Using Artificial Landmarks to Improve Revisitation Performance and Spatial Learning in Linear Control Widgets

Md. Sami Uddin
University of Saskatchewan
Saskatoon, Canada
sami.uddin@usask.ca

Carl Gutwin
University of Saskatchewan
Saskatoon, Canada
gutwin@cs.usask.ca

Alix Goguey
University of Saskatchewan
Saskatoon, Canada
alix.goguey@usask.ca

ABSTRACT

Linear interface controllers such as sliders and scrollbars are primary tools for navigating through linear content such as videos or text documents. Linear control widgets provide an abstract representation of the entire document in the body of the widget, in that they map each document location to a different position of the slider knob or scroll thumb. In most cases, however, these linear mappings are visually undifferentiated – all locations in the widget look the same – and so it can be difficult to build up spatial knowledge of the document, and difficult to navigate back to locations that the user has already visited. In this paper, we examine a technique that can address this problem: artificial landmarks that are added to a linear control widget in order to improve spatial understanding and revisitation. We carried out a study with two types of content (a video, and a PDF document) to test the effects of adding artificial landmarks. We compared standard widgets (with no landmarks) to two augmented designs: one that placed arbitrary abstract icons in the body of the widget, and one that added thumbnails extracted from the document. We found that for both kinds of content, adding artificial landmarks significantly improved revisitation performance and user preference, with the thumbnail landmarks fastest and most accurate in both cases. Our study demonstrates that augmenting linear control widgets with artificial landmarks can provide substantial benefits for document navigation.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI) → Interaction paradigms**

KEYWORDS

Spatial memory; artificial landmarks; revisitation

SUI '17, October 16–17, 2017, Brighton, United Kingdom

ACM ISBN 978-1-4503-5486-8/17/10…$15.00
DOI: https://doi.org/10.1145/3131277.3132184

ACM Reference format:

M. S. Uddin, C. Gutwin, and A. Goguey. 2017. Using Artificial Landmarks to Improve Revisitation Performance and Spatial Learning in Linear Control Widgets. In *Proceedings of ACM symposium on Spatial User Interaction, Brighton UK, October 2017 (SUI'17)*, 10 pages.

1 INTRODUCTION

Linear documents such as text, webpages, audio, video, or slideshows typically use linear interface widgets for navigation (e.g., scrollbars or sliders). These widgets provide an abstract spatial representation of the entire document (although not a visual representation), in that one dimension of the controller is absolutely mapped to document length (e.g., Y-position of a PDF, or timestamp in a video). Aside from this spatial mapping, most linear controllers do not provide any visual marks that represent document content. Often linear controllers (e.g., YouTube or Adobe Acrobat Reader) provide interactive thumbnails showing only a small portion of the document (but not the whole document) allowing users to access the content close to the focused region. As a result, using the controller to develop a spatial understanding of the document, and to remember and revisit specific document locations, can be difficult [5].

Several researchers have proposed visual augmentations to one-dimensional controllers to address a variety of navigation problems. For example, some techniques show a document's interaction history, such as Hill et al.'s edit-wear scrollbars that showed the number of edits in a text file [21]; other visualizations show notifications such as the location of syntax errors in a code editor. A few of these projects have used augmentations that can support the development of spatial memory – such as the AlphaSlider's alphabetic index [3], the document maps of the Mural Bar [29], Code Thumbnails [15], or the video thumbnail grids of the Swifter video scrubber [28]. However, although studies have shown that these augmentations can improve search and selection tasks [5, 38], there is little evidence about how well they support spatial understanding and revisitation. The exception is a technique called the Footprints scrollbar [5], which visualizes recently-visited and frequently visited document locations. A study showed that the Footprints scrollbar aided navigation back to previous locations, and that showing ten marks in the scrollbar could account for a substantial proportion of revisits [5].

A problem with techniques based on visit histories, however, is that there are many situations where a user may want to revisit a location not shown in the widget. For example, if a user watches an entire video clip, all locations are visited equally during the initial playback, making it harder to go back to a particular scene. Similarly, many visited locations may not appear in a Footprints-style augmentation: for example, some locations have not been seen often enough (or for long enough at each visit) to appear in the visualization, and some locations that go unvisited can disappear from list of recent items.

These problems arise because visit-history techniques depend on the system to remember and visualize important document locations. A different approach is to rely on the users to remember these important locations – which is possible if they are given the resources to exploit their spatial memory abilities. Human spatial location memory is highly effective, and can be both expansive and accurate if the environment is rich and spatially stable (e.g., [5, 34, 43]). Therefore, it is possible that revisitation with linear widgets can be substantially improved simply by adding a rich set of spatially-stable landmarks to the controller – allowing users to build up spatial memory of important document locations.

In this paper, we report on a study of how artificial landmarks affected spatial learning and revisitation using two different linear widgets (a horizontal slider and a vertical scrollbar) and two different kinds of content (a video and a PDF document). The study asked participants to find and then revisit different locations in the documents. For each system, we compared a standard widget with no landmarks, a widget augmented with a set of arbitrary abstract icons; and a widget with thumbnails extracted from the content. Our results showed that both kinds of artificial landmark improved users' spatial learning and revisitation performance, with the thumbnail condition performing best in both systems.

Our work provides three contributions. First, we demonstrate two designs for augmenting linear control widgets with artificial landmarks. Second, we show that that artificial landmarks can significantly improve revisitation performance, in two different contexts. Third, we provide new empirical evidence that adds to our understanding of spatial learning in user interfaces, and that helps to confirm users' ability to learn and navigate documents using spatial memory.

2 RELATED WORK

2.1 Revisitation in User Interfaces

Prior work has shown that human behavior with interactive systems is highly repetitive [23, 48] – although systems often present several ways to complete a task, users mostly choose mechanisms that they are familiar with. Revisitation is one kind of repetitive behavior in which users return to the same locations over and over, a phenomenon that has been most clearly established in use of the web [2, 12, 39]. Revisitation patterns have also been observed in menu selections [11, 16–18], document readers [5], and video players [7], and researchers have looked at several aspects of how revisitation works and how it can be supported with interface augmentations.

Prior work on revisitation support can be divided into two groups. First, there are manual techniques which rely on explicit user actions. The most common tools falling into this category are bookmarks, which are widely available in traditional web browsers, document viewers and office applications [5]. Bookmarks allow users to manually set flags to simplify future revisits of particular content. This idea was used in the Bookmark Scrollbar [25] which places bookmarks in a classic scrollbar. There are, however, some limitations associated with bookmarks' use in interfaces [1]: first, the user has to recognize that a particular location will be revisited (which may not be apparent at the time); second, users must manually place bookmarks, which can discourage users from doing so; and third, people sometimes persist with suboptimal strategies [37] even if a more efficient long-term solution is available [1] (e.g., browsing again through a document vs. using bookmarks).

The second type of revisitation support involves automatic techniques that do not require manual interventions from the user. Interface controls such as 'Forward/Back' buttons and 'Recent Items' menus monitor interactions in the background [5] and update the interface accordingly. However, as observed by Alexander et al. [5], people often misunderstand the use of the 'Recent Documents' menu and the 'Back' buttons in typical web browsers [5, 12], causing problems for efficient revisitation. Other automatic techniques are more explicitly focused on revisitation – such as Hill et al.'s [21] 'read wear' that shows histograms of a user's reading history, Alexander et al.'s [5] Footprint Scrollbar that adds transient history marks in the scrollbar, and Skopik and Gutwin's 'visit wear' augmentations that added explicit visit marks to a fisheye visualization system [38].

Much of the previous work on revisitation has involved text documents, but some research has been carried out on navigation in video players. Most media players support approximate revisitation with 'Forward' and' Rewind' buttons, but precise revisitation remains a problem. Matejka et al. [27, 28] tried to improve video scrubbing by showing thumbnails of the video frames over the slider when the cursor is on the slider knob (a technique also seen in some recent video players such as YouTube – showing only partial range of scenes as thumbnails).

2.2 Interface Augmentation

Interface augmentation is a widely-accepted method to improve computer systems expressivity and interactivity. In GUIs, it is common to augment interfaces with colors, symbols, images or icons, and some researchers have used these techniques to improve document revisitation. As discussed above, 'edit wear' and 'read wear' techniques [21] show interaction histories, and many code editors augment the scrollbar with annotations (e.g., about syntax errors [5]). Other document navigation systems have used colored marks [5] [9] or visualizations of content [3, 15, 29] to augment scrollbars and improve search (e.g., Code Thumbnails, Mural Bar, or AlphaSlider).

Media players are another type of interface where augmentation is commonly used, although previous work does not appear to explicitly focus on revisitation. To support exploration within a video, researchers augmented the timeline slider of the media player, showing visual highlights to represent personal [4] or crowd [24, 47] navigation history. Chen et al.'s

[10] Emo Player annotates the media player's slider with different colors based on the characters' emotional states. Other techniques show interactive thumbnails from the video as a storyboard on the screen [8, 22], giving users an overview of the entire video with a large grid of thumbnails which can assist exploration. Instead of extracting entire frames from the video, Schoeffmann et al.'s video explorer [35] augments the slider with the dominant colors of each frame to help users navigate and explore video content.

2.3 Use of Artificial Landmarks in Interfaces

In GUI-based systems, landmarks such as the corners of the screen can provide a strong external reference frame that helps users build spatial memory [41, 42]. Several techniques have explicitly made use of the bezel and corners of small devices (e.g., tablets or smartwatches) as landmarks to organize menus and toolbars [19, 26, 36]. In areas where there are few natural landmarks (such as the middle of a large screen), artificially placed objects (e.g., colored blocks) or even the user's own hands [42, 44] can act as landmarks and help users to navigate through the interface and recall command locations [5, 43].

This previous work indicates the usefulness of landmarks (even artificial ones), for improving document navigation and command selection. In the next section, we explore the use of artificial landmarks explicitly in linear document navigation.

3 ARTIFICIAL LANDMARKS FOR LINEAR DOCUMENT CONTROLS

To test the performance of artificial landmarks on revisitation in linear documents, and to see how different kinds of augmentation affect performance, we designed new variants of standard slider and scrollbar widgets that are augmented with landmarks. Scrollbars are used to navigate through text, pictures, or any other content in a predetermined direction (vertical or horizontal), when the display can only show a fraction of the content at once. A slider is used to set or pick a value by moving

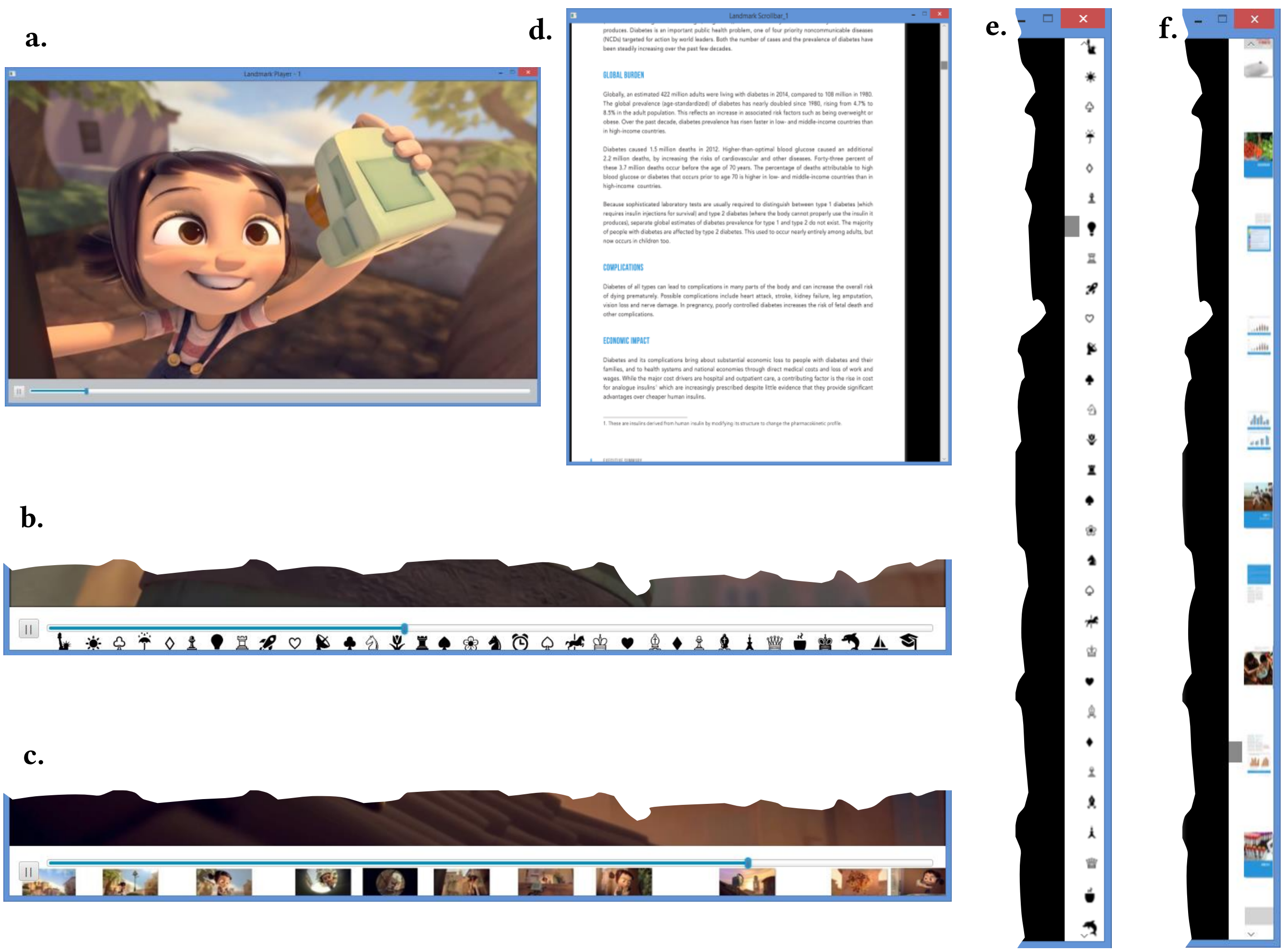

Figure 1: Study interfaces. Media player (a, b, c), PDF viewer (d, e, f). A, d: standard - with no landmarks; b, e: icon – augmented with abstract icons; c, f: thumbnail – augmented with extracted content as thumbnails. Sources [40, 46].

an indicator along a defined segment, usually in the horizontal direction. These widgets can be manipulated easily with a mouse in a standard desktop interface or with a finger in a touch interface.

We developed two versions of each augmented widget, in addition to the standard version. Fig. 1 shows our three media player interfaces: all are 1200×726px in size and all have a 1D linear control (60px tall) at the bottom of the window. All versions support the same navigation features (play/pause button and timeline slider); the only difference between the interfaces is in the augmentation of the control widget. The three versions of the slider were: *standard* (the ordinary slider with no augmentation), the *icons* version with abstract icons as landmarks that had no relation with the contents, and the *thumbnails* version with thumbnails extracted from the video (see Fig. 1: a, b, c). When the user presses the play button, the video starts, and the user can click on the slider to go to any desired location of the video. We removed 'scrubbing' functionality from all three sliders to allow us to better record the users' location choices (i.e., all navigation actions were through clicking on the slider or using the play button).

Our three versions of the PDF viewer (see Fig. 1: d, e, f) also have the same layout and size (900×890px), with a 1D control at the right of the window. Our three versions were similar to those described above, with *standard* (40px wide), *icons* (50px wide), and *thumbnails* (60px) widgets. All versions of the PDF viewer allowed users to view and navigate through the pages of a document only by clicking on the scrollbar. Clicking anywhere on the scrollbar immediately takes the user to the corresponding page; as with the video player, interactive dragging was turned off to get a more accurate measure of users' spatial location memory.

Standard

The standard versions do not provide any extra landmark other than those naturally embedded in a regular slider or scrollbar. The relative position of the thumb in the controller can be used to infer the location within the entire document. See Fig. 1: a and d.

Icons

As the length of the content increases, it is more likely that the relative location cue provided by the scrollbar thumb will become less effective in aiding spatial memorization and retrieval. The icons interface therefore augments the widget with monochrome abstract icons (arbitrary; unrelated to the contents) that are distinct from their surroundings [45], and that provide clear spatial reference points [43]. We placed 34 icons (each 26px in size) horizontally in the media player, and 30 icons (22px each) vertically in the document viewer. See Fig. 1: b and e.

Thumbnails

We augmented the control with actual images extracted from the content. Because we have limited space in the control, we selected thumbnails based on important scene transitions (in the video) or visually distinct pages (in the PDF document). We also kept the inter-thumbnail distance (e.g., time interval in video and number of pages in PDF) approximately uniform. We used 11 thumbnails (each 70x40px) horizontally in the media player and 11 thumbnails (each 32x42px) vertically in the document viewer (see Fig. 1: c and f. There are also methods to automatically extract key elements from documents [33] or video [32], which could be used in future systems.

4 STUDY: EFFECTS OF ARTIFICIAL LANDMARKS ON REVISITATION

We ran a study to examine revisitation performance with our three versions of the linear control widgets. We designed the study to answer two main questions: first, do artificial landmarks improve revisitation compared to the standard widgets; and second, do extracted thumbnails perform better than abstract icons.

4.1 Study Methods and Design

To ensure that the three widget designs were fairly compared for each content type, we used the same video and document for all three widgets. This meant that we could not have participants complete tasks with all three versions of the interface, as they would have built up experience from one interface to the next. Therefore, we chose a mixed within-participants / between-participants design for the study. Each participant used both the video player and the PDF viewer, and used a different widget design for each system. This meant that there were three groups:

- G1 (10 people): *standard* and *icons*
- G2 (10 people): *icons* and *thumbnails*
- G3 (10 people): *thumbnails* and *standard*

All groups were counterbalanced so that five people used each interface each system (e.g., in G1, 5 people used *standard* for the video player and 5 used *icons)*. We included group as a factor to check for grouping effects; because there were none (as described below), we carried out our comparisons using all ten people in each system+interface combination.

Media Player: Tasks and Stimuli

Participants started by watching a video [40] twice using one of the three custom media player interfaces (standard, icons, or thumbnails) that ran on the right screen of a dual-monitor (21-inch) environment. Each participant then went through a series of trials where they were asked to navigate to a specific target frame by clicking on the slider. Each trial began by displaying the stimulus frame on the left screen. The participant then used a mouse to locate the target frame. To make a correct revisitation, participants had to click within 10px (which corresponds to 30 frames) of the target's actual location on the slider.

The video (length 2:12) shown in the study was the same across all three media players as each participant used only one version of the media players. Eight frames from the video were manually selected and used as stimuli. They remained the same across all conditions. In the thumbnail condition, the location of three out of the eight selected target stimuli were on the visible landmarks, and the rest were between or near to the landmarks. We also made sure that the selected targets were spaced regularly throughout the video.

Document Viewer: Tasks and Stimuli

The PDF tasks were similar to those described above. Participants had two minutes to become familiar with a 42-page PDF report [46] using one of the three interfaces (standard, icons, or thumbnails). Each participant then completed a series of trials where they were asked to navigate to a specific target page by clicking on the scrollbar. Each trial began by displaying the stimulus page. The participant then used a mouse to locate the target page. To make a correct revisitation, participants had to click within 12px (which corresponds to half of a target page's height) of the target's actual location on the scrollbar. We used eight pages (spaced approximately regularly through the document) as stimuli. In the thumbnail condition, three target pages were located on the visual representation of a landmark.

Procedure and Study Design

We explored two kinds of content (*media player* and *PDF viewer*) and analysed results separately for each system. For each system, we analysed the effects of interface condition (*standard*, *icons* and *thumbnails*) on revisitation time and errors. Each participant used both systems and saw two different interfaces (but only one version of each interface for each system, as described above). The order of applications and conditions was counterbalanced.

Participants were instructed to complete the trials as fast and as accurately as possible. When using the video system, participants had 15 practice revisitations using a different video than the one used in the main experiment. Participants then completed 5 blocks of trials (each consisting of the same 8 target stimuli, presented in random order). After completed all blocks, participants completed a NASA-TLX subjective workload questionnaire [20]. When using the PDF application, participants had 15 practice revisitations using a different document, and then completed 5 blocks of trials (each consisting of the same 8 stimuli, presented in random order). The PDF task was also followed by another NASA-TLX questionnaire. Participants started with one of the two applications and then proceeded to the other.

For each trial, a selection on the slider or scrollbar displayed the corresponding video frame or PDF page. Participants could adjust their selection up to ten times for each trial. Our software recorded trial completion time, errors, and data describing every selection. At the end of the study, participants provided subjective responses through a questionnaire.

Participants and Apparatus

Thirty participants (8 female), ages 19-30 (mean 24.7), were recruited from a local university. The study took 30 minutes on average. Each participant was compensated with a $5 honorarium.

The experiment was conducted on a desktop computer running Windows 8.1, with two 21-inch 1920x1080 resolution monitors placed alongside. Software was written in JavaFX. Input was received through an optical mouse. All study interfaces ran centered in the right screen (with a white desktop background).

5 STUDY RESULTS

We report the effect size for significant between subject RM-ANOVA results as partial eta-squared: η^2 (considering .01 small, .06 medium, and >.14 large [6]), and Bonferroni correction was performed for post-hoc t-tests.

Before starting our analyses, we first checked for the effect of our grouping variable (see Section 4.1). ANOVA showed no effect of *group* (for the video player, F=0.77, p=.47; for the PDF viewer, F=2.62, p=.09), so we conducted further analyses using all participants for each system+interface combination.

5.1 Media Player: Results

For the media player tasks, 28 out of 1200 trials were discarded from analyses (either because completion time was more than two s.d. away from the respective mean of each block, or the trial could not be completed within 10 attempts).

Media Player: Trial Time

Mean completion times across blocks for the three conditions are summarized in Fig. 2. RM-ANOVA showed a significant main effect of *condition* ($F_{2,27}$=4.45, p=.02, η^2=0.2): 7746ms (s.d. 5804ms) for *thumbnail*, 10306ms (s.d. 5809ms) for *icon*, and 11588ms (s.d. 6703ms) for *standard*. Completion times decreased across *block* ($F_{4,108}$=54.86, p<.001, η^2=0.34), and as anticipated, the skill development follows a power law function [31]. There was no significant *condition* × *block* interaction ($F_{8,108}$=0.94, p=.48).

Post-hoc pairwise t-tests (Bonferroni-corrected) showed that both artificial-landmark conditions were faster than *standard* (all p < 0.01) but showed no difference between *thumbnails* and *icons*.

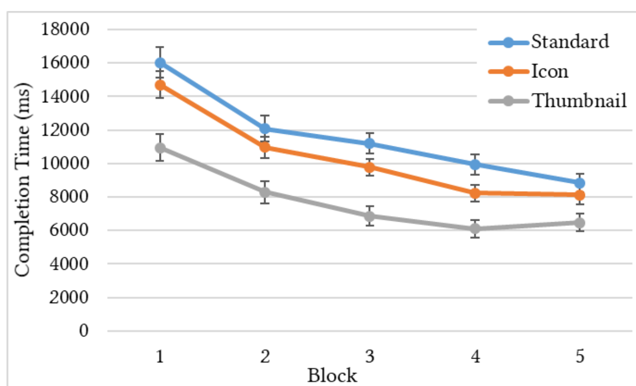

Figure 2: Mean completion time by block and interface.

Media Player: Target Proximity to Landmark

We also analyzed the trial time for *thumbnail* condition based on the proximity of the target to a landmark (3 of 8 targets were located on the thumbnails). As shown in Fig. 3, the on-thumbnail targets were faster than near-thumbnail targets, with mean completion time of 6761ms (s.d. 5366ms) and 8315ms (s.d. 5979ms). However, ANOVA did not show any significant differences between these two categories ($F_{1,18}$=0.51, p=.48).

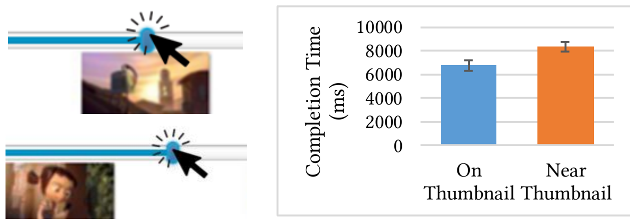

Figure 3: On-thumbnail vs near-thumbnail selections. Left-top: on thumbnail, Left-bottom: near thumbnail. Right: mean completion time.

Media Player: Error Rate – Overall

The number of errors per trial (i.e., the number of attempts to complete a trial) is summarized in Fig. 4. There was a significant main effect of *condition* ($F_{2,27}$=4.27, p=.02, η^2=0.17), but no *condition* × *block* interaction ($F_{8,108}$=1.1, p=.37). Post-hoc t-tests (all p<.001) showed *thumbnail* had the lowest error rate at 2.84 errors/trial (s.d. 3.43), compared to higher error rates of 4.21 (3.37) for icon and 4.30 (3.37) for *standard*. As expected, error rates were higher at the beginning, and decreased significantly with *block* ($F_{4,108}$=8.94, p<.001, η^2=0.1).

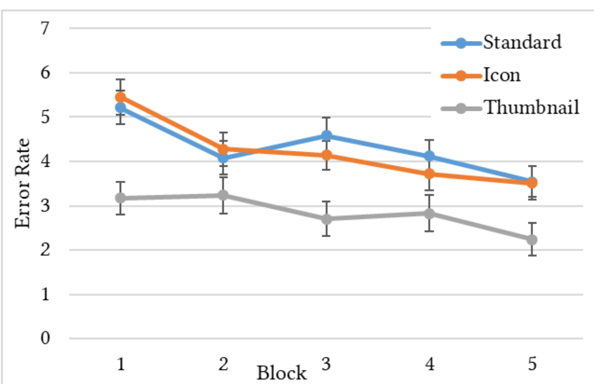

Figure 4: Error rate by block and interface condition.

Media Player: Analyses by Target

We also analyzed errors for each target (stimuli were the same in all three conditions). ANOVA showed a significant difference of *target* ($F_{7,216}$=19.04, p<.001, η^2=0.38) and *condition* ($F_{2,216}$=11.35, p<.001, η^2=0.1), but no *target* × *condition* interaction ($F_{14,216}$<1). Fig. 5 shows the errors/trial for each target and their location in the video. ANOVA also showed significant completion time differences for *target* and *condition* (all F>7, p<.001, η^2>0.2). Post-hoc t-tests showed that the targets placed at the middle (especially targets 5 and 7 for *standard* and *icon* took more time (Fig. 5; all p<.02) and was error prone (targets 4, 5 and 7; all p<.03), following the serial position effect [13]. However, *thumbnails* was significantly better than *standard* for the middle targets (2 and 4-7, all p<.01).

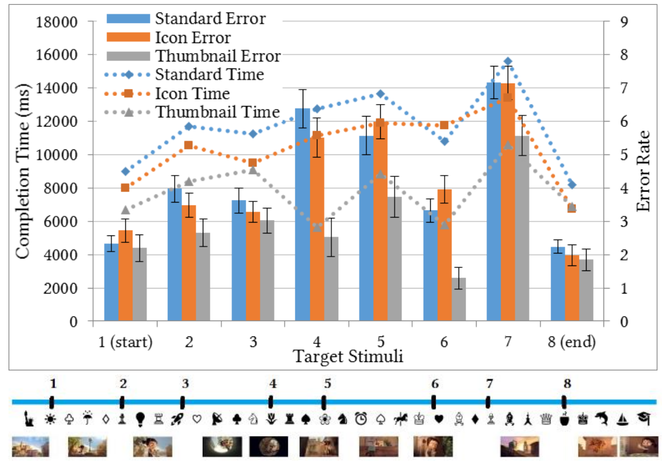

Figure 5: Analyses by target. Top: results, Bottom: locations of the target stimuli in the slider.

Media Player: Subjective Responses

Participant responses on the NASA-TLX worksheets showed significant differences for the three conditions (Kruskal-Wallis tests, Table 1: low score means better, except for Performance). Overall, both landmark conditions performed well. Post-hoc t-tests showed that *thumbnails* achieved significantly better scores for Mental, Physical, Performance, and Frustration scales (all p<.02).

Table 1: Mean (s.d.) effort scores (0-10 scale, low to high)

	Standard	Icon	Thumbnail	χ_r^2	p
Mental	8.1(1.2)	5.1(2.5)	5.2(2.44)	9.92	**.01**
Physical	7.0(1.76)	2.6(1.71)	3.9(1.85)	16.13	**.01**
Temporal	6.3(1.06)	4.7(1.83)	4.1(1.66)	8.22	**.02**
Performance	5.8(1.81)	6.7(2.41)	7.0(2.0)	1.8	.41
Effort	7.7(1.42)	5.5(2.46)	6.1(2.08)	4.97	.84
Frustration	7.0(1.56)	4.4(2.12)	4.0(2.26)	10.55	**.01**

5.2 PDF Viewer: Results

For the PDF viewer, 34 out of 1200 trials were discarded from analyses (completion time was more than 2 s.d. away from the mean, or the trial not be completed within 10 attempts).

PDF Viewer: Trial Time

Mean completion times across blocks for the three conditions are summarized in Fig. 6. RM-ANOVA showed a significant main effect of *condition* ($F_{2,27}$=7.88, p=.002, η^2=0.3): (5236ms, s.d. 3733ms) for *thumbnail*, (6820ms, s.d. 5666ms) for *icon*, and (8685ms, s.d. 5618ms) for *standard*. As with the video player, there was a significant effect of *block* ($F_{4,108}$=53.76, p<.001, η^2=0.35), but no *condition* × *block* interaction ($F_{8,108}$=1.13, p=.35).

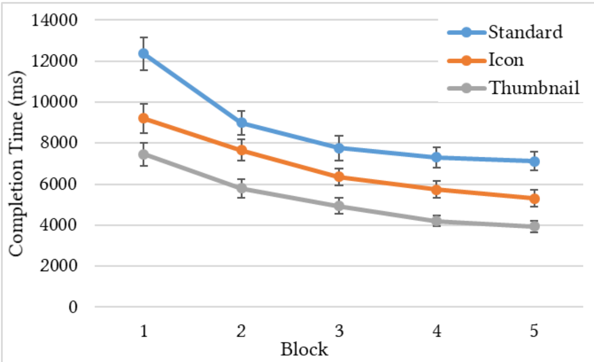

Figure 6: Mean completion time for the three PDF viewers.

PDF Viewer: Target Proximity to Landmark

Unlike the video, ANOVA showed a significant difference for the PDF viewer between on-thumbnail and near-thumbnail targets ($F_{1,18}$=9.12, p<.01, η^2=0.34), with 3859ms (s.d. 2698ms) for on-thumbnail and 6372ms (s.d. 4419ms) for near-thumbnail (Fig. 7).

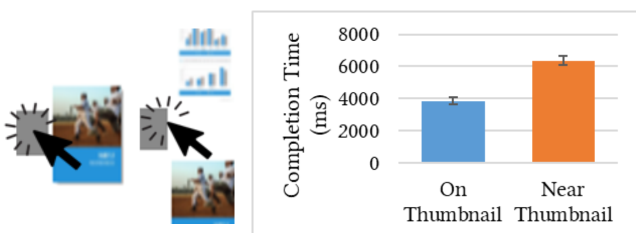

Figure 7: On-thumbnail (left) vs near-thumbnail (middle) selections. Right: mean completion time.

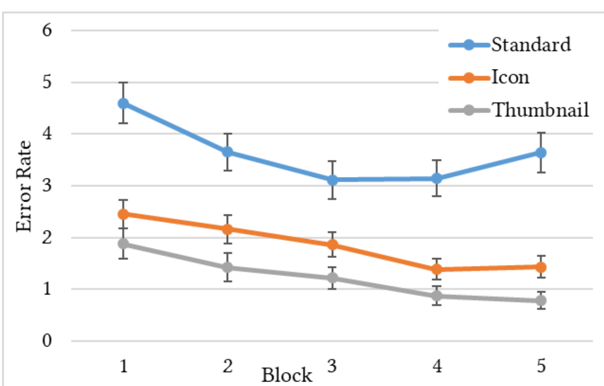

Figure 8: Error rates for document viewers.

PDF Viewer: Error Rate – Overall

There was a significant main effect of *condition* on errors ($F_{2,27}$=24.53, p<.001, η^2=0.51): *thumbnails* had the lowest error rate at 1.22 errors/trial (s.d. 2.03), compared to 3.62 (3.3) for *standard* and 1.85 (2.15) for *icons*. As shown in Fig. 8, errors were high in the first blocks and generally decreased over time (except in the final blocks of the *standard* condition), leading to a significant effect of *block* ($F_{4,108}$=12.21, p<.001, η^2=0.16), but no *condition* × *block* interaction ($F_{8,108}$=0.94, p=.49). Post-hoc t-tests showed that all three interface conditions were significantly different (all p < 0.01).

PDF Viewer: Analyses by Target

Fig. 9 shows the error rates by targets for all conditions in the PDF viewer. ANOVA showed a significant difference among *targets* ($F_{7,216}$=7.73, p<.001, η^2=0.2) and *conditions* ($F_{2,216}$=54.5, p<.001, η^2=0.34). There was also a significant interaction between *targets* × *conditions* ($F_{14,216}$=3.26, p<.001, η^2=0.17).

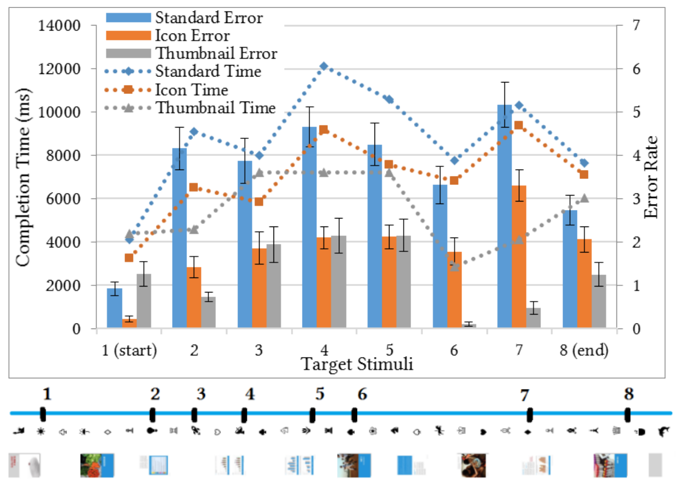

Figure 9: Analyses by target. Top: results, Bottom: locations of target stimuli on the scrollbar.

Targets from the middle areas were most error-prone for *standard* (post-hoc analysis for targets 2-7: >3.32 errors/trial, all p<.001); in contrast, landmarked versions performed better throughout, including the middle areas. ANOVA also showed significant completion time differences for *target* and *condition* (all F>11, p<.001, η^2>0.1). T-tests showed that the landmarked conditions outperformed *standard* at the middle (targets 2, 4 and 5 all with mean completion times <9188ms, p<.05; Fig. 9).

Table 2: Mean (s.d.) effort scores (0-10 scale, low to high)

	Standard	Icon	Thumbnail	χ_r^2	p
Mental	6.8(1.48)	5.7(1.89)	4.4(1.65)	7.9	**.02**
Physical	5.4(2.55)	4.4(2.55)	3.6(1.51)	2.68	.26
Temporal	5.5(0.85)	4.3(1.06)	3.1(1.2)	14.77	**.01**
Performance	6.5(1.58)	7.0(1.63)	8.3(0.95)	6.95	**.03**
Effort	6.1(2.51)	5.7(1.34)	4.8(1.32)	2.5	.29
Frustration	5.5(2.42)	4.4(1.84)	1.8(1.48)	12.49	**.01**

Subjective Responses

Table 2 summarizes mean responses to the NASA-TLX worksheets (low score means better, except for Performance).

Kruskal-Wallis tests and post-hoc t-tests (all $p<.03$) showed significant effects for Mental, Temporal workload and Frustration (*thumbnail* lowest, *standard* highest), and for Performance (*thumbnail* highest, then *icon* and *standard* lowest).

5.3 Participant Preferences and Comments

After completing both studies, participants provided their preferences between the two conditions they used. Table 3 shows that participants favored both artificial-landmark conditions, with 80% of participants preferring them across five measures. The *thumbnails* interfaces were preferred by 70% of participants between the two landmarked conditions.

Table 3: Count of participant preferences

	Group 1		Group 2		Group 3	
	Std.	Icon	Icon	Thmb.	Thmb.	Std.
Speed	2	8	2	8	8	2
Accuracy	2	8	3	7	7	3
Memorization	2	8	2	8	10	0
Comfort	3	7	2	8	8	2
Overall	2	8	3	7	8	1

Participant comments echoed our other findings. Participants made several comments on how the landmarks (especially thumbnail) helped them to develop spatial memory of the contents. One participant stated, *"Snapshot of a specific action [thumbnail] helped me remember the story sequence [of the video]."* Another mentioned *"I remembered the sequence [of the video] after seeing the closest thumbnail."* For the thumbnails in the PDF viewer, one person mentioned *"Thumbnails [of the pages] made it easy to find exact pages and guess the nearby page locations."*

Other comments indicated that the icons also helped participants to learn command locations: one mentioned *"It was easy to remember which page was beside which icon [in the PDF viewer]."* Some participants revisited locations by correlating icon with the content: as one said, *"I could correlate some of the icons with video contents."* A few participants, however, found the icons more difficult: *"The randomness of the icons was a little tough for me to remember [in PDF viewer];"* another said, *"I associated a few key locations [of the PDF document] to icons [but] it was difficult to keep track of the icons."*

6 DISCUSSION

Previous research has identified that revisitation is common in computer interfaces [5, 7], and that artificial landmarks can be useful in helping users remember item locations for future visits [41, 43]. Our study results suggest that spatially-stable artificial landmarks can help people to remember locations and can improve revisitation performance. Our study provides two main results:

- Both artificial-landmark interfaces were faster than the standard widget for both applications
- Of the two types of artificial landmarks, the *thumbnails* condition was fastest and most accurate, and was strongly preferred by participants.

6.1 Explanation and Interpretation of Results

Landmarked Interfaces can Improve Spatial Revisitation

During the study, mean completion time decreased significantly across blocks for all conditions in both applications; and as anticipated, they followed the power-law of learning curve [31]. The number of errors also decreased across blocks. These can be indications that users transitioned quickly from slow visual search revisitation to more rapid spatial-memory-based revisitation. This transition occurs in all conditions for both applications, confirming that users developed spatial memories, regardless of the interface. However, Figs. 2, 4, 6 and 8 reveal that participants were successful in forming and exploiting these memories more rapidly and accurately when interfaces were augmented with artificial landmarks (*icons* and *thumbnails*) for both applications.

Figs. 2 and 6 indicate that although the *standard* interface for both applications allowed participants to build a certain degree of spatial memory of the contents, participants were less able to rely on their spatial memory to perform rapid actions [34, 43]. As shown in Figs. 4 and 8, participants in the *standard* condition made substantially more mistakes in revisitation, especially in the later blocks of the PDF viewer, likely because the standard widgets provided no clear reference frame. As a result, *standard* interfaces with no landmarks were slowest and most error prone.

Why did *Thumbnails* Outperform *Icons*?

Our results show that in both applications, several measures favoured the *thumbnails* interface over *icons* (trial time, error rates, workload measures, and preference). We see two key factors responsible for this advantage. First, the thumbnail landmarks in the *thumbnails* interfaces were actual representations of the contents as images in miniature scale. The thumbnails showed a clear mapping between the control widget and the content, compared to *icons* more abstract icons that had no connection with the actual content. Participants had to manually form the mappings between icons and contents for later visits. Though participants were successful in learning the mappings and revisited target stimuli, additional time was required to learn the connections, which may explain the slower performance of *icons* compared to *thumbnails.* The clearer indication of document content may have helped participants to exploit their spatial memory and perform rapid revisitations with better accuracy. Performance of the *icons* condition could potentially be improved by choosing icons that are more meaningful to the content (although this may not always be possible).

Second, the number of landmarks available in the interfaces may have caused the performance difference between the two interfaces. *Icons* interfaces had more landmarks (more than 30) compared to *thumbnails* (11). Overloading the 1D controller of an interface with more items means users require to learn and remember more mappings. Learning and remembering only 11 items may have been easier – determining the ideal number of landmarks is an interesting area for future study. Additionally, the available contents of the document, especially in PDF document where colorful charts and images were present in more than 60% pages, could have helped users to form spatial memory. Exploring the influence of document-contents (e.g., pages with images or only texts) on spatial memory formation is a compelling area of future work.

Landmarks can Help Overcome the Serial Position Effect

The serial position effect [13] suggests that people best recall the first (primacy effect) and the last (recency effect) items in a series compared to the middle items [14, 30]. During our study, we showed the entire content (video and document) to the participants first, then asked them to revisit or recall items shown as target stimuli. Analyses by targets (see section 5.1 Fig. 5 and section 5.2 Fig. 9) show that standard interfaces for both applications followed the serial position effect, meaning that the targets used from the beginning and the ending of the documents were recalled more accurately than the middle targets. However, we see that landmark-augmented interfaces appeared to overcome this effect: participants were able to revisit targets in the middle with better accuracy in both applications, especially with *thumbnails*. We believe this was possible because of the presence of the artificial landmarks in *icon* and *thumbnail* interfaces. Landmarks provided spatial anchors for the participants that helped to remember middle area's contents and allowed them to revisit those locations accurately.

6.2 Implications for Design

Our findings provide additional evidence that artificial landmarks can assist in developing users' ability to build spatial memories for rapid revisitation [5, 43]. Many interactive interface components often limit users from utilizing the full potential of their spatial memory because they lack landmarks. For example, the blank trough of a slider or scrollbar may lead to inefficient revisitations. Past research has attempted to temporarily augment scrollbars with interaction histories [5, 21] or usage information of a video in media player's slider [24]. Our results suggest that static embellishments of the interface with artificial landmarks can substantially improve users' revisitation experiences.

There are, however, a few challenges involved in augmenting interfaces with landmarks. First, there is a chance of interference from artificial landmarks in the interface and from the primary content of the interface. While the user is focusing on the main content, the static landmarks on the controller might catch unwanted attention from the user. Similarly, focus on the main content might push the landmarks out of the user's focus. Careful design of the landmarks and proper placements can overcome this issue. For instance, landmarks can be highly transparent to reduce unwanted focus; this way, users only switch their attention on the landmarks when they need to remember content. Further studies are needed to explore real-world issues in use of artificial landmarks.

A second challenge is the length of the content (e.g., video duration or document length). For a relatively large document, small movements of the slider/scroller result in large navigation actions. A possible solution to this problem is adding additional controllers – each controller could be responsible for specific range (e.g., 30-minute video or 50-page document) of the full content, and still provide landmarks to support revisitation.

Third, the success of the *thumbnails* interface may be related to our ability to choose appropriate and meaningful thumbnails from the document content. Although previous work in automatic extraction of summary information from video has proven to be successful (e.g., [32, 33]), a clear direction for future work is to replicate our study using an algorithm for automatic extraction of thumbnails. In addition, less work is available on the best method for creating memorable and meaningful thumbnails from text or PDF documents. It is important to note that even if automatic extraction is difficult for some documents, our abstract icons also can provide a significant improvement compared to standard un-augmented linear control widgets.

In future, we plan to explore the use of artificial landmarks in realistic application setting with large contents, explore the automatic extraction of meaningful landmarks, and determine the ideal number of landmarks for different document lengths. We will also carry out studies to analyze interference between landmarks and content, and explore new designs to determine how different levels of visual salience interact with the development of spatial memory.

7 CONCLUSIONS

Linear control widgets can provide a stable spatial representation of a document, and enable efficient navigation and revisitation. We explored the potential of artificial landmarks in improving spatial learning and revisitation of locations for linear documents. Linear control widgets for two applications (a media player and a PDF viewer) were augmented with artificial landmarks – either arbitrary abstract icons or thumbnails extracted from the content – to help users form spatial understanding of the document while using the interfaces. We compared standard widgets (with no landmarks) to widgets augmented with artificial landmarks. Both artificial-landmark conditions improved performance, with the *thumbnails* condition performing best in terms of revisitation time, errors, perceived effort, and preference. Our studies show that artificial landmarks are a simple and valuable method for improving navigation in linear documents.

ACKNOWLEDGMENTS

This work was supported by Natural Sciences and Engineering Research Council of Canada. Our special thanks to Miranda Miller who recruited and ran participants.

REFERENCES

[1] D. Abrams, R. Baecker, and M. Chignell. 1998. Information archiving with

bookmarks. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '98*. New York, NY, USA, 41–48.

[2] E. Adar, J. Teevan, and S. T. T. Dumais. 2008. Large scale analysis of web revisitation patterns. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. New York, NY, USA, 1197.

[3] C. Ahlberg, and B. Shneiderman. 1994. The alphaslider: a compact and rapid selector. In *Conference companion on Human factors in computing systems - CHI '94*. New York, New York, USA, 226.

[4] A. Al-Hajri, G. Miller, M. Fong, and S. S. Fels. 2014. Visualization of personal history for video navigation. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA, 1187–1196.

[5] J. Alexander, A. Cockburn, S. Fitchett, C. Gutwin, and S. Greenberg. 2009. Revisiting read wear: analysis, design, and evaluation of a footprints scrollbar. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI '09*. New York, NY, USA, 1665–1674.

[6] O. Bau, and W. E. Mackay. 2008. OctoPocus: a dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st annual ACM symposium on User interface software and technology - UIST '08*. New York, New York, USA, 37–46.

[7] F. Bentley, and J. Murray. 2016. Understanding Video Rewatching Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX '16*. New York, NY, USA, 69–75.

[8] J. Boreczky, A. Girgensohn, G. Golovchinsky, and S. Uchihashi. 2000. An interactive comic book presentation for exploring video. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00*. New York, NY, USA, 185–192.

[9] D. Byrd. 1999. A scrollbar-based visualization for document navigation. In *Proceedings of the fourth ACM conference on Digital libraries - DL '99*. New York, NY, USA, 122–129.

[10] L. Chen, G.-C. Chen, C.-Z. Xu, J. March, and S. Benford. 2008. EmoPlayer: A media player for video clips with affective annotations. *Interacting with Computers*. 20, 1 (Jan. 2008), 17–28.

[11] A. Cockburn, C. Gutwin, and S. Greenberg. 2007. A predictive model of menu performance. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*. New York, NY, USA, 627–636.

[12] A. Cockburn, and B. Mckenzie. 2001. What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies*. 54, 6 (Jun. 2001), 903–922.

[13] A. Coleman. 2006. *Dictionary of Psychology*. Oxford University Press.

[14] J. DEESE, and R. A. KAUFMAN. 1957. Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of experimental psychology*. 54, 3 (Sep. 1957), 180–7.

[15] R. DeLine, M. Czerwinski, B. Meyers, G. Venolia, S. Drucker, and G. Robertson. 2006. Code Thumbnails: Using Spatial Memory to Navigate Source Code. In *Visual Languages and Human-Centric Computing (VL/HCC'06)*11–18.

[16] L. Findlater, and J. McGrenere. 2004. A comparison of static, adaptive, and adaptable menus. In *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*. New York, NY, USA, 89–96.

[17] S. Greenberg, and I. H. H. Witten. 1993. Supporting command reuse: empirical foundations and principles. *International Journal of Man-Machine Studies*. 39, 3 (Sep. 1993), 353–390.

[18] S. Greenberg, and I. H. H. Witten. 1993. Supporting command reuse: mechanisms for reuse. *International Journal of Man-Machine Studies*. 39, 3 (Sep. 1993), 391–425.

[19] C. Gutwin, A. Cockburn, J. Scarr, S. Malacria, and S. C. Olson. 2014. Faster command selection on tablets with FastTap. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA, 2617–2626.

[20] S. G. Hart, and L. E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*. 52, (1988), 139–183.

[21] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. McCandless. 1992. Edit wear and read wear. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*. New York, NY, USA, 3–9.

[22] D. Jackson, J. Nicholson, G. Stoeckigt, R. Wrobel, A. Thieme, and P. Olivier. 2013. Panopticon:a parallel video overview system. In *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*. New York, New York, USA, 123–130.

[23] J. M. Juran. 1951. *Quality Control Handbook*. McGraw-Hill.

[24] J. Kim, P. J. Guo, C. J. Cai, S.-W. (Daniel) Li, K. Z. Gajos, and R. C. Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*. New York, New York, USA, 563–572.

[25] S. A. Laakso, K. P. Laakso, and A. J. Saura. 2000. Improved scroll bars. In *CHI '00 extended abstracts on Human factors in computing systems - CHI '00*. New York, NY, USA, 97–98.

[26] B. Lafreniere, C. Gutwin, A. Cockburn, and T. Grossman. 2016. Faster Command Selection on Touchscreen Watches. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA, 4663–4674.

[27] J. Matejka, T. Grossman, and G. Fitzmaurice. 2012. Swift: reducing the effects of latency in online video scrubbing. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. New York, NY, USA, 637–646.

[28] J. Matejka, T. Grossman, and G. Fitzmaurice. 2013. Swifter: improved online video scrubbing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. New York, NY, USA, 1159–1168.

[29] D. S. McCrickard, and R. Catrambone. 1999. Beyond the scrollbar: an evolution and evaluation of alternative navigation techniques. In *Proceedings 1999 IEEE Symposium on Visual Languages*270–277.

[30] J. Murdock, Bennet B., and B. B. 1962. The serial position effect of free recall. *Journal of Experimental Psychology*. 64, 5 (1962), 482–488.

[31] A. Newell, and P. S. Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*. 1, (1981), 1–55.

[32] A. Pavel, D. B. Goldman, B. Hartmann, and M. Agrawala. 2015. SceneSkim: Searching and Browsing Movies Using Synchronized Captions, Scripts and Plot Summaries. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. New York, NY, USA, 181–190.

[33] A. Pavel, C. Reed, B. Hartmann, and M. Agrawala. 2014. Video Digests: a browsable, skimmable format for informational lecture videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*. New York, NY, USA, 573–582.

[34] J. Scarr, A. Cockburn, C. Gutwin, and A. Bunt. 2012. Improving command selection with CommandMaps. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. New York, New York, USA, 257–266.

[35] K. Schoeffmann, M. Taschwer, and L. Boeszoermenyi. 2010. The video explorer: a tool for navigation and searching within a single video based on fast content analysis. In *Proceedings of the first annual ACM SIGMM conference on Multimedia systems - MMSys '10*. New York, NY, USA, 247.

[36] K. Schramm, C. Gutwin, and A. Cockburn. 2016. Supporting Transitions to Expertise in Hidden Toolbars. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, NY, USA, 4687–4698.

[37] H. A. A. Simon. 1987. Satisficing. *The New Palgrave Dictionary of Economics*. P. Eatwell,J., Milgate, M., and Newman, ed. Stockton Press. 243–245.

[38] A. Skopik, and C. Gutwin. 2005. Improving revisitation in fisheye views with visit wear. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*. New York, New York, USA, 771–780.

[39] L. Tauscher, and S. Greenberg. 1997. How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies*. 47, 1 (Jul. 1997), 97–137.

[40] Thesis Film - Last Shot: A. Widodo.: *http://awidodoportfolio.blogspot.ca/p/film.html*. Accessed: 2017-05-01.

[41] M. S. Uddin. 2016. *Improving Multi-Touch Interactions Using Hands as Landmarks*. University of Saskatchewan.

[42] M. S. Uddin, and C. Gutwin. 2016. Rapid Command Selection on Multi-Touch Tablets with Single-Handed HandMark Menus. In *Proceedings of the SIGCHI Conference on Interactive Surfaces and Spaces - ISS '16*. Niagara Falls, ON, Canada, 205–214.

[43] M. S. Uddin, C. Gutwin, and A. Cockburn. 2017. The Effects of Artificial Landmarks on Learning and Performance in Spatial-Memory Interfaces. In *Proceedings of the 2017 SIGCHI Conference on Human Factors in Computing Systems - CHI '17*. Denver, CO, USA, 3843–3855.

[44] M. S. Uddin, C. Gutwin, and B. Lafreniere. 2016. HandMark Menus: Rapid Command Selection and Large Command Sets on Multi-Touch Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA, 5836–5848.

[45] N. G. Vinson. 1999. Design guidelines for landmarks to support navigation in virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*. New York, NY, USA, 278–285.

[46] WHO Global report on diabetes.: 2017. *http://www.who.int/diabetes/global-report/en/*. Accessed: 2017-05-01.

[47] B. Yu, W.-Y. Ma, K. Nahrstedt, and H.-J. Zhang. 2003. Video summarization based on user log enhanced link analysis. In *Proceedings of the eleventh ACM international conference on Multimedia - MULTIMEDIA '03*. New York, New York, USA, 382–391.

[48] G. K. Zipf. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley.