Supporting Early Document Navigation with Semantic Zooming

Tom Owen[1], George Buchanan[2], Parisa Eslambochilar[1], and Fernando Loizides[2]

[1] Future Interaction Laboratory, Swansea University, Swansea, UK
[2] Centre for HCI Design, City University, London, UK
{cstomo,p.eslambochilar}@swansea.ac.uk
{george.buchanan.1,fernando.loizides.1}@city.ac.uk

Abstract. Traditional digital document navigation found in Acrobat and HTML document readers performs poorly when compared to paper documents for this task. We investigate and compare two methods for improving navigation when a reader first views a digital document. One technique modifies the traditional scrolling method, combining it with Speed-Dependent Automatic Zooming (SDAZ). We also examine the effect of adding "semantic" rendering, where the document display is altered depending on scroll speed. We demonstrate that the combination of these methods reduces user effort without impacting on user behaviour. This confirms both the utility of our navigation, and the minimal use information seekers use of much of the content of digital documents.

1 Introduction

Within–document navigation is a common action performed by users when reading texts in many circumstances: initial skimming, quick readings, the deep analysis of a selected document, and checking remembered details. In this paper, we investigate the support of navigation from the primary perspective of initial navigation of a document. The user may seek specific targets (e.g. keywords), but the exact location of these is unknown, and they even may be absent from the document. In this environment, users seek to obtain a quick overview of the document to determine its structure and judge its content, usually in the context of an immediate information need. This initial and brief reading is usually termed "document triage".

Cathy Marshall [11] has demonstrated that digital document reader software significantly impedes users' interaction with documents. Our work addresses the challenge, raised by Marshall, of providing "library materials that not only capture the affordances of paper, but also transcend paper's limitations".

We build upon recent research [4,5] that indicates that users actually use only limited parts of a document's content during interactive triage. This selective attention appears to be even more pronounced in digital documents when compared to printed texts. We follow the natural corollaries of this knowledge to

G. Chowdhury, C. Khoo, and J. Hunter (Eds.): ICADL 2010, LNCS 6102, pp. 168–178, 2010.

minimise the information presented to the user when scrolling, and to maximise the visibility of the remaining content. In principle, this will reduce visual clutter and improve a user's visual search performance.

However, there is contrary evidence that non–realistic renderings of documents undermine user's acceptance of documents [8]. Therefore, a focussed study is required to assess whether the use of a novel presentation results in a lower subjective user evaluation of the techniques.

This paper commences with a general introduction to the "state of the art" in digital document navigation. We proceed to describe the design and implementation of a set of novel techniques for supporting within–document navigation. These are subsequently evaluated in a user study. The results of the study are described in detail, and discussed in relation to previous research. We conclude with an outline of future avenues for research and a summary of the main contributions of the work.

2 Navigation in Digital Documents

To understand the problems experienced during digital document navigation, we initially ran a laboratory–based observational study [3]. We interviewed academics who were regular users of both digital and paper texts. The aim of these interviews was to provide qualitative data that would provide a more detailed picture of the current interaction problems with digital documments. We identified specific issues from our own observational data [3] and the work of researchers such as Cathy Marshall [11] and Kenton O'Hara [12], including: temporary placeholding; permanent bookmarking [6]; navigation; skimming and overview; search–within–documents and reading effort. In this paper we focus on the issues of navigation, skimming and overview.

We conducted semi–structured interviews with nine humanities researchers, and repeated this with three computer science researchers. Ratings were obtained on a scale of 1 to 10 for specific interactions, and qualitative feedback elicited to complement each rating. When asked to compare the ease of navigation in digital and paper documents, the users gave an average rating of 5.7 (digital) versus 7.4 (paper). This difference may appear small. However, the detailed picture revealed a stark contrast between positive and negative features.

Scrolling proved to be a complex issue: while participants expressed a positive inclination towards it, in practice their experiences were mixed: e.g. "I like scrolling but I find it annoying when it pauses and just stops on you". Another user was more specific: "when you want to scroll quickly and the computer can't keep up, it is really confusing, you end up not knowing where you are". Render speeds are a key underlying problem here, where larger documents often cannot be displayed sufficiently quickly to create a smooth interaction. One academic observed that "You just cannot flick through the PDF".

Simple "go-to" navigation for a specified page was seen as an advantage to electronic media, though three participants reported problems with PDF documents having different printed and logical page numbers. Books downloaded

from Google Books were cited by a historian as one particular problem: "the page numbers seem just to start at one, regardless of the original front matter or numbers, so I have to keep on making up the numbers or just guessing it." In contrast, though printed paper lacked that specific problem, it had a more prevalent one "it gets real hard when the pages seem to stick together and you go back and forth and blow on it to get ... the right page ... I get so distracted by that."

Other findings triangulated well with Marshall and O'Hara: e.g. reading was rated as requiring more effort on digital displays. We also corroborated our earlier observations in revealing a low rate of use of positively rated electronic tools such as within–document search. One literature scholar said: "I guess, in truth, I think I use it when I get desperate, when I can't see what I think...expect should be there."

At the end of each session, we introduced a number of experimental document reader applications for comparison. User responses to these – particularly to one overview tool – informed the design of novel navigations that we introduce in the next section.

3 Design

Scrolling is a major component of navigation in digital documents. Existing literature demonstrates that scrolling is the main method for moving through an electronic text [2,9]. It is therefore worthwhile improving this particular element of document reader software.

Our interviewees reported that slow and erratic response to scrolling was one irritation experienced in using Acrobat and similar software. This in part stems from a relatively simple rendering paradigm in document reader software: regardless of scrolling speed, all content is rendered. However, it is worth considering why people scroll rapidly through digital documents.

Data from many sources indicates that rapid scrolling is caused by a desire to obtain a quick overview of a document. Liu [9] and Marshall [11] both note this behaviour, and relate its occurrence to gaining a general impression of a document's content and structure. O'Hara and Sellen [12] report similar patterns of users' document handling nearly a decade earlier. This suggests that this pattern is a deeply ingrained and persistent one, and good reader software will be designed with this behaviour in mind.

In this section, we articulate four alternative designs for scrolling across PDF documents. We combine two different approaches: adjusting scrolling strategies on one side, and using different rendering techniques on the other. Using two different methods for each of the two approaches results in four different combinations. We will now discuss the scrolling and rendering methods we have used in turn, before demonstrating the resulting designs in use. In the following section, we will report on a four–way comparison of these different methods in a laboratory user study.

3.1 Speed–Dependent Automatic Zooming (SDAZ)

Traditional scrolling presents a simple direct–manipulation method for moving across a document. One novel method that has been popularised in recent years is SDAZ. In SDAZ, as a user scrolls across a surface more rapidly, the view zooms out, to give a wider overview of the content (see Fig. 1). This seems to be a reasonable candidate for document navigation, and indeed has been tested by Cockburn et al [1] for such purposes.

Fig. 1. Document display with traditional (left) and SDAZ (right) scrolling. In SDAZ, the view has shrunk the text and content, providing a less detailed but broader view.

One of the findings in Cockburn's study suggests that SDAZ provides a more rapid method on targeting "large document features" such as major headings and images than traditional linear scrolling. However, users were cued with a known target object and an indicated direction of movement in the study. This did not match the problem we are studying: in our targeting task, users will be often moving towards a target that is at best partially anticipated, and where the direction is at best partially assumed (i.e. to be below the first page). We therefore needed to retest this proven method for our new task.

We developed an extension of our existing document reader software [5] that encompassed both linear scrolling and SDAZ scrolling. The two different modes are represented in Figure 1.

3.2 Semantic Rendering

Our second design consideration is that of document rendering. Our previous study and the interview data both revealed that slow rendering directly impacted upon user's document navigation. We also had discovered that visual clutter negatively impacts user performance during visual search [5].

There are multiple techniques used in the desktop publishing industry to alleviate rendering problems. These can be observed in design software such as Adobe Illustrator: first, there is full rendering, which yields a high quality view of the document at a high time cost; second, threading can be used to progressively render a complex view with increasing fidelity as the same document fragment remains visible; thirdly, a reduced content display can be used (e.g. outlines only) which produces a general impression of the content very rapidly.

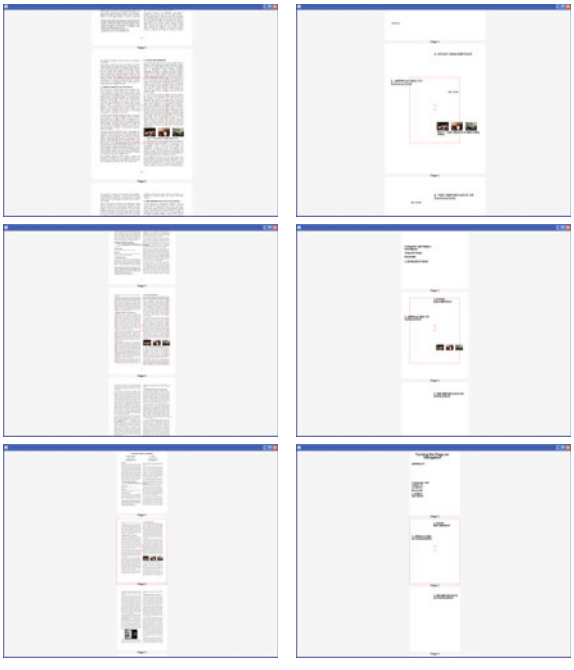

Fig. 2. SDAZ with different levels of zoom: full (left) and semantic (right) rendering

We developed a semantically (or, more precisely, typographically) informed rendering method that combines progressive and outline rendering. Instead of reducing filled volumes to mere outlines, we instead select heading text to be displayed first, and then the remainder of the content is progressively rendered in decreasing order of significance: e.g. heading, sub–heading, caption, and emphasised text. This results in a relatively rapid display during rapid scrolling, and minimises interactional lags, where a user's scrolling position is far in advance of the current displayed text.

3.3 Implementation

We developed four different interfaces built upon the same underlying software: 1) a traditional, linear zoom, full rendering display; 2) linear zoom with semantic rendering; 3) SDAZ with full document content; 4) SDAZ with semantic rendering. The software was built on JPedal, a Java–based PDF rendering suite.

4 User Study

We undertook a user study to evaluate the four different interfaces. In this section, we first report the study design, before presenting the results from the experiment.

4.1 Study Design

To provide a testbed for the study, we selected eight different PDF documents on a single computer–science related topic. For each document, we created a set of simple navigational tasks, informed by our knowledge of typical goals that users seek to satisfy during document triage. Examples would include "find the conclusions section to identify the main contribution of the paper", or "locate where the paper discusses accessibility for haptic interfaces". For each task, we had a known target point in the document. There were four types of target: a literal–match on a heading (e.g. "conclusions section"); a non–literal heading match; a match on the body text, and a false (absent) target. By varying literal and visibility properties of the target across different tasks, the results would provide a discrimination between such factors in our results.

The basic form of the study followed a conventional pattern of pre–study questionnaire, the main experimental session, and a post–study semi–structured interview. The induction questionnaire obtained basic demographic information and the participant's experience with using document reader software such as Acrobat. The post–study interview gleaned subjective information about the participant's experience with the different systems that they encountered.

The main part of the experiment consisted of the participant using all four example systems we described above. For each system, the observer demonstrated the software in use on a standard example document. The participant was then given an open–ended time to familiarise themselves with the system. The study then progressed to 32 small–scale tasks over eight documents (one task of each target type per document). A simple questionnaire was taken for each separate interface, and a summative evaluation to compare all four interfaces.

To balance for ordering, learning and dependency factors, a latin–squared design was used throughout. Combinations of system, task, document and order were used to minimise biases in the data.

The questionnaire on each system captured the participant's immediate subjective ratings of the that system. The concluding interview probed the participant's assessment of each system in detail, to gain further insight into the comparative advantages and disadvantages of the interfaces.

To ensure that learning effects were not introduced into the study, we sought to recruit users who were experienced with digital document navigation and the tools and concepts behind scrolling techniques. Therefore for the study we recruited participants who were undertaking research and likely to be making regular use of digital documents.

For the experiment, we recruited a total of sixteen computer science students, aged from 20 to 37. Participants were either in the final semester of a bachelor's degree, or engaged in postgraduate study. Potential participants were vetted for dyslexia and uncorrected sight defects, and we present here only data from participants who passed these criteria (one dyslexic candidate was excluded).

4.2 Results

The experiment captured a volume of data, from which we will only extract the most authorative findings. We commence with an evaluation of differences in user behaviour during the study, before progressing to analyse the participant's subjective feedback elicited through the questionnaires and interview.

Log Data. Each participant completed 8 tasks in every interface mode. Tasks were of four types, as described in the experimental design. We analysed the data for each type of task separately, using a two–way ANOVA in each case. For tasks where the body text matched the task, or the task was to find content not in the document, there were no discernable effects. Only when the targets were headings did any significant results emerge.

The most striking result occurred in the exact heading match tasks: the two–factor ANOVA produced three significant results. Comparing SDAZ versus linear scrolling produced p=0.0360, F=4.543, F–cit=3.954. This indicated a significant difference between the two modes, with linear scrolling proving superior (average times of 17.50s vs. 24.20s). The comparison of full text versus semantic rendering modes resulted in p=0.0089 (F=7.171, F–crit=3.954), revealing an even stronger outcome, this time with the full rendering proving inferior (average of 25.05s vs 16.52s). Finally, the interaction test revealed p=0.0176 (F=5.861, F–crit=3.954). (df=1,1,1,84 for the test as a whole).

The underlying reason for these outcomes is readily understood when it is stated that the SDAZ full–text mode produced an average time of 32.40s against averages of between 16.17s and 17.92s for the other three modes. Put plainly, there was strong evidence that the SDAZ full text mode meant that even heading text was unreadable for most participants. When semantic zooming was applied, the SDAZ method performed at a level comparable with linear scrolling.

Task Completion Data. The rate of completion on heading match tasks was very high. Only 5 out of 128 exact heading match tasks and 8 topical heading match tasks were not completed. Of the 128 body–text match tasks, however, there was a very high failure rate: 69 out of 128. These were fairly distributed across all modes, and applying a Chi–square test to the data gave a result indicating p=0.985. This is clearly very far from suggesting any significant effects from any mode.

Subjective Ratings. The participants rated each interface for nine task–focussed factors on a seven point Likert scale. These were subsequently evaluated through a two–factor ANOVA to determine where perceptible effects were noted. The first item that returned a significant result was the rating "clarity of document display". Here, we encountered p=0.038 (df=1,1,1,44; F=4.566,0.943,0.604; F–crit=4.062). The significant difference was produced from the comparison of the rendering method: semantic presentation was rated lower than full text (avg.=4.21,var=2.95 vs avg.=5.125,var=1.42).

Interestingly, no reliable difference emerged in the rating "clarity of document display when scrolling" (F=0.030,1.921,0.120; F–crit=4.062). Nor was any effect uncovered when comparing "ease of use".

The rating "Ease of seeing document structure" produced an interaction effect (p=0.008; F=0.028,0.713,7.31; df=1,1,14, etc.). The semantic SDAZ and traditional linear scrolling rated inferior to the SDAZ full–text and linear semantic presentations. This may suggest that users are able to benefit from either SDAZ or semantic presentation, but the combination is too unfamiliar or demanding. Further investigation will be required.

The second factor the produced a signficant result was "Time to evaluate document size". Here, the average ratings given varied from 2.83 (linear full–text mode) to 5.25 (semantic SDAZ). ANOVA yielded p=0.0005 (df=1,1,1,44; F=1.715,14.094,0.191). This reveals a significant impact from the scrolling method. SDAZ based methods were faster than the linear scroll methods.

A related third factor also produced a conclusive result "Time to identify document structure". ANOVA resulted in p=0.021 (df=1,1,1,44; F=1.033,5.587,1.560; F crit=4.062). The significant difference was again between the SDAZ and linear scrolling modes. No significant effect was computed from the application of semantic zoom.

These tests paint contrasting picture, with the SDAZ methods yielding a better overview of the document in a short time, while linear methods were preferred in terms of the quality of display. We will now report the subjective feedback from our participants.

Subjective Feedback. Our participants gave plentiful subjective feedback on their experience of using the four different systems. The impact of the different presentations were pronounced. Criticisms of the semantic presentation were commonplace: with the clear message being the low acceptability of losing document text "this is completely unnatural", "I'm not seeing the real document" are just two examples of a consistent hostility to the presentation format. The "non–natural" effect was specifically highlighted by ten participants.

The SDAZ method also produced many reservations from our users. Concerns here were about the visibility of headings. These reservations were – unsurprisingly – more consistent in the full–text mode, where they were shared by 11 participants. However, even in the semantic presentation, six participants expressed the same problem.

5 Discussion

The results of our topic raised a number of different issues. Subjectively, users preferred the familiarity of the traditional linear scrolling technique, and the display of full text. However, there was only minimal evidence that there was actually any performance advantage to this mode. In contrast, though the combined SDAZ and semantic rendering interface was subjectively poorly rated, its users attained a similar level of completion rate to every type of task, and in very similar times. The combination of SDAZ with full text display was, however, noticeably inferior for simple heading–match tasks with all other interfaces. It was a closer match in all other tasks.

Our experiment demonstrates that the SDAZ method, while comparable in performance for many tasks, is not well accepted in the context of reading text. Beneficially, the zooming feature of SDAZ can give a quick impression of the size of the document more readily than linear scrolling. On the other hand, the low resolution of the text makes reading even headings very difficult when using full–content rendering.

Semantically–enhanced SDAZ outperforms traditional SDAZ when a user is searching for content that appears in headings, but there is no difference for smaller content (e.g. body text). Omitting the document content when scrolling is rapid proved to be unsatisfactory for most of our participants.

The role of body text in skim reading is problematic: while users are dissatisfied when it is not displayed, they nonetheless are known to rate it as relatively unimportant when deciding document relevance in an information seeking task [7] and our own data shows that visual search is rather ineffective, with over 50% of visual searches failing. This raises rather significant questions for future research in the field, particularly as how to support a user's visual search. While within–text search features are commonplace in document readers, studies repeatedly demonstrate a very low rate of use (e.g. [10]). If our participants are typical of the general user population, artificial renderings of content, even if more task–effective, will face considerable resistance and low subjective ratings.

This last point is a significant issue for the development of digital libraries. If users are demonstrably poor at picking out key text for their information need from the document body, assistance will be required. However, users' subjective dependence on seeing the "authentic" document will strongly constrain the range of designs intended to support their visual search. Our data clearly demonstrates that visual search performance is poor, and as this activity is a key step in the information seeking of most users, progress is urgently required to improve the efficacy of DLs as a whole.

6 Conclusion and Future Work

Reading digital documents is a commonplace task where there are clear issues outstanding with user interaction. This has direct relevance to digital libraries: as users increasing read digital documents "online" rather than on paper, these deficiencies become ever more important to address. This paper has demonstrated that novel techniques for visual search, even within digital documents, cannot be used uncritically when the user is scanning text. We also identified that using document content and presentation can close usability deficits where they exist. Applying semantic rendering did improve many parts of the reading process, but does not yet perfect the SDAZ method.

SDAZ provides a route for improving the rapid overview of a document, but it proves less effective for detailed reading. The limited resolution of digital displays, combined with persistence effects commonplace even on modern LCD computer screens, result in poor legibility of small or briefly viewed text. These limitations are unlikely to disappear in the near future.

Our research into SDAZ as a method for supporting the rapid navigation of text shows that a naive use of the method leads to difficulties in user satisfaction and also some problems with effectiveness. Nonetheless, there are benefits and successes that suggest persistence is required to actually uncover its long–term potential. Our semantically–enhanced SDAZ presentation performed much closer to the traditional interface than its full–text rendering sibling. The main drawback was the elimination of body text. One avenue for research is how best to reconstitute the presentation of body text in the semantic rendering mode.

An additional strand of research would be to investigate semantic rendering on a set of users who are likely to have different reading habits to that of those tested in our study. Based on the findings of the research presented in this paper it would be sensible to introduce a broader audience to the concepts we have introduced. A different set of participants may yield yet more information as to how to improve SDAZ.

Both SDAZ and linear scrolling methods proved ineffective in supporting a user locate information hidden in the main body, and this is confirmed as a major interaction challenge. SDAZ and other presentation modes have shown that interactions can be improved and user needs both better understood and satisfied through systematic research. Similar diligence must now be applied to visual search across the plain text of digital documents.

Acknowledgments

This research is supported by EPSRC Grant EP/F041217.

References

1. Andy Cockburn, J.S., Wallace, A.: Tuning and testing scrolling interfaces that automatically zoom. In: CHI 2005: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, pp. 71–80. ACM, New York (2005)
2. Buchanan, G.: Rapid document navigation for information triage support. In: Proc. ACM/IEEE Joint Conference on Digital Libraries, June 2007, p. 503. ACM Press, New York (2007)
3. Buchanan, G., Loizides, F.: Investigating document triage on paper and electronic media. In: Kovács, L., Fuhr, N., Meghini, C. (eds.) ECDL 2007. LNCS, vol. 4675, pp. 416–426. Springer, Heidelberg (2007)
4. Buchanan, G., Owen, T.: Improving navigation in digital documents. In: Procs. ACM/IEEE Joint Conference on Digital Libraries, pp. 389–392. ACM, New York (2008)
5. Buchanan, G., Owen, T.: Improving skim reading for document triage. In: Proceedings of the Symposium on Information Interaction in Context (IIiX), pp. 83–88. British Computer Society (2008)
6. Buchanan, G., Pearson, J.: Improving placeholders in digital documents. In: Christensen-Dalsgaard, B., Castelli, D., Ammitzbøll Jurik, B., Lippincott, J. (eds.) ECDL 2008. LNCS, vol. 5173, pp. 1–12. Springer, Heidelberg (2008)

7. Cool, C., Belkin, N.J., Kantor, P.B.: Characteristics of texts affecting relevance judgments. In: 14th National Online Meeting, pp. 77–84 (1993)
8. Flanders, J.: The body encoded: Questions of gender and electronic text. In: Sutherland, K. (ed.) Electronic Text: Investigations in Method and Theory, pp. 127–144. Clarendon Press (1997)
9. Liu, Z.: Reading behavior in the digital environment. Journal of Documentation 61(6), 700–712 (2005)
10. Loizides, F., Buchanan, G.R.: The myth of find: user behaviour and attitudes towards the basic search feature. In: JCDL 2008: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, pp. 48–51. ACM, New York (2008)
11. Marshall, C.C., Bly, S.: Turning the page on navigation. In: JCDL 2005: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries, pp. 225–234. ACM, New York (2005)
12. O'Hara, K., Sellen, A.: A comparison of reading paper and on-line documents. In: CHI 1997: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 335–342. ACM, New York (1997)