhanced scrollbars (TES), which performed well in the first experiment. SFT was faster than TES across all document types and lengths, particularly when tasks involved revisitation. In addition, SFT was strongly preferred by participants.

Author Keywords

Scrolling, zooming, thumbnails, document navigation.

ACM Classification Keywords

H5.2 [User Interfaces]: Interaction styles.

INTRODUCTION

Scrolling is the standard interface control for navigating through almost every type of digital document, yet several researchers have observed that it causes a performance bottleneck. Byrne et al [5] state that scrolling is an “obvious case where widget design could make a difference” (p. 550), and O’Hara and Sellen [24] observe that scrolling is “irritatingly slow and distracting” and that users need “quicker, more effortless navigation” (p. 341). They also note that scrolling and paging hamper the user’s incidental memory of the location of document features, impeding their ability to exploit powerful capabilities for spatial memory in support of navigation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2006, April 22-28, 2006, Montréal, Québec, Canada.
Copyright 2006 ACM 1-59593-178-3/06/0004...\$5.00.

Perhaps as a result of these problems, commercial systems offer many scrolling enhancements: rate-based scrolling [35] directly controls scroll speed; dynamic zooming [2] allows users to control the proportion of the document shown in one window; semantic scrolling allows users to jump between document objects such as headings and figures; and thumbnail-enhanced scrollbars show page miniatures beside the regular page view. Research systems have also made improvements to scrolling. For example, Igarashi and Hinckley [18] introduced ‘speed-dependent automatic zooming’ (SDAZ), which eliminates motion blur by automatically zooming out as the scroll rate increases. Sun and Guimbretiere [30] further enhanced SDAZ by combining it with ‘rapid serial visual presentation’ (RSVP) to replace scrolling with page flipping at high scroll speeds.

These commercial and research systems do not change the basic idea of scrolling; rather, their aim is to allow users to scroll more quickly. Even the tools based on zooming and thumbnails are tightly coupled with scrolling because users must scroll when the thumbnails do not fit within one window.

This paper advocates a simple interface that changes the way that users navigate documents, and eliminates scrolling altogether for medium- and long-distance movement. The new technique is called Space-Filling Thumbnails (SFT). With SFT, users switch between their normal reading view and a thumbnail view in which all of the document pages are scaled and tiled to fit the window. There are two potential benefits of this interface. First, in tasks where the user relies on visual features for target identification, they can see the entire document at once, rather than having to scroll just to see the document’s contents. Second, SFT allows users to develop and exploit a spatial understanding of the document more quickly than with scrolling interfaces. These hypotheses are tested by experiments reported below.

The following section describes our SFT system. We then review related work on scrolling, spatial memory, and scale-based interfaces. After the review we report on three studies of SFT: a comparative study of seven alternative scrolling techniques, and two studies that compare navigation performance with thumbnail-enhanced scrollbars and SFT. The studies show that SFT is faster than all other types of scrolling across a variety of document types and lengths.



Figure 1. The Space-Filling Thumbnails system displaying 150 pages. The page under the cursor is magnified, and the most recently viewed page is highlighted in red.

SPACE-FILLING THUMBNAILS (SFT)

Space-Filling Thumbnails is an approach to document navigation in which users switch between a detailed view of a single page and an overview of the entire document.

Our prototype implementation of SFT is shown in Figure 1. It has two display modes: a page view and a thumbnail view. The page view displays one entire page within the window. The thumbnail view (Figure 1) shows all pages as miniatures, laid out as a space-filling matrix in row-major order: again, there is no scrolling. Users move between page and thumbnail views by pressing the middle mouse button. The mode transition is animated (100ms) to help indicate the relationship between the two views. The current page in the thumbnail view is shaded red, with a 3-pixel-wide red border. Clicking on a thumbnail displays that page in the page view.

Each thumbnail has a one-pixel grey border for delineation, and when thumbnails are smaller than 200 pixels in height an enlarged version (154x200 pixels) follows the user's cursor. The location of each thumbnail in the matrix view remains constant during window resizing, provided the aspect ratio of the parent window remains constant.

The visual distinctiveness of individual thumbnails clearly depends on thumbnail size, raising concerns that performance with SFT will deteriorate as document length increases. As a worst-case scenario, we used SFT to display all 1179 pages of Tolstoy's *War and Peace* (www.

bookstacks.org) on a 1600x1200 screen, which produces 34x44 pixel thumbnails. Even at this small size the thumbnails provide a cue to page layout (examples shown at right).



As stated above, we anticipate two main benefits from SFT. First, in visual search tasks, where the user has little prior knowledge of the location of document features, we believe that SFT will allow faster target acquisition than scrolling systems. The thumbnails allow users to visually scan all pages without manipulating the interface, while scrolling requires that individual pages be moved into the display before the eye can scan them. Second, we anticipate that SFT's thumbnail layout will allow users to more quickly construct and exploit a spatial understanding of their documents.

Standard zooming tools, including those of Microsoft Word and Adobe Reader, can produce similar views to SFT's matrix page layout. SFT's important difference, however, is that it constrains users to *only* full-page or space-filling views, and this constraint is intended to promote spatial learning because page locations remain constant. Word and Reader, in contrast, support matrix views as part of a zoom-based continuum that does not provide spatial constancy.

RELATED WORK

Three areas of related work are important for this investigation: scrolling interfaces, spatial location memory, and scale-based zooming and overview+detail techniques.

Scrolling Interfaces

Scrolling has been a navigation method since the earliest visual editors [31], and there have been many studies investigating augmented scrollbars (e.g., [4, 15]) and scroll-based reading (e.g., [17, 20, 22, 27]). The most relevant work for this investigation, however, is on scroll-based target acquisition.

Target acquisition in scrolling tasks means manipulating the scroll interface to move a particular page of the document into the main view. Current techniques use one of two basic control functions for scroll movement in acquisition tasks [35]: position control (also called 'zero-order' control) maps the user's input signals directly to document location, and rate control (also called 'first-order' control) maps the user's input to the rate of movement (i.e., the speed). Hinckley et al [16] showed that Fitts' law [12] accurately models scroll-based target acquisition with both position control (using a scroll wheel) and rate control input.

When scrolling to off-screen targets, users encounter a trade-off between increasing scroll speed and their ability to identify targets in the underlying information space. Faster scrolling potentially allows more rapid acquisition, but it causes motion blur and reduces the time that the target is on-screen, both of which hinder target identification. Igarashi and Hinckley [18] proposed 'speed-dependent

automatic zooming' (SDAZ) to ease these problems. SDAZ couples the document's zoom level to the scroll speed—the document automatically zooms out as the scroll speed increases, reducing motion blur and increasing on-screen time. Cockburn & Savage [8] showed that SDAZ is faster than position-controlled scrollbars [9].

Sun and Guimbretiere [30] modified Cockburn & Savage's systems to create 'Flipper', which switches from SDAZ scrolling to 'rapid serial visual presentation' (RSVP) [2] when the system zooms out far enough that more than one page is visible. Flipper's RSVP mode causes whole pages to be 'flipped' onto the screen for a short period of time, eliminating motion blur. In their evaluation, Flipper outperformed both SDAZ and thumbnail-enhanced scrollbars. Finally, Spence et al [28] compared image retrieval with three different presentation modes: static, which displayed all 64 images on the screen at once for time period t ; mixed RSVP, which presented four images for $t/16$ time, and slide-show RSVP, which showed one image for $t/64$ time (equivalent to Flipper's page-display mode). Participants preferred mixed-mode to static-mode, and performance measures also supported mixed-mode presentation.

Spatial memory

Spatial memory research is pertinent to our investigation because it has been shown that users regularly form a spatial understanding of documents [24], and because scrolling systems can impede that understanding [25].

Several studies agree that there is a strong relationship between spatial aptitude and performance across distinct styles of computer use: for example, text editors [11], computer games [13], and file managers [33]. Others have shown that a lack of positional constancy in icon and menu-item placement harms performance [23, 32].

Users' capacity for exploiting spatial memory is successfully demonstrated by evaluations of the Data Mountain [26], which allowed users to create spatial arrangements of web-page thumbnails on a receding 3D plane. An initial evaluation showed that the spatial layout allowed faster web-page retrieval than a scrolling list [26], but it was a follow-up evaluation [10] that revealed the pervasiveness of the spatial cue. Three months after creating their initial thumbnail layout, participants were able to retrieve web-pages with only a slight reduction in performance; more surprisingly, performance was similar (after a temporary performance drop) when the visual thumbnails were replaced with blank images.

These studies show that spatial location can be a powerful cue for information retrieval. Jones and Dumais [19], however, issue cautions on over-reliance on spatial cues.

Spatial location memory is exploited to different degrees in current scrolling techniques. The scroll thumb of a standard scrollbar passively portrays relative spatial location,

allowing users to learn spatial information if they attend to the cue. Traditional scrolling (by dragging the thumb) actively communicates location, but rate-based techniques do not because they encourage users to focus on the document surface rather than on the thumb's spatial cue.

Scale-based interfaces (zooming and overview+detail)

Several types of interfaces use changes in scale to assist navigation. Two of these are zooming (temporal scaling) and overviews (scaling of a separate view of the space). For example, the overview of Space-Filling Thumbnails is similar to Bederson's photograph image layout in PhotoMesa [1]. PhotoMesa allows several hundred image thumbnails to be shown in one window without scrolling, using a 'quantum treemap' layout to depict how photographs are stored in different folders.

Zooming requires a temporal separation between pre- and post-zoom display states. An alternative approach is to spatially separate focus and context using an 'overview+detail' display [6]. This approach has been extensively studied in previous work. For example, Hornbaek & Frokjaer [17] show that sidebar thumbnail interfaces can aid document comprehension, and O'Hara, Sellen and Bentley [25] show that an overview+detail approach can improve learning of a document's spatial properties, compared to traditional scrolling systems.

Semantic zooming is another technique that can be used to reduce the readability problems of miniaturized documents. Woodruff et al. [34] demonstrate the advantages of a thumbnail enhancement technique that enlarges semantically important text such as headings and search-terms. This technique was further developed in the overview+detail 'Popout Prism' system [29], which uses a similar enlargement technique to highlight search terms in both the overview and detail region. Lam and Baudisch [21] also show performance advantages for their 'summary thumbnails' system, which produces readable text fragments while largely preserving the original layout of web pages. Finally, thumbnails can be augmented to reveal other semantic properties, such as 'dog-ears' that show the number of times a page has been visited [7].

EXPERIMENTS WITH SFT

We carried out three experiments to test SFT's performance for document navigation. The first study compared several techniques to gain an initial understanding of how SFT differed from current scrolling methods. The second and third studies carried out a more detailed comparison of SFT and the best competitor chosen from the first experiment.

Experimental Tasks

Selecting suitable tasks for evaluating scrolling systems is complicated by the wide variety of navigation activities for which scrolling is used. These tasks include searching for a particular document feature (e.g., a section heading, table,

or figure), quickly referencing another part of a document then returning to the original location, browsing, ‘nudging’ the document to centre text in the screen, and many more.

Our evaluations focus on SFT’s support for visual and spatial searches. An example visual task is “The page with a big table at the top”, supplemented with “around half way through the document” when spatial understanding is available. It seems reasonable to suspect that these tasks are indicative of many scrolling activities.

To emulate visual and spatial tasks, all experiments involved repeatedly finding the same set of pages within documents. When finding a page for the first time, the participant’s search is purely visual, but as they repeatedly search for the same item they are better able to exploit spatial awareness of the target’s location.

Scrolling Techniques

The diversity of related work means that there are many competing interfaces that could be included in an evaluation. Candidate systems include traditional scrollbars, thumbnail-enhanced scrollbars, rate-based scrolling, speed-dependent automatic zooming, Flipper, and multi-page variants of Flipper.

To gain an initial understanding of the different techniques’ performance, and to provide an empirical basis for selecting interfaces for a full-scale evaluation, we constructed several candidate systems (described below) and compared their performance in an initial experiment. All of the systems were constructed from the same C++/OpenGL core program, allowing identical display frame rates, data logging, and task cueing facilities. All interfaces were controlled by a Logitech three-button mouse.

Scrollbars and Thumbnail-Enhanced Scrollbars (TES)

Traditional scrollbars were included because they remain the standard interface for document navigation. The scrollbar interface was similar to normal scrollbars, consisting of up/down arrows at the top and bottom of the scroll trough, an active trough for paging, and a scroll thumb that both shows and controls the document region that is displayed in the window.

In our Thumbnail Enhanced Scrollbars (TES) system, the thumbnail panel showed ten pages in one column, which is similar to the default setting in Adobe Reader. Our implementation behaved like standard desktop applications: the thumbnail window followed the user’s scrolling actions in the main window, but the main window did not follow actions in the thumbnail overview until a particular thumbnail was selected.

Rate-based scrolling (RBS)

The rate-based scrolling interface was visually identical to the scrollbar interface, but the scrollbar was simply a passive location display that could not be used to control

scrolling. Like many Microsoft applications, the scroll rate was controlled by vertical dragging with the middle mouse button. During rate-based scrolling, the cursor  changed to the mode indicator shown at right. There was a linear relationship between vertical drag distance and scroll speed, to a maximum document velocity of 47cm/sec at a drag displacement of 170 pixels [9] (one page is 27cm).

Speed-Dependent Automatic Zooming (SDAZ)

SDAZ scrolling is rate-based scrolling with the addition of automatic zooming to reduce motion blur and increase the on-screen time of document features. Users controlled and viewed the SDAZ system identically to rate-based scrolling, except that the underlying document’s zoom level was automatically adjusted as the scroll speed changed. We used the SDAZ calibration settings determined in [9].

RSVP-based Flipping and Multi-Page Flipping (MRSVP)

We implemented two page-flipping interfaces based on RSVP. Sun and Guimbretiere’s Flipper blends SDAZ and RSVP to produce an interface that scrolls with SDAZ behaviour until a threshold velocity, beyond which RSVP page-flipping is used. To isolate the performance of RSVP, our systems eliminated the initial SDAZ scroll mode.

The two RSVP-based systems display either one page (RSVP) or four pages (MRSVP) at a time (Figure 2). Our RSVP system is equivalent to the ‘slide-show’ mode of Spence et al’s study [28]. MRSVP is equivalent to ‘mixed-RSVP’, which out-performed ‘slide-show’ mode [28].

Scrolling is controlled by dragging with the middle button. A passive scrollbar displays document location. To help users predict the pace of page flipping, an animated cursor shows a filling circle which indicates the remaining time before the next page flip. This feedback is particularly important for the first few flips as it provides the sole cue to the page-flipping pace. The maximum page-flipping rate is ten flips per second (attained at a drag displacement of 170 pixels), and is based on Spence et al’s measure of the point at which image identification deteriorates markedly.

In MRSVP, scrolling is initiated in two ways: the user can either drag the middle mouse button (like RSVP) or they can click with the left mouse button. Either action causes the display to rapidly animate between the normal single-page view and the quadrant view shown in Figure 2 (see video figure). Middle-button dragging then controls the rate of page flipping, with all four pages replaced simultaneously. The user returns to the single-page view by clicking the target page in the quadrant view with the left mouse button.

EXPERIMENT ONE: SEVEN METHOD SCROLL-OFF

We conducted an initial study to compare the performance of, and preferences for, these seven different scrolling systems. Thirteen volunteer Computer Science graduate

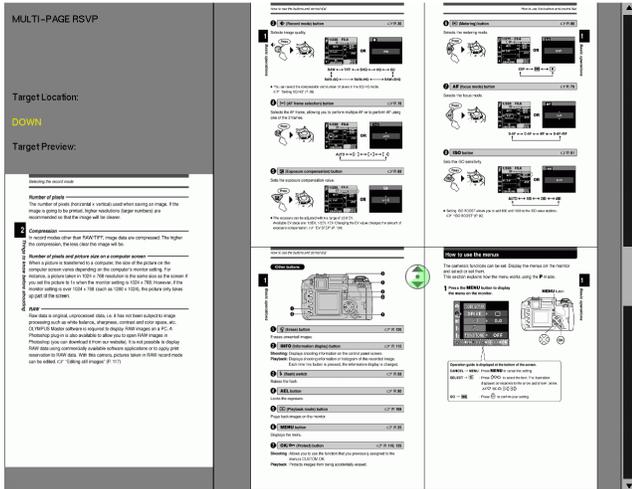


Figure 2. The experimental interface in MRSVP mode. The task cueing interface is on the left.

students took part in the study. The participants' tasks involved navigating to specific pages in thirty-page journal papers: first finding a page based on a large visual preview of the page, then immediately repeating the same task to emulate task performance when users have a strong spatial memory of the page's location.

The task cueing interface

Figure 2 shows the experiment system with the MRSVP interface. Each task began with the display of a new target page in the task-cueing sub-window at the lower left of the screen. The target page preview was sufficiently large to make the body text legible (420x594 pixels). The sub-window also showed the navigation direction to the target from the starting page (up or down). Page numbers were removed from the page images. Participants viewed the target preview for seven seconds, (during which time the main task window was inactive), and then began the task. The preview window remained visible through the task.

Task completion was automatically detected by software when more than one-third of the target page was stationary on the screen with the mouse button up (only in page view with SFT). The system then displayed a "Show next task" dialogue, which allowed participants to rest momentarily before starting the next task. A video of an experimental session can be seen in the video figure for the paper.

Experiment One: Design and Procedure

The experimental design is a 7x2 repeated measures analysis of variance (ANOVA) with two factors: *interface-type* (scrollbars, TES, RBS, SDAZ, RSVP, MRSVP, and SFT), and *task-type* (visual and spatial search). The dependent measure was task completion time.

Participants had a one-minute free practice session with each interface before completing four training tasks and ten testing tasks with that interface (five tasks for each level of *task-type*). The order that each participant used the

interfaces was balanced using an incomplete Latin square. Eight different 30-page journals were prepared for the experiment: one was used for all training tasks, and the remaining seven were used for the testing tasks, one per interface, in the same order by all participants. All of the documents (in all of the experiments) were converted to greyscale images.

The five target pages for each document were randomly selected from the first or last nine pages of the document. Starting locations for these tasks were then randomly selected to be between 15-18 pages from the target.

The ten testing tasks with each interface were presented in pairs. The first task of each pair was a 'visual search' task, followed by a 'spatial search' task that repeated exactly the same start and target locations. Participants were informed that tasks would be immediately repeated, and they were encouraged to try to memorise the target's location to optimise their performance in the repeated tasks. Training tasks familiarised the participants with this procedure in each interface. The explicit request to memorise the target was intended to encourage a strong spatial understanding of the document. The repeated tasks ask the question "if the user had a good spatial memory of page locations, how would the interface exploit that memory?"

On completing all tasks with all of the interfaces, participants rated the efficiency of each interface for the visual and spatial tasks (using a five-point Likert scale), and provided an overall preference ranking.

Experiment One: Apparatus

The experiment ran on Intel Pentium 4 2.8GHz computers, equipped with 1GB RAM and NVIDIA GeForce FX5200 graphics cards connected to 19-inch CRT displays at 1280x1024 resolution and 75Hz refresh rate. The task-cueing region of the window was 420x1024 pixels and the scrolling interfaces were displayed in the remaining 860x1024 region of the display (Figure 2). Input was received through a Logitech three-button opto-mechanical mouse with a one-to-one control-display gain setting.

Experiment One: Results

Of the 910 tasks across all conditions, six outlier tasks, which took more than thirty seconds to complete, were discarded: three with TES, and one each with scrollbars, MRSVP and SFT. The participants completed all other tasks rapidly, with an overall mean of 7.3 seconds (standard deviation 3.2).

Across both visual and spatial tasks, SFT's mean task time was fastest at 4.1 seconds (s.d. 3.0), followed by MRSVP (6.1, s.d. 1.8) and TES (6.4, s.d. 2.6), then scrollbars, RSVP, SDAZ and RBS (slowest at 10.4, s.d. 2.9), giving a significant main effect for interface-type: $F_{6,72}=24.7$, $p<.01$. Figure 3 summarises these results. A post-hoc Tukey test ($\alpha=.05$) gives an honest significant difference of 2.47s.

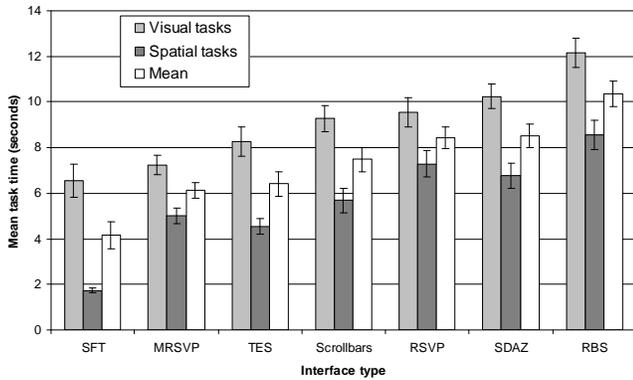


Figure 3. Mean times for visual and spatial tasks in Experiment One. Error bars show mean \pm one std error.

As expected there was a significant main effect for task-type, with ‘spatial’ tasks (mean 5.7 seconds, s.d. 2.6) being completed more quickly than ‘visual’ tasks (mean 9.0, s.d. 2.7): $F_{1,12}=132.6$, $p<.01$

There was a significant interaction between factors interface and task type ($F_{6,72}=2.6$, $p<.05$). One cause of the interaction, shown in Figure 3, is the rapid completion of ‘spatial’ tasks with SFT (mean 1.7, s.d. 0.3). TES allowed the second fastest mean task completion time for spatial tasks (mean 4.5, s.d. 1.2), but at more than two and a half times SFT’s mean. Another factor contributing to the interaction is the relatively poor performance of RSVP and MRSVP in ‘spatial’ tasks. Participants’ comments helped explain this effect—the RSVP interfaces demanded constant and close attention, reducing the users’ ability to attend to spatial cues while completing tasks.

The participants’ subjective assessments of the interfaces provide interesting insights into their use of the systems. Although the MRSVP system allowed the second fastest mean task time, it was unpopular, with a median 6th ranking among the seven interfaces. RSVP also had a median 6th ranking. Participants’ comments showed a clear dislike of RSVP’s high visual demands with statements including “Hard to scan” and “It felt like I was about to get a headache... or a seizure”.

SFT was the clearly preferred interface with a median 1st ranking (ten of thirteen ranked it first). TES was ranked second, followed by SDAZ, scrollbars and RBS. Comments about SFT were almost uniformly positive: “Awesome – so easy!” and “Very easy to remember rough location of thumbnail”. There was, however, some conjecture about SFT’s limitations: “More pages means thumbs need to be smaller, which means [it would be] harder to find pages”.

Mean responses to the 5-point Likert scale question “The interface was efficient for the task” (1-disagree, 5-agree) ranged from 2.8 (s.d. 1.2) for RSVP in the visual search task to 4.9 (s.d. 0.3) for SFT in the spatial task. SFT received the highest mean rating in the visual search task at 4.2 (s.d. 0.9). TES received the second-highest mean scores

at 3.8 (s.d. 0.6) and 4.0 (s.d. 0.8) in the visual and spatial tasks respectively. The Likert responses showed significant differences between interfaces in both visual (Friedman $\chi^2=17.9$, $p<.01$) and spatial tasks ($\chi^2=30.2$, $p<.01$).

Experiment One: Discussion

The experiment produced encouraging results. In 30-page journal papers, SFT allowed visual search tasks to be completed in 66% of the time of the next fastest interface (MRSVP), and in spatial tasks the difference was even greater (38% of runner-up TES). Subjective responses confirmed that SFT is both efficient and popular, with most participants ranking it as their favourite.

The results for both of the RSVP interfaces are interesting as there is a conflict between its fast task completion (MRSVP was second fastest overall) and its poor subjective assessment. Our results agree with Flipper’s evaluation [30], which showed the efficiency of RSVP-based scrolling, but we additionally show that this can be further improved by using Spence et al’s multi-page design. Despite its efficiency, our subjective responses strongly suggest that RSVP-based scrolling will be unpopular due to high demands for visual attention.

Previous evaluations have shown that SDAZ significantly outperforms traditional scrollbars and rate-based scrolling [8, 9]. In this experiment, however, scrollbars had a faster mean task time than SDAZ (although the difference is outside the Tukey honest significant difference for post-hoc comparison). We have two possible explanations for failing to replicate this result. First, when using scrollbars the participants directly interact with the scroll-thumb, which provides a spatial cue to document location. SDAZ, in contrast, does not involve direct interaction with any spatial cue, possibly explaining the better performance with scrollbars in ‘spatial’ tasks. Second, the training periods were short, allowing participants only one minute of free practice followed by four training tasks with each interface. It may be that this time was insufficient for participants to become familiar with the novel behaviour of interfaces such as SDAZ and RSVP. It is worth noting, however, that SFT also provides an unusual interface, yet with equitable training, participants were able to use it efficiently.

One of the main objectives of the first experiment was to decide which scrolling interface would provide the best control comparison for further evaluations of SFT. We selected thumbnail enhanced scrollbars (TES) for three reasons: first, TES is implemented in several contemporary user interfaces, including Microsoft Word/PowerPoint and Adobe Reader; second, although MRSVP had a slightly lower mean task completion time than TES, it was uniformly disliked; third, results indicated that TES better supports the users’ spatial memory than MRSVP.

EXPERIMENT TWO: SFT VERSUS TES

Experiment One used 30-page journal papers and it artificially ‘implanted’ the participants’ spatial knowledge of tasks by explicitly asking them to memorise page locations. The second experiment used documents of differing lengths and types, and it does not explicitly implant spatial knowledge. It compares how well thumbnail-enhanced scrollbars and space-filling thumbnails support visual search tasks, incidental learning of spatial locations, and exploitation of spatial memory. It also examines the participants’ subjective preferences and workload assessments of the interfaces.

The task-cueing interface and apparatus were identical to that used in Experiment One. There were 32 participants: all undergraduate Computer Science students with an age range of 18 to 36 years (mean 22 years).

Experiment Two: Design

The experiment is designed as a $2 \times 3 \times 3$ repeated measures ANOVA with the following factors:

- *Interface-type*: TES and SFT.
- *Document type*: conference paper (10 pages), journal paper (30 pages), product manual (150 pages).
- *Search iteration*: first, second, third search.

The *Document type* factor is intended to expose differences between the interface types across different document lengths. Two different conference papers, journal papers and manuals were used, and the interface order with the documents was rotated across participants. The conference papers were both extracted from the proceedings of CHI 2005. The journal papers were taken from ACM ToCHI, and the manuals were for Sony and Olympus digital cameras. To remove the possibility of ‘cheating’ in the visual search tasks, page numbers, tables of contents and indexes were removed from all documents.

The *Search iteration* factor is intended to expose differences between the two interfaces in the degree to which they allow users to develop and exploit a spatial understanding of the documents. In theory, scrollbars provide a strong spatial cue through the location of the scroll-thumb in the scroll-trough, but we wanted to observe whether this cue translates into effective spatial learning. SFT, however, demands that users navigate through a 2D spatial layout of thumbnails, and we therefore predict comparatively high levels of incidental spatial learning. The participants were not informed that the same targets would be re-used, the intention being to investigate incidental rather than intentional spatial learning.

The primary dependent variable is task completion time. Software controlled the participants’ exposure to the training and experimental tasks, cueing the tasks, logging their actions, and presenting dialogues to gather subjective assessments at the end of all tasks with each interface. Participants subjectively assessed workload with each

interface using the NASA Task Load Index worksheets [14]. They also used 5-point Likert scales to state how well the interfaces supported the tasks and their spatial memory. Finally, they stated which interface they preferred overall.

Experiment Two: Procedure

Participants completed all tasks with one interface before proceeding to the other. Interface order was balanced with half using SFT first, and half TES first. Participants also completed all tasks with the shorter document types before proceeding to longer ones.

Three distinct target pages were automatically generated for each interface and document combination. The targets were always in the middle third of the document, and the starting locations for the tasks were randomly generated to be between 23-33% of the document length above or below the target. Participants were unaware of these constraints on the target and start locations.

Experiment Two: Results

The mean task time for the 1728 tasks across all conditions was 8.5 seconds, s.d. 9.2 seconds—the high standard deviation is due to the wide range of times between 10-page and 150-page documents. Tasks were capped at one minute, resulting in 26 incomplete tasks with TES and 25 with SFT—data from incomplete tasks were discarded.

Across all page lengths and search iterations, tasks were completed significantly faster with SFT (mean 6.5, s.d. 7.8) than TES (10.4, s.d. 10.0): main effect $F_{1,31}=66.6$, $p<.01$. As expected, both other factors also showed significant main effects. Task times increased with document length from 2.8s with ten pages, through 4.8s with thirty pages, to 17.7s in 150 page documents: $F_{1,62}=348.7$, $p<.01$. Task times decreased for subsequent search iterations ($F_{2,62}=98.1$, $p<.01$). Figure 4 summarises the results.

There were significant interactions between interface type and page length ($F_{2,62}=12.4$, $p<.01$) and between interface type and search iteration ($F_{2,62}=3.6$, $p<.05$). The interface by

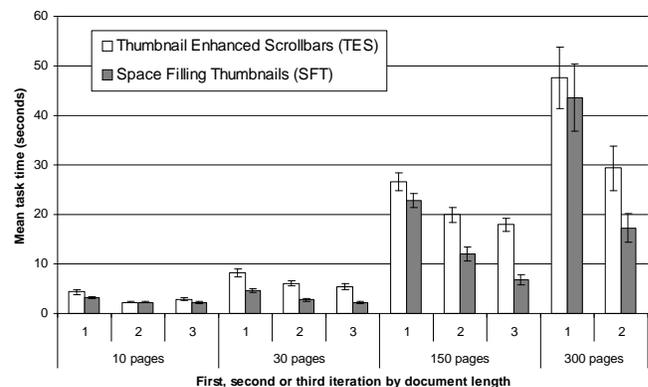


Figure 4. Mean times for TES and SFT by document-length and search iteration in Experiments 2 and 3. Error bars show mean ± 1 standard error.

page-length interaction is explained by SFT’s increased comparative efficiency in longer documents. SFT’s mean task times were 20% lower than TES in ten-page documents, but in 30-page and 150-page documents this value was 52% and 35% respectively. The interface \times iteration interaction is also explained by the increasing comparative benefits of SFT on successive iterations, with task time reductions on the first, second, and third iterations increasing from 22% through 40% to 58%. Both of these interactions support the hypothesis that SFT’s thumbnail layout enables users to better use their spatial memory.

Subjective responses also favoured SFT. The participants answered four Likert-scale questions (1 for disagree, 5 for agree) after completing all tasks with both interfaces. Results are summarized in Table 1.

Responses to the NASA Task Load Index questions showed little difference in subjective workload between the interfaces, with no significant differences for any of the 5-point Likert-scale measures. Finally, twenty-two of the participants stated that they preferred SFT, and ten preferred TES; the difference is significant ($\chi^2=3.8, p<.05$).

The participants’ comments focussed on three issues: the difficulty of the tasks as the page length increased (with both interfaces); the reduction in difficulty for subsequent searches, particularly with SFT; and problems with SFT’s cursor-following thumbnail magnifier. We had anticipated the first two comments, but not the third. One participant summarised the problem as “the pop up view of the thumbnail got in the road of scanning the thumbnails”. We observed several participants trying to scan the unmagnified thumbnails with their eyes, using the cursor to keep track of the current thumbnail. This style of use was frustrated by the cursor-following magnified view, which obscured part of the original thumbnail as soon as the cursor reached it. This problem would be easy to rectify by using PhotoMesa’s strategy of only revealing the magnified overview when the cursor is stationary.

EXPERIMENT THREE: PROBING SFT’S WEAKNESS

The third experiment probes SFT’s potential weaknesses by comparing SFT and TES when navigating through a 300 page novel. The long document produces small thumbnails (45x63 pixels, a size that would allow 650 pages in a 1600x1200 display) that lack distinctive features. It was necessary to separate experiment two and three because pilot studies showed that the high mean task times would cause excessive participation time and fatigue. Although real users would almost certainly use Search tools for navigating through long documents, these tasks are intended to expose interface differences in extreme cases of use. Experimental concerns are discussed later in the paper.

There were 14 participants, all undergraduate Computer Science students, with an age range from 18 to 36 (mean 21). Participation lasted approximately 30 minutes.

Question	SFT	TES	p
	mean (s.d.)	mean (s.d.)	
SFT/TES was efficient for the tasks	3.9 (1.0)	3.4 (1.0)	<.05
I liked using SFT/TES	4.0 (1.0)	3.4 (0.9)	<.05
It was easy to return to pages	3.7 (1.4)	3.2 (1.1)	0.1
I could remember page locations	4.1 (1.3)	3.1 (1.1)	<.01

Table 1. Questionnaire responses and significance values (Wilcoxon Signed Ranks tests).

Experimental tasks, training and software were the same as Experiments One and Two.

The experiment is designed as a 2 \times 2 repeated measures ANOVA for factors *interface-type* (SFT and TES) and *search-iteration* (first ‘visual’ search, second ‘spatial’).

Leo Tolstoy’s *Anna Karenina* (www.bookstacks.org) was used for all tasks. Page numbers and tables of contents were removed. Four targets were randomly generated according to the same rules as Experiment Two (targets in the middle third and starting locations 23-33% of the document length above or below). The participants’ tasks with each interface involved finding the first target, then finding the second target (‘visual’ searches), then repeating the first and second tasks (‘spatial’ searches). Participants were not informed that the tasks would be repeated.

Experiment Three: Results

Tasks were limited to four minutes, resulting in six incomplete tasks: three with SFT and three with TES. The mean time to complete the tasks was 34.5s, with a standard deviation of 22.9. The high variance in task completion time was anticipated due to the substantial task difficulty. Therefore, the data in the analysis of variance are log-transformed to stabilise the variance.

There is a significant main effect for interface type, with tasks being completed more quickly with SFT (mean 30.5, s.d 23.4) than TES (mean 38.4, s.d. 22.1): $F_{1,13}=7.6, p<0.05$. There is also a significant main effect for search iteration ($F_{1,13}=51.9, p<0.01$), with first searches taking an average of 45.6s (s.d. 24.1) and second searches 23.3 (s.d. 15.2). These results are summarised in Figure 4.

As in experiment two, there is a significant interaction between interface type and search iteration ($F_{1,13}=7.1, p<0.05$) which is caused by the marked reduction of task time with SFT in second iterations: it improves from 92% of TES’s task time in the first iteration to only 59% of TES’s time in the second iteration.

DISCUSSION

In all of our three experiments, SFT was significantly faster than its competitors. In experiments two and three, which directly compare performance using SFT against the ‘best of the rest’ interface TES, SFT outperformed TES by up to 58%, and was strongly preferred by participants. The results support the hypotheses that SFT allows faster visual search than other scrolling methods and that it better enables users

to exploit their spatial memory. SFT's performance benefits also appear to be robust to variations in document type and length – from ten page conference papers with many distinct visual features to 300 page novels with few obvious landmarks. In sum, SFT appears to be a very promising technique for improving document navigation.

In the next paragraphs we consider four issues related to the validity and generality of our results: the realism of the tasks, the realism of the cueing mechanism, potential problems in deploying SFT in real-world interfaces, and the problem of changing document sizes.

Realism of the tasks. The tasks used in our experiment are necessarily artificial, particularly with the 300 page document. While it is reasonable to imagine users navigating between regions in conference papers, journal articles and manuals (our 10, 30 and 150 page documents), it is unlikely that they would need to do so in a novel, and it is even less likely that they would use scrolling to do so; 'Find' utilities would be much more likely. However, our explicit objective in Experiment Three was to probe SFT's weaknesses, and even here it outperformed TES.

Realism of the cueing interface. There are experimental concerns regarding our task cueing mechanism, which presented a scrolling direction (up/down) and a picture of the target page. Despite the fact that our cueing mechanism is unlikely to accurately reflect a user's natural conception of navigational goals, we see no reason to believe that the cueing mechanism is biased towards SFT.

Deploying SFT in a real-world system. There are several potential concerns in generalising these results to large-scale deployments of SFT in everyday desktop environments. When completing tasks with SFT, the thumbnail layout provided the sole interface control for completing the tasks. Yet if traditional scrollbars had also been available, it is possible that participants would have used them because of their long-term 'conditioning'. The interface inertia of scrollbars may hinder wide-scale use of SFT if it were available in standard desktop systems. There are also questions regarding users' acceptance of SFT's strong modal separation between thumbnail and page views: while participants liked it in our 'page-chasing' tasks, it is not clear that they would feel the same in everyday document browsing. In further work we intend to investigate how participants choose to use SFT when it is available as a functionally rich document browsing system.

This third concern leads to questions of SFT's commercial deployment. A radical deployment would constrain scrollbars to movement within the current page, relying on simple interface actions such as key-presses or mouse-clicks for movement to the next/previous page and on SFT's thumbnails for all long distance 'scrolling'. Several systems already make such a constraint on scrollbar range: e.g. the GSView PostScript previewer (ghostgum.com.au).

A less radical implementation would add SFT's functionality as yet another tool in the suite of options already available for document navigation. This option could be easily incorporated into current systems such as Adobe Reader and Microsoft Word because they already support thumbnail views like SFT's, but they do not yet include an interface shortcut to immediately access a space-filled representation. It remains to be seen whether users would adopt the SFT view, or whether the inertia of scrollbars would prevail.

Changing the number of pages. The documents used in the evaluations were not editable, producing constant thumbnail locations in SFT. While these static locations generalise to document browsers (such as Adobe Reader), there are risks in generalising to editors because the thumbnails will move as material is added and deleted. To begin investigating this effect we included two final tasks with each interface at the end of Experiment 3 (not fully reported due to space constraints). In these tasks, the participants were informed that editing had changed the document content by up to five pages. Analysis of these tasks showed no significant difference between SFT and TES ($p=0.5$). Although further work is needed, at present we see no reason to suspect that SFT will hinder target acquisition when pages move.

CONCLUSION

We have presented the design and evaluation of Space Filling Thumbnails (SFT), a simple interface technique that allows rapid document navigation. SFT replaces scrollbar-based document movement with two views: a single-page view, and a matrix-layout thumbnail view that shows all pages at once in miniature.

The first evaluation compared user performance and preferences in document navigation tasks across seven different systems. SFT was the most preferred system and it out-performed all others. A system based on Multi-page RSVP was second fastest, but unpopular due to its high demands for visual attention. Thumbnail-enhanced scrollbars also allowed rapid task completion and was ranked second in preference.

The second and third evaluations directly compared SFT with TES, demonstrating that SFT's performance and preference advantages are robust across different types and lengths of document, including a 300-page novel. Results suggest that SFT's thumbnail layout improves both the user's ability to visually scan targets and their spatial memory for target locations.

In future work we will investigate how users employ SFT when they have the freedom to choose between different means for document navigation. We will also inspect the performance improvements enabled through semantic enhancements to the thumbnails such as those provided by the Popout Prism [29].

REFERENCES

1. Bederson, B. (2001). PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps. *Proc. UIST01*, 71-80.
2. Bederson, B. and Hollan, J. (1994). Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. *Proc. UIST'94*, 17-26.
3. de Bruijn, O. and Spence, R. (2000). Rapid Serial Visual Presentation: A Space-Time Trade-Off in Information Presentation. *Proc. AVI2000*, 189-192.
4. Byrd, D. (1999). A Scrollbar-based Visualization for Document Navigation. *Proc. Digital Libr. '99*, 122-129.
5. Byrne, M., John, B., Wehrle, N. and Crow, D. (1999). The Tangled Web We Wove: A Taskonomy of WWW Use. *Proc. CHI'99*, 544-551.
6. Card, S. Mackinlay, JD and Shneiderman, B. (1999). *Readings in Info. Visualization*. Morgan-Kaufmann.
7. Cockburn, A. and Greenberg, S. (2000). Issues of Page Representation and Organisation in Web Browser's Revisitation Tools. *Aus. J. of Info. Sys.*, 7 (2), 120-127.
8. Cockburn, A. and Savage, J. (2003). Comparing Speed-Dependent Automatic Zooming with Traditional Scroll, Pan and Zoom Methods. *Proc. British HCI*, 87-102.
9. Cockburn, A., Savage, J. and Wallace, A. (2005). Tuning and Testing Scrolling Interfaces that Automatically Zoom. *Proc. CHI'05*, 71-80.
10. Czerwinski, M., van Dantzich, M., Robertson, G. and Hoffman, H. (1999). The Contribution of Thumbnail Image, Mouse-Over Text and Spatial Location Memory to Web Page Retrieval in 3D. *Proc. Interact'99*, 163-170.
11. Egan, D. and Gomez, M. (1985). Assaying, Isolating, and Accommodating Individual Differences in Learning a Complex Skill. in Dillon, R. ed. *Individual Differences in Cognition*, 173-217.
12. Fitts, P. (1954). The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *J. Experimental Psychology*, 47, 381-391.
13. Gagnon, D. (1985). Videogames and Spatial Skills: An Exploratory Study. *Educational Communication and Technology*, 33 (4), 263-275.
14. Hart, S. and Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. in Hancock, P.M., ed. *Human Mental Workload*, 139-183.
15. Hill, W. and Hollan, J. (1992). Edit Wear and Read Wear. *Proc. CHI'92*, 3-9.
16. Hinckley, K., Cutrell, E., Bathiche, S. and Muss, T. (2002). Quantitative Analysis of Scrolling Techniques. *Proc. CHI'2002*, 65-72.
17. Hornbaek, K. and Frokjaer, E. (2003). Reading Patterns and Usability in Visualizations of Electronic Documents. *ACM Trans. CHI*, 10 (2), 119-149.
18. Igarashi, T. and Hinckley, K. (2000). Speed-dependent Automatic Zooming for Browsing Large Documents. *Proc. UIST'00*, 139-148.
19. Jones, W. and Dumais, S. (1986). The Spatial Metaphor for User Interfaces: Experimental Tests of Reference by Location versus Name. *ACM Trans. OIS*, 4 (1), 42-63.
20. Klein, C. and Bederson, B. (2005). Benefits of Animated Scrolling. *Ext. Abstracts CHI'05*, 1965-1968.
21. Lam, H. and Baudisch, P. (2005). Summary Thumbnails: Readable Overviews for Small Screen Web Browsers. *Proc. CHI'05*, 681-690.
22. Mills, C. and Weldon, L. (1986). Reading text from computer screens. *ACM Comp. Surveys*, 19, 329-358.
23. Moyes, J. (1994). When Users Do and Don't Rely on Icon Shape. *Ext. Abstrs. CHI'94*, 283-284.
24. O'Hara, K. and Sellen, A. (1997). A comparison of reading paper and on-line documents. *Proc. CHI'97*, 335-342.
25. O'Hara, K., Sellen, A. and Bentley, R. (1999). Supporting Memory for Spatial Location While Reading from Small Displays. *Ext. Abs. CHI'99*, 220-221.
26. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. and van Dantzich, M. (1998). Data Mountain: Using Spatial Memory for Document Management. *Proc. UIST'98*. 153-162.
27. Schwarz, E., Beldie, I. and Pastoor, S. (1983). A Comparison of Paging and Scrolling for Changing Screen Contents by Inexperienced Users. *Human Factors*, 25 (3), 279-282.
28. Spence, B., Witkowski, M., Fawcett, C., Craft, B. and de Bruijn, O. (2004). Image Presentation in Space and Time: Errors, Preferences and Eye-Gaze Activity. *Proc. AVI'04*, 141-149.
29. Suh, B., Woodruff, A., Rosenholtz, R. and Glass, A. (2002). Popout Prism: Adding Perceptual Principles to Overview+Detail Document Interfaces. *Proc. CHI'02*, 251-258.
30. Sun, L. and Guimbretiere, F. (2005). Flipper: a new method of digital document navigation. *Ext. Abs CHI'05*, 2001-2004.
31. Sutherland, I. (1963). Sketchpad: A Man-Machine Graphical Communication System. *AFIPS Spring Joint Computer Conference*, 329-346.
32. Teitelbaum, R. and Granda, R. (1983). The Effects of Positional Constancy on Searching Menus for Information. *Proc. CHI'83*, 150-153.
33. Vicente, K., Hayes, B. and Williges, R. (1987). Assaying and Isolating Individual Differences in Searching a Hierarchical File System. *Human Factors*, 29 (3), 349-359.
34. Woodruff, A., Faulring, A., Rosenholtz, R., Morrison, J. and Pirolli, P. (2001). Using Thumbnails to Search the Web. *Proc. CHI'01*, 198-205.
35. Zhai, S., Smith, B. and Selker, T. (1997). Improving Browsing Performance: A Study of Four Input Devices for Scrolling and Pointing Tasks. *Proc. INTERACT'97*, 286-292.