

Readability of Scanned Books in Digital Libraries

Alexander J. Quinn¹, Chang Hu¹, Takeshi Arisaka², Anne Rose¹, Benjamin B. Bederson¹

¹University of Maryland
Department of Computer Science
Human-Computer Interaction Lab
College Park, MD 20742 USA
{aq, changhu, rose, bederson}@cs.umd.edu

²Hitachi, Ltd., Software Division
Advanced Information Systems R&D Department
Hitachi System Plaza Shinkawasaki
Kawasaki, Kanagawa, 212-8567 Japan
takeshi.arisaka.vh@hitachi.com

ABSTRACT

Displaying scanned book pages in a web browser is difficult, due to an array of characteristics of the common user's configuration that compound to yield text that is degraded and illegibly small. For books which contain only text, this can often be solved by using OCR or manual transcription to extract and present the text alone, or by magnifying the page and presenting it in a scrolling panel. Books with rich illustrations, especially children's picture books, present a greater challenge because their enjoyment is dependent on reading the text in the context of the full page with its illustrations. We have created two novel prototypes for solving this problem by magnifying just the text, without magnifying the entire page. We present the results of a user study of these techniques. Users found our prototypes to be more effective than the dominant interface type for reading this kind of material and, in some cases, even preferable to the physical book itself.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces

General Terms

Human Factors, Experimentation, Design

Author Keywords

Digital libraries, online document display, readability, document navigation, scanned books, children's books

INTRODUCTION

When illustrated books and layout rich documents are scanned and displayed wholly on computer screens, a variety of display problems compound to make reading and comprehending the materials difficult. When viewed on small and medium size displays, the images must typically be scaled down in order to fit, or scaled up to make the text large enough to read. Since web browsers, operating systems, and other software often place fixed toolbars and other interface components at the top and bottom of the screen, a substantial fraction of the total screen space is often unavailable for the actual content.

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 2008, April 5–10, 2008, Florence, Italy.

Copyright 2008 ACM 978-1-60558-011-1/08/04...\$5.00



Figure 1: Standard reader showing entire page. Text is part of the image scan.

Estimating Web Users' Display Sizes

If all computer users were using large monitors, then this would be less of a problem. However, a large number of computer users are using small monitors, either because of the portability of laptops and other mobile devices or because large displays are not available to them.

There is a broad spectrum of display sizes in use. We can gain some understanding of what is out there by instrumenting web sites with JavaScript code to record the screen resolution reported by visitors' web browsers. Of the last 200,000 visitors to our own International Children's Digital Library, 62% came with display resolutions of either 1024x768 or 800x600. The web development information web portal, w3schools.com, reports that in January 2007, 68% of their site visitors had one of those two resolutions [36]. 60% of visitors to the University of Texas home page were set to one of those two resolutions [35]. In all three cases, the vast majority of other resolutions reported were greater than 1024x768. Assuming that most of these displays are about 90 dots per inch, a typical value, that works out to a screen size of approximately 11 inches (28 cm) wide by 8.5 inches (22 cm) tall for the 1024x768 resolution. That is the same as a "14 inch" monitor, a common size for notebook computer LCD screens.

The exact amount of screen space taken by toolbars and interface components will vary depending on the configuration. As an example, when viewing any book

using Google's online book viewer with Internet Explorer maximized at 1024x768 resolution, the space available for the book is only 744x542, or 6 inches high assuming 90 DPI. In that configuration, viewing a letter-size book page without scrolling would require shrinking the page down to 54% of the original height and width. 10-point type becomes 5.4 points, well below the range of type sizes that can be comfortably read [4]. Furthermore, the lower resolution of computer displays, compared with paper, makes even comparably sized displays harder to read than paper [28]. In the example shown in Figure 2, as originally shown at 1024x768, the cap height (distance from the baseline to the top of a capital letter) is 6 pixels and each page occupies only 22% of the total screen area. The text is barely legible and cannot be read comfortably. Even when the web site and browser were both set to their full screen modes, the cap height increased to only 8 points.

Displaying scanned books online

The most common method of viewing scanned books online is to display the scanned page at an easily readable size, but require the user to use scrollbars to see different regions of the page. That has the drawback that simply viewing a page requires effort to move the scrollbars and that some short term memory is required to remember where the current view fits into the context of the whole page. Furthermore, scrolling over a 2D surface with two scrollbars is notoriously difficult and time consuming. Perhaps most importantly, enjoyment of some kinds of materials (e.g. children's books) is dependent on seeing the entire page, with illustrations and text in the same view. For such materials, scrolling is not a good option.

A related issue is that most current web browsers use image resizing algorithms that favor speed over quality. This becomes an issue because many web-based digital libraries use JavaScript or similar technology to resize the images so they precisely fit the viewable area in the user's screen. This is done because the variability in users' system configurations makes the viewing area size unpredictable.

This research was initiated to address problems experienced with the International Children's Digital Library (ICDL, www.childrenslibrary.org) [20], an online digital library of children's books. We are especially focused on solving the problem with picture books. We define a picture book as one for which the illustrations and text have roughly equal importance. With picture books, the standard solution of scrolling a view pane is not adequate because the illustrations and text need to be seen simultaneously in order to fully appreciate the literary value. Furthermore, picture books often come in especially large formats or odd shapes, so the differences in aspect ratio can be especially challenging.

The problem of displaying scanned picture books in a way that is readable and easy to navigate is an ongoing concern for the ICDL project. In the current ICDL web site, users are first shown an image that is dynamically resized to just



Figure 2: Space constraints in Google's book reader

fit in the browser content area while preserving the original aspect ratio (although this resizing will be removed soon). The initial view is shown in Figure 1. Then, if that is not readable, users can click a button to zoom in, resulting in the view shown in Figure 4. At that point, scrolling is needed in order to view the entire book page. Comments from users indicate the current solution is problematic. Solving this problem is the core motivation for this work.

Main contributions

This paper describes the prototypes that we implemented to address the issues just raised and presents a controlled user study that evaluates them.

The important contributions of this paper include the following:

- We provide two fresh approaches to the problem of showing scanned, illustrated books in web browsers.
- We show how computer vision can be used to offer an interface that eliminates the need for scrolling.
- We illuminate the relationship between the dynamicity of the content layout and the ease of navigation.
- We convey the user's perspective on what makes a given approach suitable.

RELATED WORK

Solving the problem with the ICDL would require improving the readability of the text while maintaining easy navigation through the books. Thus, most of the related work can be divided between general readability of static text and document navigation techniques.

Readability

As early as 1971, Baldwin and Baily studied the differences between reading materials in print and reading them on a microfiche screen, concluding no difference in preference or performance between positive and negative contrast [3]. This result is relevant to the current work because many children's books have light colored letters on a dark background, such as a night sky. Concerns with readability of on-screen text have continued since then, as summarized by Mills and Weldon in 1987 [27]. Their comprehensive review shows that on-screen readability is affected by many

variables, such as text spacing, font style, contrast, and screen resolution. They found that readability is influenced by the quality of the computer screen. Dillon has written extensively on the differences between reading from computer monitors and reading from paper [13], but his comprehensive literature review concluded that the many confounding variables make it difficult to create a solid comparison [11]. Shneiderman's textbook on human-computer interaction summarizes research on the differences between reading from computer monitors and reading from paper. The conclusion is that screen resolution is the sole reason for performance differences, provided that all layout and font variables are kept constant [31]. Similarly, Ziefle found that eye fatigue is reduced when screen resolution is increased [39].

More recently, research in readability of text on computer screens has focused on typography. A 1995 study found that as the text size drops below 9.75 points on a 15-inch 1024x768 CRT monitor, preference measures declined, but performance was optimal between 8.25 and 9.75 points, regardless of the font family used [34]. Another more recent study found that although font sizes of less than 6 points generally lead to reduced accuracy in proofreading tasks, small differences such as 10-point to 12-point yield little or no change [4]. In general, when comparing typefaces commonly encountered on modern desktop computers, they found 12-point non-antialiased Arial (a sans serif typeface) to be most preferred, while 10-point antialiased Times New Roman (a serif typeface) was least preferred. However, a separate study found that users strongly preferred Georgia (a serif typeface) to Verdana (a sans serif typeface) [8]. In that case, reading performance was also better with Georgia. This apparent contradiction might be because Georgia and Verdana were originally designed for viewing on screen but Arial and Times New Roman were designed for print.

Dyson studied the effect of physical layout of text on readability, finding that reading performance can be optimized by considering the movements of the eye through the text [14]. For example, narrower characters may improve performance because the eyes can take in more text in each fixation. Also, although study participants tended to prefer dual columns (e.g. the current document) and narrower columns, performance actually increases with wider columns because they require fewer focal sweeps from right to left.

Recently, subpixel font rendering has become common in commercial software created by Microsoft, Adobe, and Apple. For example, ClearType is a technology built into Windows XP and Windows Vista that takes advantage of the way LCD monitors are constructed with separately addressable red, blue, and green subpixels in order to render curves and diagonal lines more smoothly [6]. Although the technology effectively increases the

resolution at which fonts can be rendered, beyond the native resolution of the monitor, it is generally only effective when reading black text on a white background on an LCD monitor in landscape orientation [16]. Although patents cover all of these commercial implementations, SubLCD, a free and open-source project, provides a similar rendering technology, free of patent entanglements [32].

Although readability has been studied extensively with adults, not as much work about on-screen readability by children has been done. One study attempted to gauge the effects of varying font family, style, and size on reading performance and subjective preference with children aged 9 to 11 [5]. Concrete results were limited because of variability in the children's reading strategies, reading abilities, and levels of motivation. They did however find that the children strongly preferred sans serif fonts to serif fonts. The children also preferred 14-point fonts to 12-point fonts. That is equivalent to an x-height (physical height of a lowercase "x") of about 2.5 mm versus 3.0 mm on the 17-inch 1024x768 CRT monitor used.

In interpreting all of these results, it is important to consider that reading is not one universally consistent activity, but has very different characteristics depending on the setting and task at hand. Reading an email in an email program is a qualitatively different experience from reading a scanned book or filling out a form on a web page. Depending on the task and its underlying purpose, the required level of content comprehension will differ. Furthermore, different ways of presenting text afford different strategies of moving through the material.

Document Navigation

Due to the prevalence of electronic document handling tasks in everyday computer use, extensive research has been done to study efficient methods of viewing and navigating documents on computer screens.

One commonly studied problem is that of navigating a large document when limited screen space is available. The scrolling window remains the dominant approach but several projects have explored alternatives. Many of these projects are based on generalized fisheye views, which was formalized by Furnas [15]. The idea is that the user controls an area of focus in which the text is magnified to a readable level. The rest of the document is shown in a compact, but usually illegible representation. One study compared reading text documents using the fisheye interface with the usual scrolling interface and one more where a compact "overview" is given alongside a scrollable, readable "detail" window. It found fisheye to be the fastest but overview+detail to be most suitable overall [18]. Another study by the same authors found that overview+detail can be distracting because the entire document is visible all the time, but that it also led to higher scores on an essay assignment, compared to essays written while using other interfaces [19].

Many other novel approaches have been tried for navigating long documents. One such experiment found that by filling the screen with small, clickable thumbnail images of pages in the document, users were able to quickly find specific information in the document and reported greater satisfaction [9]. Another project tried using a card stack idiom [7]. Robertson's and Mackinlay's "document lens" used a 3D visualization to view a page of a collection at full magnification at the top of a virtual hill, with the other pages seemingly sliding down the sides of the hill [30].

Another study compared several methods for displaying text dynamically using a variety of methods, including scrolling horizontally, scrolling vertically, flashing one word at a time (RSVP), and others [24]. They did evaluations using several kinds of mobile devices with varying screen sizes. They found that the presentation method had a greater effect than screen size on reading speed. This reinforces our belief that by redesigning the way that scanned book pages are presented, we can have a greater impact than we would have by simply increasing the amount of free screen space.

Similar to the document navigation problem is the problem of viewing an image that is too large to fit on the screen. This is essentially what many users of the ICDL need to do to read a scanned image of a book page on a small to medium sized screen. Several novel techniques exist. For example, Liu et al use a method similar to RSVP to show successive sections of an image, emphasizing salient regions, such as text or human faces [26]. Jones et al explore the tradeoff with time and space efficiency in a different way, using speed-dependent automatic zooming, a technique that varies the magnification level depending on how fast the user pans across the image [23]. This technique was shown to be especially effective for locating specific targets, especially on medium sized screens.

Based on the techniques explored in academic research, industry has developed some compelling solutions to the same problems. The Adobe Reader desktop application, for example, comes standard with three tools for document navigation [1]. First, it implements the overview+detail interface by providing a left navigation panel with thumbnails of each of the pages, with the full-size page in the center of the screen. It also provides two kinds of lens tools. One provides an small window with a magnified version of a portion of the screen. The other provides a small overview window with the current page. Adobe also has a product for PDAs that takes a PDF file as input and automatically reflows the text and resizes the images so that a full-size PDF file can be easily viewed on a small device [2].

Active reading

Commercial products such as Adobe Reader also provide facilities for digital annotation of reading material. Digital ink annotations on tablet computers have been used to



Figure 3: ClearText reader. Text is rendered with operating system text rather than as an image.

explore active modes of reading by Schilit [32] in the XLibris system. That system lets users create annotations to keep track of information learned while reading and to go back and search later. Although our proposed solutions do not involve active reading, similar features might be a natural extension in the future, provided they could be designed to suit young readers. Kaplan and Chisik have explored this design space. They found that social reading can be supported by digital libraries to enhance satisfaction and educational objectives for preteens [24].

CANDIDATE SOLUTIONS

Given the complexity of the problem and the diversity of users and reading styles, we opted to design two solutions and then run a comparison study. The solution currently deployed on the ICDL, which uses traditional geometric magnification, is also described for comparison.

ClearText

In the current ICDL interface, when a page is displayed wholly on the screen, two issues are apparent. First, the portion of the screen that is occupied by the text is very small, typically around 25% for picture books. If the entire browser content area could be utilized, the font size could be increased significantly. However, that space is used for interface components, such as the title bar, browser toolbars, the title of the book, scroll bars, and the task bar at the bottom of the screen. Second, artifacts from the compression and resizing of the images cause further readability problems beyond what can be attributed to the size of the type.

Prototype #1, called ClearText, addresses both issues by partially decoupling the text from the image of the page background. From the user's perspective, ClearText makes it possible to resize the text without resizing the background. When the user views a page or pair of pages of the book, the image(s) are scaled to fit the available space in the browser content area. Unlike other solutions, the text shown is not the original book's text, but rather a reproduction using the computer's built-in fonts. Initially,

the text is drawn at a size that closely matches the printed book. At the top of the screen, there are magnifier icons that can be used to increase or decrease the size of the type. Increasing the font size may obscure the illustrations, but makes the text easier to see. Whether larger text obscures the illustrations depends on the layout of the original page. The text is initially centered at the same place as the original text and is never allowed to extend beyond the edges of the page. In that way, the computer-rendered text seems like an integral part of the book page.

Choosing the possible font sizes was the subject of some early design iterations. Initially, we set the minimum font size to 10-point and set no maximum font size, other than to constrain that the text must fit inside the page image. However, that seemed confusing because a very text-dense page would only be displayable up to a modest font size, but a page with very little text would have much more flexibility. However, for the sake of consistency, we wanted the font size to remain constant when the user switched to a different page. If, for example, the user set the font size to 24-point on a sparse page and then moved to a more dense page, there would be no good way to display the new page at the selected size. Therefore, it became evident that each book would need to have a global maximum font size. That global maximum is set to the maximum point size at which the text can be displayed in the most text-dense page of the book. The ClearText interface is shown in Figure 3.

The technical means of enabling ClearText are beyond the scope of this report. Here, we provide just a brief summary. The objective in preparing books to be viewed using ClearText is to discover where the text resides on the page, remove it, and fill in with a background that matches the area around it. Using a combination of computer vision techniques and some human validation, we find the bounds of the text on the page and identify which pixels contribute to the text. Then, we use an infilling algorithm to fill in those pixels with color and texture that matches the surrounding area. That leaves a page with only the illustrations, but not the text. Then, we use volunteers to transcribe the text, so that it can be displayed on top of the page, in approximately the same region that previously contained the original text. Text is rendered on top of the background image using JavaScript and cascading style sheets (CSS). The new text is centered on the same location where the original text was.

One of the key advantages of ClearText is that it utilizes the computer's native font rendering capabilities. This avoids displaying text that has undergone compounded degradation through scanning, compression, and resizing to fit the browser window. Instead, the quality of the text is bounded only by the system's native font rendering capabilities. Thus, it can even potentially take advantage of subpixel font rendering [16] in current generation operating systems for smoother, more readable text than would be possible with even a perfect image of text.

Furthermore, once real text is drawn by the browser, the text can be searched. Accessibility is also improved because screen readers can now work. Further, the ability to control how text is drawn offers the possibility of displaying translations of the books in the original context.

Using the computer's native fonts provides flexibility, but sometimes at the cost of the author's creative intent, especially in the case of some children's books that feature slanted or shaped text to convey feeling. Since ClearText renders all text in rectangular blocks of horizontal text, that feeling may be lost.

Another challenge in designing ClearText was that the text may overlap with salient parts of the image, making the text difficult to read. Our prototype solves this by adding a semitransparent white rectangle between the page background and the text. This reduces contrast in the underlying image while increasing the contrast of the text. If the text overlaps with an important part of the image, the user can use the controls to temporarily reduce the font size.

PopoutText

In order to preserve the styling of the original text and its relationship to the background, a second novel prototype was created. Prototype #2, called PopoutText, selectively magnifies just the portion of the image that contains the text, leaving the original text as part of the image. The PopoutText interface is shown in Figure 5. As with ClearText, when the user initially views a page or pair of pages of the book, the whole width and height of the page(s) is visible. When the mouse is hovered over an area that contains text, a dashed blue outline appears around the text, indicating that the area is clickable. When the user clicks on the text area, it "pops out" of the page at a higher magnification level that is easily readable. Clicking on the image again causes it to pop back in, leaving the original image of the entire page(s). Alternatively, clicking on a second block of text while the first block is still expanded causes the first block to pop back in and the second block to pop out.



Figure 5: PopoutText reader. Text is part of the scanned image.

Compared to ClearText, the technical needs of PopoutText are much less. Neither transcription nor identification of pixels contributing to text are needed. Only the bounding box of each block of text is needed. The popping out is handled by JavaScript and CSS code. Currently, when the text pops out, it is displayed immediately. However, we ultimately plan to implement smooth transitions similar to the Highslide thumbnail image JavaScript library [17], an early inspiration for PopoutText.

The main advantages of PopoutText over ClearText are believed to be its preservation of the original text styling and relative ease of implementation. Furthermore, PopoutText handles pages with multiple, disjoint blocks of text more gracefully than ClearText does. However, while it preserves the styling and relative positioning of the text, it sacrifices the ability to appreciate the illustrations at the same time the text is being read – a key goal of this research. PopoutText also does not support searching, read-aloud for accessibility, or translation.

Standard

The interface currently deployed on the ICDL web site is similar to that of other common web-based digital libraries. Thus, we consider this the baseline from which we hope to improve. The standard viewer initially presents the entire page or pair of pages. By clicking on a magnifier icon, the user can magnify the entire image. The magnified image is displayed in a scrollable view panel. As deployed today, there are 5 magnification levels available. When the user navigates to a different page, the magnification level is preserved. The standard interface is shown in Figure 4.

Use case comparison

To understand the differences in navigation among the three prototypes, we consider the following example: “A user has loaded the first page of a book, reads that page, and then reads the next page.”

ClearText: The user clicks the magnifier icon twice at the top of the screen to change the font size from 10-point to 14-point. Next, the user reads the page. To go to the next

page, the user clicks the arrow icon at the top of the screen. The font size is preserved, so the user continues reading.

PopoutText: The user notices a block of text, hovers the mouse over it, and notices the dashed blue outline. The user clicks the text once and the text pops out of the page as shown in Figure 5. To see the rest of the illustrations again, the user clicks the popped out text to cause it to shrink back into the page. To move to the next page, the user clicks the arrow icon at the top of the screen. To read the next page, the user again clicks the text once to cause it to pop out.

Standard: The user clicks once on the magnifier icon at the top of the screen. The entire page is magnified. At that point, the text is no longer completely visible. The user clicks and drags the horizontal scrollbar to scroll horizontal scrollbar, and then again on the vertical scrollbar to find the text and center it in view. This may need to be repeated if the user can't immediately find the text. To see the full illustrations again, the user clicks the other magnifier icon to return to the original magnification level. To move to the next page, the user clicks the arrow icon at the top of the screen. Reading the next page requires the same procedure as for the previous page.

EVALUATION

In order to understand the factors contributing to applicability of the three interfaces, a controlled user study was conducted. Since picture books come in a wide variety of layouts and styles, it would have been impossible to adequately test every type of book with all interfaces. Therefore, the focus of the evaluation was on discovering the specific factors contributing to success, rather than simply selecting a winner.

The general strategy was to have each participant read a significant section of a picture book using each of the three computer interfaces plus the physical book. Thus, there were four sections, one for each of the four reading mechanisms.

Study design

Fliers and email announcements were used to recruit 27 respondents of ages ranging from 18 to 62, with an average age of 25. All participants had some connection to the university. Children were not used because of the variability found in previous readability studies with children [5] and because we were interested in the general problem of presenting layout rich materials online, and not just improving the ICDL. Furthermore, our experience with the ICDL shows that somewhere around half of its use is with adults (parents, teachers, librarians, learners of a second language, etc.). Thus, adults are an important segment of our users. Requirements for participation in the study included normal reading vision with or without correction, and ability to read English children's books at a native speaker's level. On average, participants reported being very comfortable (4.2 on a 5-point Likert scale) reading documents on a computer. They also reported using computers an average of 6.8 days per week.



Figure 4: Standard reader zoomed in using the interface to magnify the text.

Experiments were conducted in an office setting with overhead fluorescent lighting. The interfaces were accessed via the Internet Explorer version 6.0 web browser on an IBM Thinkpad R50 laptop computer with a 14-inch LCD screen set at the native resolution of 1024x768. The choice of screen resolution and web browser reflect the majority of visitors to the ICDL web site. Participants were free to use a mouse, the laptop touchpad, or the built-in TrackPoint pointing stick. Actions such as zooming, panning, and page loads were logged using JavaScript. All participants were told to adjust the angle and distance of the screen to whatever was most comfortable and natural.

The biggest challenge was in choosing appropriate reading material. If study participants were to read the same passage for all four methods, perceptions of the material would change over time and likely affect perceptions of the reading mechanisms, as well. However, a perfect assessment would use passages of equal font size, style, color, and contrast. The tone of the story and the number of words per line and words per page would also be equal, in the ideal case. We searched through the ICDL database and at local bookstores for books or compilations of books that might have these characteristics. Creating synthetic material was considered but abandoned because such material would probably not have the same artistic appeal and would thus skew attempts to assess the effect of a particular interface on the author's creative intent.

As a compromise, *The Blue Sky* [21] and *Ciconia Ciconia* [22] were chosen. These books were written by the same author (Andrea Petrlík Huseinović), and they have the same physical page dimensions and approximately the same total number of pages. The tone of the stories is also very similar. Both books have rich, colorful illustrations upon which the text is set. However, there are significant typographical differences between the two books. *Ciconia Ciconia* is set in a sans serif font. *The Blue Sky* is set in a slightly larger serif font. Also, *Ciconia Ciconia* contains a few pages with negative contrast (white letters on a black background) but *The Blue Sky* does not. These differences represent a potential confounding factor in the study. However, this threat is mitigated by the much more significant differences between the interfaces themselves. Recent readability studies have shown that small differences in font size make little or no difference [4]. In the case of ClearText, the font used to overlay the text was the same for both books. Also, counterbalancing ensured that each section of a book was paired with a given interface roughly the same number of times. The order in which the interfaces were presented was also counterbalanced as much as possible, although perfect

counterbalancing was infeasible due to the number of participants.

Procedure

The two books were each divided into two sections equal to half the length of the book, either 8 or 10 text pages. The first 2 pages of each section were reserved for training. The following procedure was followed for each of the 4 reading mechanisms:

1. Researcher reads a brief set of instructions and gives the participant an opportunity to ask questions.
2. Participant reads the first 2 text pages as a training phase.
3. Participant is again allowed to ask questions, if any.
4. Participant reads the rest of the section, terminated by a verbal command by the researcher.
5. Participant answers two written, multiple choice questions about the content of the story, just to verify attentiveness. (No participant answered more than one question wrong during the entire test.)
6. Participant fills out a subjective assessment questionnaire about that interface, including 5 questions on a 6-point scale. Questions covered legibility, perceived reading speed, ease of manipulation, preservation of the author's creative intent, and attractiveness.
7. Participant is given an opportunity to rest or stretch, if desired.

After completing all 4 sections, a final questionnaire was given, wherein the 4 interfaces were ranked together by the same 5 criteria used in the previous questionnaires. At the end, participants were encouraged to give comments and explain the reasons for their rankings.

Elimination

Initially, there were 31 study participants, but 4 were eliminated due to mistakes made in filling out the questionnaires. The preceding section refers only to the participants that were *not* eliminated. As a quality measure, we eliminated all questionnaire data for any participant who answered "strongly disagree" for the physical book to the question, "This book reading method preserves the book author's creative intent." Since the book *is* the author's creative intent, we assumed that such participants did not read the questionnaire carefully. There were 2 such participants. All other participants answered "strongly agree" (26) or "agree" (1), corresponding to 6 and 5 on the 6-point scale. We also eliminated all questionnaire data for 2 participants who completed the reading exercises but did not completely answer the final questionnaire.

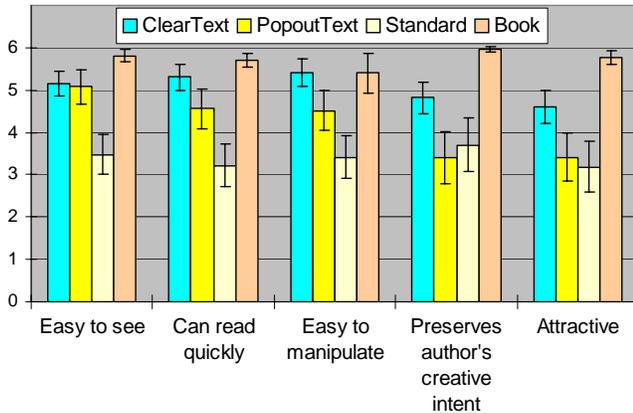


Figure 6: Average ratings (High values are good. Error bars represent a 95% confidence interval.)

RESULTS

Subjective assessments of individual interfaces

We performed an RM-ANOVA on each of the five measures (readability, perceived speed, ease of manipulation, adherence to creative intent, and attractiveness), with the 6-point subjective rating as the dependent variable and the book reader (standard, ClearText, PopoutText, or physical book) as the independent variable. In order to run the analysis, we substituted the average value for one missing value, due to a question left blank by one study participant.

- *Readability:* A significant main effect ($F(3,78)=37.938$, $p<0.001$) was observed with significant differences ($p<0.001$) between Standard and each of the other readers.
- *Perceived Reading Speed:* A significant main effect ($F(3,78)=35.253$, $p<0.001$) was observed with significant differences between Standard and each of the other readers ($p<0.001$), and also between PopoutText and the physical book ($p=0.001$).
- *Ease of Manipulation:* A significant main effect ($F(3,78)=18.427$, $p\leq 0.002$) was found, with ClearText and Standard being significantly different ($p<0.001$).
- *Adherence To Author's Creative Intent:* A significant main effect ($F(3,78)=25.356$, $p<0.001$) was observed with significant differences between the physical book and each of the other readers ($p<0.001$), and also between ClearText and PopoutText ($p=0.008$).
- *Attractiveness:* A significant main effect ($F(3,78)=34.641$, $p<0.001$) was observed with significant differences between the physical book and each of the other readers ($p<0.001$), and also between ClearText and each of the other readers ($p\leq 0.01$).

These results showed that most (but not all) participants preferred the physical book on all criteria. Among computer interfaces, ratings of individual interfaces preferred ClearText for all criteria. For legibility (“easy to see”), ease of manipulation (“easy to manipulate”), and perceived reading speed (“can read quickly”), PopoutText was rated higher than Standard. Standard was preferred

over PopoutText for preservation of author’s creative intent (“preserves author’s creative intent”) and attractiveness (“attractive”), probably because it makes the least modifications to the page appearance. Results for the 6-point scale questionnaires are shown in Figure 6.

Rankings

Rankings, shown in Figure 7, followed a similar trend to the individual ratings, except that Standard was ranked above ClearText for its closeness to the author’s creative intent. Because these values are interdependent, we did not perform a statistical analysis on the rankings.

DISCUSSION

By far, the most surprising result is that 61% of users actually preferred an electronic method to the physical book for some criteria. 37% ranked an electronic method higher than the physical book for ease of manipulation. In interviews, participants revealed that turning pages of a physical book can be cumbersome, but clicking through an interface, especially ClearText, is much more convenient. 33% ranked either ClearText (7%) or PopoutText (26%) #1 for legibility (“easy to see”). In interviews, several of these users explained that the physical book provides no magnification mechanism, making it more difficult to see the detail of the illustrations. Three participants ranked ClearText #1 for both perceived reading speed and ease of manipulation, probably because page turning affects the amount of time required to finish the book.

Somewhat less surprising was that ClearText was ranked #3 or #4 for its preservation of the author’s creative intent by 41% of the participants, probably because it imposes the greatest amount of change on the book’s layout by removing the original text and replacing it with a different typeface in a slightly different location. More importantly, ClearText does not preserve the styling of non-horizontal text. We expect that if the study were run using just one book, the results for this would vary significantly depending on the style of the original text. For an illustrated business document, this would probably be less significant.

FUTURE WORK

Despite variability in some criteria, ClearText appears to be the most promising direction for the next book viewing

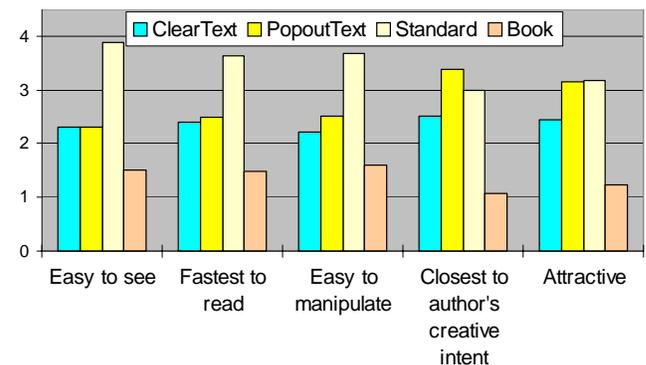


Figure 7: Average rankings (Low values are good.)

interface for the ICDL. However, deploying ClearText for the entire collection of books presents several challenges. First, we need to secure permission to do the transcriptions. Although the ICDL has already secured permission to display images of the books on the web, that permission does not extend to the raw text for all of the books. Furthermore, removing the text from the pages may be legally construed as a derivative work, further requiring special legal permission from the publishers.

Assuming we can secure the needed permissions, we will also need to further develop the technical means and human network required to process and check the removal of the text and to transcribe the text.

All of this is currently being worked on. Ultimately, the goal is to offer readers the choice of all three readers for most books in the ICDL. Some books will be exempt either because of a lack of permission or because the content does not lend itself well to the ClearText approach. Since PopoutText requires no image manipulation and no transcription, no additional permissions will be necessary. Based on the results of this study, we expect to be able to offer ClearText and PopoutText for all users who have JavaScript enabled in their browsers. Standard may continue to be the only option for users who choose to disable JavaScript.

As stated above, the improvements to the ICDL are only an intermediate goal of this work. The display of scanned books in a web browser is a broader issue affecting many digital libraries. The expectation is that by demonstrating the effectiveness and feasibility of ClearText and PopoutText in the ICDL, they will become attractive options for other digital libraries.

ACKNOWLEDGEMENTS

We appreciate all the members of the International Children's Digital Library team (especially the children!) for their help in designing and understanding these interfaces. We also appreciate the statistical advice of François Guimbretière and the ability to use the beautiful books of Andrea Petrlík Huseinović.

Finally, this work could not have been done without the support of the National Science Foundation (contract #0205082).

REFERENCES

- [1] Adobe Reader desktop application home page. www.adobe.com/products/acrobat. September 1, 2007.
- [2] Adobe Reader Mobile home page. September 1, 2007. www.adobe.com/products/acrobat/acrrmobiledevices.html
- [3] Baldwin, T. S. and Baily, L. J. 1971. Readability of technical training materials presented on microfiche versus offset copy. *J. Appl. Psychol.* 55, 37-41.
- [4] Bernard, M. L., Chaparro, B. S., Mills, M. M., and Halcomb, C. G. 2003. Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *Int. J. Hum.-Comput. Stud.* 59, 6 (Dec. 2003),

- 823-835. DOI= [http://dx.doi.org/10.1016/S1071-5819\(03\)00121-6](http://dx.doi.org/10.1016/S1071-5819(03)00121-6)
- [5] Bernard, M., Chaparro, B., Mills, M., Halcomb, C. Examining children's reading performance and preference for different computer-displayed text. *Behaviour & Information Technology*, 2002, Vol. 21, No. 2, 87-96. DOI=<http://dx.doi.org/10.1080/01449290210146737>
- [6] Betrisey, C., Blinn, J.F., Dresevic, B., Hill, B., Hitchcock, G., Keely, B., Mitchell, D.P., Platt, J.C., Whitted, T.. Displaced Filtering for Patterned Displays. *Proc. Society for Information Display Symposium*, pp. 296-299, (2000). DOI=<http://dx.doi.org/10.1889/1.1832941>
- [7] Björk, S. and Redström, J. 1999. An alternative to scrollbars on small screens. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, May 15 - 20, 1999). CHI '99. ACM Press, New York, NY, 316-317. DOI= <http://doi.acm.org/10.1145/632716.632908>
- [8] Boyarski, D., Neuwirth, C., Forlizzi, J., and Regli, S. H. 1998. A study of fonts designed for screen display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Los Angeles, California, United States, April 18 - 23, 1998). C. Karat, A. Lund, J. Coutaz, and J. Karat, Eds. *Conference on Human Factors in Computing Systems*. ACM Press/Addison-Wesley Publishing Co., New York, NY, 87-94. DOI= <http://doi.acm.org/10.1145/274644.274658>
- [9] Cockburn, A., Gutwin, C., and Alexander, J. 2006. Faster document navigation with space-filling thumbnails. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006). R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, and G. Olson, Eds. *CHI '06*. ACM Press, New York, NY, 1-10. DOI= <http://doi.acm.org/10.1145/1124772.1124774>
- [10] Combs, T. T. and Bederson, B. B. 1999. Does zooming improve image browsing?. In *Proceedings of the Fourth ACM Conference on Digital Libraries* (Berkeley, California, United States, August 11 - 14, 1999). DL '99. ACM Press, New York, NY, 130-137. DOI= <http://doi.acm.org/10.1145/313238.313286>
- [11] Dillon, A. 1992. Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics*, 35(10), 1297-1326. DOI=10.1080/00140139208967394
- [12] Dillon, A. (1999) TIME - A multi-level framework for the design and evaluation of digital libraries. *International Journal of Digital Libraries* 2(2/3), 170-177.
- [13] Dillon, A. 1994 *Designing Usable Electronic Text: Ergonomic Aspects of Information Usage*. Taylor & Francis, Inc.
- [14] Dyson, M.C. How physical text layout affects reading from screen. *Behaviour & Information Technology*, Volume 23, Number 6, November-December 2004. Taylor & Francis (2004), 377-393. DOI=<http://dx.doi.org/10.1080/01449290410001715714>
- [15] Furnas, G. W. 1986. Generalized fisheye views. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, Massachusetts, United States, April 13 - 17, 1986). M. Mantei and P. Orbeton, Eds. *CHI '86*. ACM Press, New York, NY, 16-23. DOI= <http://doi.acm.org/10.1145/22627.22342>

- [16] Gibson Research Corporation. www.grc.com/ctwhat.htm, accessed 3/1/2007.
- [17] Highslide JS thumbnail image viewer. <http://vikjavev.no/highslide/>, accessed 9/1/2007.
- [18] Hornbæk, K. and Frøkjær, E. 2001. Reading of electronic documents: the usability of linear, fisheye, and overview+detail interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Seattle, Washington, United States). CHI '01. ACM Press, New York, NY, 293-300. DOI= <http://doi.acm.org/10.1145/365024.365118>
- [19] Hornbæk, K. and Frøkjær, E. 2003. Reading patterns and usability in visualizations of electronic documents. ACM Trans. Comput.-Hum. Interact. 10, 2 (Jun. 2003), 119-149. DOI= <http://doi.acm.org/10.1145/772047.772050>
- [20] Hourcade, J. P., Bederson, B., Druin, A., Rose, A., Farber, A., & Takayama, Y. The International Children's Digital Library: Viewing Digital Books Online, in Proceedings of the International Workshop 'Interaction Design and Children' (Eindhoven, The Netherlands, August 2002), 125-129.
- [21] Huseinović, Andrea Petrlik. 2001. *The Blue Sky*. Kašmir Promet (Croatia, 2001).
- [22] Huseinović, Andrea Petrlik. 2003. *Ciconia Ciconia*. Kašmir Promet (Croatia, 2003).
- [23] Jones, S., Jones, M., Marsden, G., Patel, D., and Cockburn, A. 2005. An evaluation of integrated zooming and scrolling on small screens. Int. J. Hum.-Comput. Stud. 63, 3 (Sep. 2005), 271-303. DOI= <http://dx.doi.org/10.1016/j.ijhcs.2005.03.005>
- [24] Kaplan, N. and Chisik, Y. 2005. Reading alone together: creating sociable digital library books. In *Proceeding of the 2005 Conference on Interaction Design and Children* (Boulder, Colorado, June 08 - 10, 2005). IDC '05. ACM, New York, NY, 88-94. DOI= <http://doi.acm.org/10.1145/1109540.1109552>
- [25] Laarni, J. 2002. Searching for optimal methods of presenting dynamic text on different types of screens. In Proceedings of the Second Nordic Conference on Human-Computer interaction (Aarhus, Denmark, October 19 - 23, 2002). NordiCHI '02, vol. 31. ACM Press, New York, NY, 219-222. DOI= <http://doi.acm.org/10.1145/572020.572050>
- [26] Liu, H., Xie, X., Ma, W., and Zhang, H. 2003. Automatic browsing of large pictures on mobile devices. In Proceedings of the Eleventh ACM international Conference on Multimedia (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, 148-155. DOI= <http://doi.acm.org/10.1145/957013.957045>
- [27] Mills, C. B. and Weldon, L. J. 1987. Reading text from computer screens. ACM Comput. Surv. 19, 4 (Dec. 1987), 329-357. DOI= <http://doi.acm.org/10.1145/45075.46162>
- [28] Osborne, D. J. and Holton, D. Reading from screen versus paper: there is no difference, *International Journal of Man-Machine Studies*, v.28 n.1, p.1-9, January 1988
- [29] Plaisant, C., Carr, D., and Shneiderman, B. 1995. Image-Browser Taxonomy and Guidelines for Designers. IEEE Software. 12, 2 (Mar. 1995), 21-32. DOI= <http://dx.doi.org/10.1109/52.368260>
- [30] Robertson, G. G. and Mackinlay, J. D. 1993. The document lens. In Proceedings of the 6th Annual ACM Symposium on User interface Software and Technology (Atlanta, Georgia, United States). UIST '93. ACM Press, New York, NY, 101-108. DOI= <http://doi.acm.org/10.1145/168642.168652>
- [31] Shneiderman, B., Plaisant, C. Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition). Addison Wesley (2004).
- [32] Schilit, B. N., Golovchinsky, G., and Price, M. N. 1998. Beyond paper: supporting active reading with free form digital ink annotations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Los Angeles, California, United States, April 18 - 23, 1998). C. Karat, A. Lund, J. Coutaz, and J. Karat, Eds. Conference on Human Factors in Computing Systems. ACM Press/Addison-Wesley Publishing Co., New York, NY, 249-256. DOI= <http://doi.acm.org/10.1145/274644.274680>
- [33] SubLCD. <http://www.oyhus.no/SubLCD.html>
- [34] Tullis, T. S., Boynton, J. L., and Hersh, H. 1995. Readability of fonts in the windows environment. In Conference Companion on Human Factors in Computing Systems (Denver, Colorado, United States, May 07 - 11, 1995). I. Katz, R. Mack, and L. Marks, Eds. CHI '95. ACM Press, New York, NY, 127-128. DOI= <http://doi.acm.org/10.1145/223355.223463>
- [35] University of Texas home page. Screen Resolution Statistics. Accessed September 1, 2007. www.utexas.edu/teamweb/reports/screen_resolution/index.php
- [36] W3Schools. Browser Display Statistics. Accessed 9/1/2007. www.w3schools.com/browsers/browsers_display.asp
- [37] Waycott, J. and Kukulska-Hulme, A. 2003. Students' experiences with PDAs for reading course materials. Personal Ubiquitous Comput. 7, 1 (May. 2003), 30-43. DOI= <http://dx.doi.org/10.1007/s00779-002-0211-x>
- [38] Yee, K. 2003. Peephole displays: pen interaction on spatially aware handheld computers. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM Press, New York, NY, 1-8. DOI= <http://doi.acm.org/10.1145/642611.642613>
- [39] Ziefle, M. Effects of display resolution on visual performance. Human Factors. Volume 40, number 4. Human Factors and Ergonomics Society, Santa Monica, CA (1998), 554-568. DOI=10.1518/001872098779649355