PageSpark: An E-Magazine Reader with Enhanced Reading Experiences on Handheld Devices

Jiajian Chen
Georgia Institute of Technology
Atlanta, GA 30318 USA
jchen30@mail.gatech.edu

Jun Xiao
HP Labs
Palo Alto, CA 94304 USA
jun.xiao2@hp.com

Jian Fan
HP Labs
Palo Alto, CA 94304 USA
jian.fan@hp.com

Eamonn O'Brien-Strain
HP Labs
Palo Alto, CA 94304 USA
eamonn.obrien-strain@hp.com

ABSTRACT
In this paper we present PageSpark, a system that automatically converts static magazine content to interactive and engaging reading apps on handheld reading devices. PageSpark enhances the reading experience in three general aspects: page layout reorganization, page element interactions and page transitions. We explored and implemented several design variations in each aspect with the prototype running on the iPad. Participants from our initial user study showed strong interest of using PageSpark over existing magazine reading apps.

Author Keywords
PageSpark, document analysis, page segmentation, page transition, page layout.,

ACM Classification Keywords
H.5.1. Information interfaces and presentation (e.g., HCI): Multimedia Information Systems.

General Terms
Algorithms, Design, Human Factors.

INTRODUCTION AND RELATED WORK
The fast development in both mobile services and digital publishing is transforming people's way of consuming media content. A growing range of e-readers and tablets, such as the iPad and Kindle, are available for people to read digital magazines, newspaper and books. These reading appliances are handheld, lightweight, and have superior displays compared to traditional PC monitors. However, interaction design for such reading devices is still an active research area in need of attention to make the reading experience as relaxing and engaging as reading prints.

Various researches have been conducted to develop better user interaction mechanisms for common tasks such as emailing and web-browsing on mobile devices with small displays. For example, Wobbrock et al. presented WebThumb and discussed interaction techniques for a small screen browser [6]. Lank and Phan described a *focus+context* sketching on Pocket PC devices [4]. The main focus of our research, however, is on e-reader devices that are close in size and resolution to traditional books, and on improving reading experience for e-magazines.

EICS'11, June 13–16, 2011, Pisa, Italy.

Figure 1. PageSpark UI snapshots for PC World Magazine.

Not surprisingly, the prevailing interaction metaphor for e-readers models after paper and their appearance and affordance. Seminal studies, such as the one published by Adler et al. looked at how people use paper documents in their everyday life and discussed design implication for digital reading devices [1]. However, such studies often centered on professional and work related scenario and the comparison between reading on paper and reading on screen often dwelled on reading speed and comprehension.

Reading magazines, we argue, is often more for recreational purpose than for information gathering. The readers' task goal, if there is any, is not to study or learn the content but to follow and enjoy the storyline. Digital publishing brought a range of features to books that would not be possible in print, such as built-in dictionary and search. However, because of the linear, narrative nature of the magazine content, we believe presentation and appearance play crucial roles rather than hyper-textual navigation structures. Wilson et al. suggested that visual cues should be adapted to exploit the potential of the medium [5].

PageSpark offers an attempt to achieve that. It automatically converts static PDF magazines to interactive multimedia apps running on the iPad. The iPad is our platform of choice because of its form factor and display, and iOS programming interface, but our design concepts and developed technologies may well be extended to other tablets and e-reader platforms. Similarly, our analysis

engine only works on PDF files now because the PDF format is prevalent in digital distribution of magazine content, but other document analysis technologies that work on web content or scanned document images may well be incorporated into the PageSpark system.

SYSTEM ARCHITECTURE

There are three main components in PageSpark. Given input of a PDF magazine (e.g., National Geographic), first we use our PDF segmentation algorithms to analyze the content of each page to extract its text and image. Each text and image ("information block") is labeled for its semantics, such as an article title and author name, and stored separately. Second we compute the reading order of the text content of the articles, link multi-page images and articles and generate an XML description file for this magazine. After these two offline steps, PageSpark parses the XML and maps the semantics of the mark up in runtime into interactive behaviors in the app that runs on the iPad. The system architecture of PageSpark is shown below.

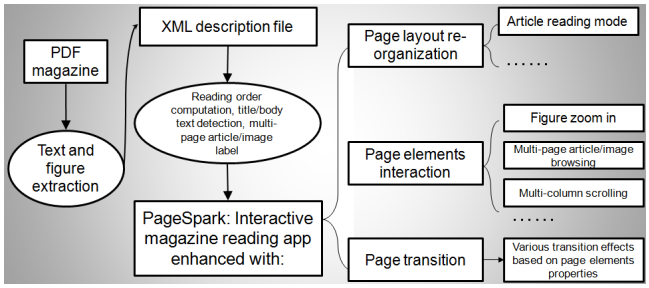


Figure 2. System architecture.

PDF SEGMENTATION ALGORITHM

Even though PDF (portable document format) files truthfully preserve the intended visual appearance of electronic documents, they do not maintain logical and semantic structures such as paragraphs, titles and captions that are more meaningful for readers to interact with. The goal of our PDF segmentation is thus to identify these semantic structures from unstructured internal PDF data utilizing some visual properties.

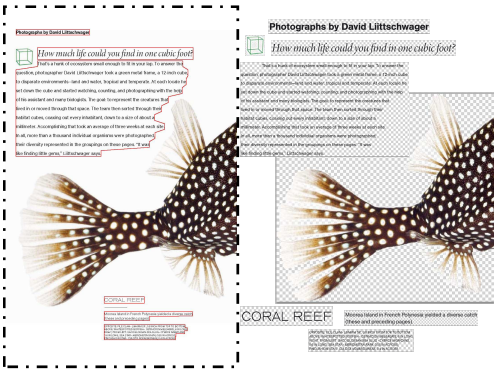
Text Grouping

In PDF files, the text is represented as words with attributes of font name, font size, color and orientation [3]. The task of text grouping is to group words into text lines, and group text lines to text segments/paragraphs. Currently we only deal with text of horizontal and vertical orientations.

To group words into lines, we start a new line and add an available word to the line. We then look for available words to add to the line on both the left and the right ends. The conditions for a word to be added to the line are: 1) the difference between the font size of the words and the font size of the line must not exceed one point, 2) the horizontal distance between the word's bounding box and the line bounding box must be less than the nominal character space for the font and it is the smallest among all available words, 3) the vertical overlap between the word bounding box and the line's must be more than 40%. When no word can be added to the current line, we start a new line. It should be pointed out that we do not require words in a line to have the same font style. This is based on our observation that many documents include URL links and names that have different font styles. For each text line, we compute and maintain two key metrics of font size and central location weighted by lengths of words. To group text lines into segments, we sort the text lines in top-down fashion. We start with a new segment and add an available line to it. We grow the segment by adding lines to it (see Figure 3 (a)).

Image Object Segmentation

The need for image segmentation arises due to two reasons: 1) A PDF image object may contain multiple semantic image objects,2) the bounding box of an image object may overlap with other image objects or text regions and thus limits the display and interaction options. We classify foreground/background pixels by computing color distance. The connected component analysis is followed to identify image objects from foreground pixels (see Figure 3(b) as an example).



(a) **(b)**

Figure 3. An example of the PDF segmentation.

XML DESCRIPTION

We proposed an XML based description format to organize the results of PDF segmentation. In this format every information block in each page is stored as a node in a hierarchical tree structure in an XML file. For each information block we store its position, size, text content and reference images as attributes of its corresponding node. An example of a node is shown in Figure 4.

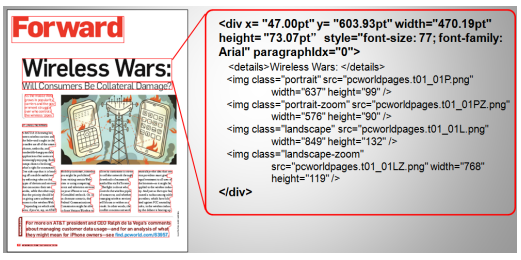


Figure 4. Segmentation result of an magazine page and its XML description of the 'wireless war' information block.

INTERACTION DESIGN

PageSpark enhances user's reading experiences beyond simple zooming and paging in three general aspects based on page segmentation analysis: page layout reorganization, page elements interaction and page transition. Page layout reorganization intelligently computes and reorganizes article content for better reading. Page elements interaction allows users to interact with every piece of text and image content of the article. Page transitions add visually appealing effects to increase reader engagement.

Page layout reorganization

The first enhancement we designed is page layout reorganization. Our goal is increasing readability by reorganizing the layout of page contents. Magazine articles usually have a multi-column style. We found that the font size in the columns is usually too small to read easily even on handheld devices with middle-size displays in portrait view. Traditional PDF readers allows user to zoom in to look, but it is not a good solution from readers' perspective. Modern e-readers such as Kindle or Sony Reader provide specially designed format with proper font size for e-pubs suitable for reading on these devices, but it requires a format redesign of the content. We designed an 'Article Reading Mode' to cope with this problem. In this mode, PageSpark leverages the result of PDF segmentation and automatically puts all text content of an article together to form a clear single reading scroll, as shown in Figure 5.

To form a single reading column in the correct order we use the following rule-table based heuristic algorithm to compute the reading order for each text block in the article.

Rule Set	Rank
Font size and style	1
TextBlock.origin.x	2
TextBlock.origin.y	3
TexBlock Column Width	4

Table 1. Rule table for computing reading order.

Given a set of text blocks of an article, we run a two pass algorithm to compute the reading order for each block. In the first pass, based on the rule table we distinguish titles, footnotes with the main body text. We also create several buckets based on the width of information block to find a group of blocks that has smallest variation in width. Combining these two steps we can distinguish main body text from other types of information blocks. In the second pass, we compute the reading index of each main body text block based on its position in the original page layout.

We animate the transition between the original page layout and the article reading mode. When use taps inside the content of an article on the original PDF page, text blocks of this article pop up and fly to form the long article reading scroll. It also zooms in and scrolls to the exact location in the article where user's finger touches.

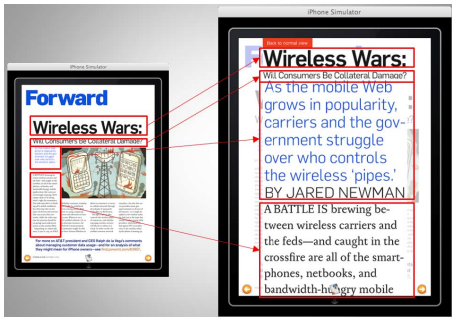


Figure 5. Article reading mode of a page in portrait view.

The article reading scroll takes 75% of width of the screen (576 pixels) in the portrait mode on the iPad. In the meantime the original page is rendered semi-transparently as the background. User can tap the background to switch back to the ordinary page view. We choose these parameters because we find that these settings make reading multi-column articles much easier not only on the iPad, but also on other devices with middle-size displays.

Page elements interaction

In general the goal of page elements interaction is make every piece of the magazine page interactive. There are many possible design variations for page elements interactions. We implement three enhancements to demonstrate the possibilities: multi-column scrolling, multi-page article/image browsing, and single figure zooming in.

In the previous section we presented the article reading mode. We found this mode is very useful for reading article in portrait. However, we found in general a multi-column article is readable on the iPad in landscape mode. In this case, we designed and implemented a multi-column scrolling mechanism to enhance reading experiences. Braganza et al. concluded from their studies that horizontal paging with multi-column layout is superior to scrolling in terms of performance and user preference [2]. But we found that the main problem is that by the time the user finishes reading the last line of the previous column, s/he has to scroll the whole page all the way back to the top to read to the next column. The continuity of reading, as a result, is lost. In PageSpark we render each column of the article independently in landscape mode. Therefore, each column of the article is independently scrollable to provide continuous reading experiences for the users.

Figure 6. User can scroll each single column of the article.

The second feature, multi-page article/image browsing allows user to get a quick overview of an article or image than spans multiple pages. When the user taps the margin

area of a page that belongs to a multi-page article or image spread, the current page zooms out and its adjacent article/image pages slide in to form an overview of the whole article or image. User can quickly jump to any page of the article by tapping a thumbnail in this mode, as shown in Figure 7.

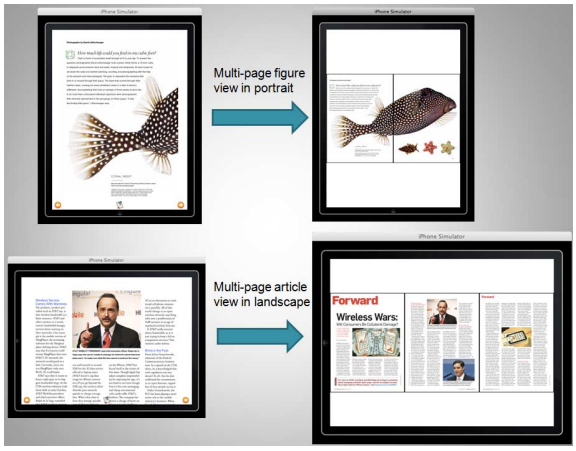


Figure 7. Multi-page article/image browsing mode.

Lastly, single figure zooming, is a feature that zooms in a figure when user touches it, fits the figure to the screen size and fades out the background page.

Page transition

Another key enhancement of PageSpark compared with existing PDF magazine readers is that we apply various visually-appealing transition effects intelligently to different document elements. The purpose of applying these transitions is to better convey the content structure to the readers by distinguishing text from images, and headings and titles from body text and callouts in animations and transitions. We pre-defined many types of transition effects. When user switches between article pages each information block (e.g., text, figure, title) has its corresponding transition effect (e.g., fade in/out, slide in/out, and cross-dissolve). Besides the article page transition, we also designed and implemented several special transitions for ads pages, such as highlighting.

USER STUDIES

In a preliminary user study we interviewed and observed 10 people to gather information of their reading behaviors on handheld devices. We discussed with them about extra value for digital magazines as well as any discomfort of reading them on screen vs. on paper. This preliminary study provided us many insights to build PageSpark prototype.

In a post user study, we tested our system with a full 155 page national geographic magazine. We asked 10 participants to freely interact with the magazine in PageSpark, and in a traditional PDF reader on the iPad. We interviewed them after this session and asked open ended questions. People expressed a strong preference of using PageSpark over the PDF reader to view the magazine. Using PageSpark, participants showed a greater level of engagement. One explanation is simply the fun of seeing how the system reacts and the novelty of the system. But tangibility and responsiveness are often mentioned by the users as the main reason they like the system. They commented that PageSpark made the magazine "live".

DISCUSSIONS AND FUTURE WORK

PageSpark enhances reading experiences in three general aspects. We have implemented one instance for each aspect, but we just touched the tip of an iceberg. There are many design variations that can be explored under the PageSpark's general framework for different types of digital content.

For example, besides the article reading mode, we can apply other types of page layout reorganization, such as removing unrelated content or adding additional content. This may not be suitable for professionally designed magazines, but it may be useful for web articles since web pages often contain a large number and area of unrelated content, such as ads.

For the page elements interaction, we can break the linear structure of the magazine and provide a more natural non-linear navigational help for the readers. For example, since we can extract the semantic meaning of any page entities, names and keywords found on the pages can be indexed for search. User can then simply tap a photographer's name to retrieve all the photos taken by this photographer across the entire magazine collection.

Finally, for the page transition, we may apply different transition templates or styles for different types of content. For example, static print advertisement can be automatically converted into animated display ads using the PageSpark system. The key, however, is to find the right balance between attention grabbing and being annoying.

REFERENCES

1. Adler, A., Gujar, A., Harrison, B. L., O'Hara, K., and Sellen, A. A diary study of work-related reading: Design implications for digital reading devices. *CHI' 98*, pp. 241-248. New York, NY.
2. Braganza, C., Marriott K., Moulder P., Wybrow, M. and Dwyer T. Scrolling behavior with single- and multi-column layout, Proceedings of *WWW'09*.
3. Chao H. and Fan J. Layout and Content Extraction for PDF Documents, *DAS*'04, *LNCS* 3163, pp. 213–224.
4. Lank, E. and Phans, S. Focus+context sketching on pocket PC. *CHI '04*. pp.1275-1278.
5. Wilson, R., Landoni, M. and Gibb, F. (2002) A user-centred approach to e-book design. The Electronic Library, 20 (4).
6. Wobbrock, J., Forlizzi, J., Hudson, S. and Myers, B. WebThumb: interaction techniques for small-screen browsers. *UIST'02*. pp.205-208.