

Using Thumbnails to Search the Web

Allison Woodruff^{*}, Andrew Faulring^{**‡}, Ruth Rosenholtz^{*},
Julie Morrison^{*}, Peter Pirolli^{*}

^{*}Xerox Palo Alto Research Center
3333 Coyote Hill Rd.
Palo Alto, CA 94304 USA

[‡]Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA

ABSTRACT

We introduce a technique for creating novel, textually-enhanced thumbnails of Web pages. These thumbnails combine the advantages of image thumbnails and text summaries to provide consistent performance on a variety of tasks. We conducted a study in which participants used three different types of summaries (enhanced thumbnails, plain thumbnails, and text summaries) to search Web pages to find several different types of information. Participants took an average of 67, 86, and 95 seconds to find the answer with enhanced thumbnails, plain thumbnails, and text summaries, respectively. We found a strong effect of question category. For some questions, text outperformed plain thumbnails, while for other questions, plain thumbnails outperformed text. Enhanced thumbnails (which combine the features of text summaries and plain thumbnails) were more consistent than either text summaries or plain thumbnails, having for all categories the best performance or performance that was statistically indistinguishable from the best.

Keywords

Thumbnails, Web search task

INTRODUCTION

Internet users spend a significant amount of time examining search engine results; one commercial search engine vendor claims to answer 40 million search queries each day [16]. The user must page through lists of Web documents, briefly evaluating each for possible relevance to a particular information need. Improving the efficiency of this tedious process directly benefits the end-user and, by improving end-user satisfaction, indirectly benefits the search engine vendor.

The search engine can increase user efficiency by either (1) returning higher-quality document lists (e.g., through better index coverage and ranking algorithms) or (2) providing information that allows the user to evaluate the results more quickly. Search engine vendors attack both problems. The

standard practice with regard to approach (2) is to provide brief textual summaries of the Web documents. We believe that this latter practice can be improved upon.

We have performed a quantitative comparative study of textual and graphical summarization mechanisms applied to search engine results. We argue that graphical summaries of the documents – thumbnail images – can greatly increase the efficiency by which end-users process search engine result sets. For example, thumbnails allow users to classify a Web page's genre very rapidly. Most interestingly, our empirical results suggest that, if properly designed, *textually-enhanced thumbnails* deliver the efficiency benefits of both textual summaries and unenhanced thumbnails.

To understand why this might be the case, one must understand the relative advantages and disadvantages of presenting information in textual and graphical form. We now turn to a brief discussion of the relative tradeoffs, with particular attention paid to the specific application of Web search results.

Text summaries are terse but are verbal rather than visual. They can be quickly downloaded, and often contain a great deal of valuable information about each document. For example, search engines commonly provide the document's URL, title and size, as well as a few sentences that either summarize the document or contain some of the search keywords. On the other hand, the user must evaluate the document's relevance by reading the text summary. Text summaries do not provide much information about the page layout or any images contained in the page. Reading lists of search results is tiring, and empirical studies show that the average search engine user is unwilling to read through more than a few pages of such listings.

Simple graphical summaries have largely complementary strengths and weaknesses. Thumbnails are typically larger and therefore slower to download than text summaries. Textual content in simple thumbnails is less accessible, as it is difficult to read and is not conveniently summarized. In contrast, graphical summaries do provide information about the layout, genre, and style of the page. If the user has previously seen the page, or one like it, the visual representation may aid in recognizing or classifying it. This becomes even more compelling in view of the fact that the human visual system can process images more quickly than

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.

SIGCHI'01, March 31-April 4, 2001, Seattle, WA, USA.
Copyright 2001 ACM 1-58113-327-8/01/0003...\$5.00.

text. Graphical information can speed many tasks tremendously. We can get the “gist” of an image in 110 ms or less, changing fixation roughly every 300 ms [5]. In that 110 ms, we can on average read less than 1 word, or skim 2 words. (The average reader of English reads about 4.2 words per second, and can skim or scan at roughly 17 words per second [3].) This is borne out by the fact that searching for a *picture* of a particular object among other pictures is faster than searching for the *name* of that object among other words [12].

In this paper we compare text summaries with plain thumbnails (simple reduced-size images), as well as with enhanced thumbnails, which we have designed in the hopes of capturing some of the advantages of both text summaries and plain thumbnails. We make several contributions:

- We present designs for enhanced thumbnails that work to combine the advantages of both text summaries and plain thumbnails. These designs involve presenting a reduced image of the document along with various forms of emphasis of information in the document. Previous work has generally presented plain thumbnails [e.g. 1, 8, 9, 14], and/or represented properties of the document in an abstract form [e.g. 4, 20]. In some cases we modify the document, by changing the HTML, prior to reduction. In particular, our enhanced thumbnails enforce readability of certain parts of the document within the thumbnail and display highlighted keywords transparently overlaid on the reduced document.
- Much of the previous work on using thumbnails has emphasized using them for recall of previously seen documents. We focus on using them in an application in which the user is unlikely to have seen many of the documents before.
- We present a study comparing the effects of text summaries, plain thumbnails, and enhanced thumbnails on realistic search tasks. Enhanced thumbnails had better and more consistent performance than the other summary types.

In the next section we discuss related work. In the subsequent sections, we discuss our system for generating thumbnails, our study to compare text summaries with plain and enhanced thumbnails in a search task, and future work and conclusions.

RELATED WORK

Previous work has used a number of different designs for thumbnails. A number of programs use plain thumbnails. These include many graphical editors, recent versions of Microsoft® Windows®, and the systems described by [8, 9, 14], among others. Ayers and Stasko’s thumbnails consist of a reduced view of the upper left corner of a document [1]. Cockburn et al. [4] generated thumbnails that consist of reduced images plus “dogears” that indicate bookmarked and frequently visited pages.

Rather than rendering a reduced image of a page, Wynblatt and Benson [20] produce Web page “caricatures.” These caricatures contain select features of a page, often rendered in an abstract form: title, representative image, number of images, abstract, etc. These caricatures do not preserve layout and lack some of the visual information that might be naturally available in a reduced image of the page. For example, link density of a Web page is represented by the background color of the caricature rather than allowing the user to judge the density from an image of the page.

TileBars [7] are abstract representations of documents that graphically indicate the text segments in which search terms appear. Our enhanced thumbnails show the relationship between the occurrence of search terms in the context of the document, and at a finer granularity, but do not provide as compact an overview of the relationship between search terms as TileBars.

A number of systems employ thumbnails. Much of the previous work in this area involves previously viewed documents, in the hope that a thumbnail preview may help the user’s memory and thus aid in the task. A common task has been navigating through previously viewed Web pages [e.g. 1, 2, 4, 8, 14]. In addition, a number of systems use thumbnails to aid in the management and retrieval of files on a user’s computer, for which it is reasonably likely that the user already would have seen the document or image represented by the thumbnail. Graphical editors, for instance, allow the user to preview an image or a collection of images. Recent versions of Microsoft® Windows® provide a thumbnail view of documents within a folder, supporting file formats such as HTML as well as image formats.

Kopetzky and Mühlhäuser [9] describe a system in which links from a Web page are represented by a document thumbnail that appears temporarily upon a mouseover of the link. Though in many cases the user would not have previously seen the documents represented by these thumbnails, the authors again justify the use of thumbnails as a memory aid.

This focus on thumbnails as an aid to memory in retrieving previously seen documents leads us to ask whether thumbnails are only useful when the user has already seen the corresponding documents. In this paper we examine the use of thumbnails in a Web search task, in which few, if any, of the documents are likely to have been previously viewed.

In addition to creating applications that use thumbnails, researchers have studied the utility of thumbnails in a memory task. Czerwinski et al. [6] ask users to spatially lay out 100 Web pages in Data Mountain, and then measure their performance at retrieving those documents a few months later. After a brief learning period, users were just as good at retrieval whether thumbnails were present, or only plain white boxes representing the documents. This

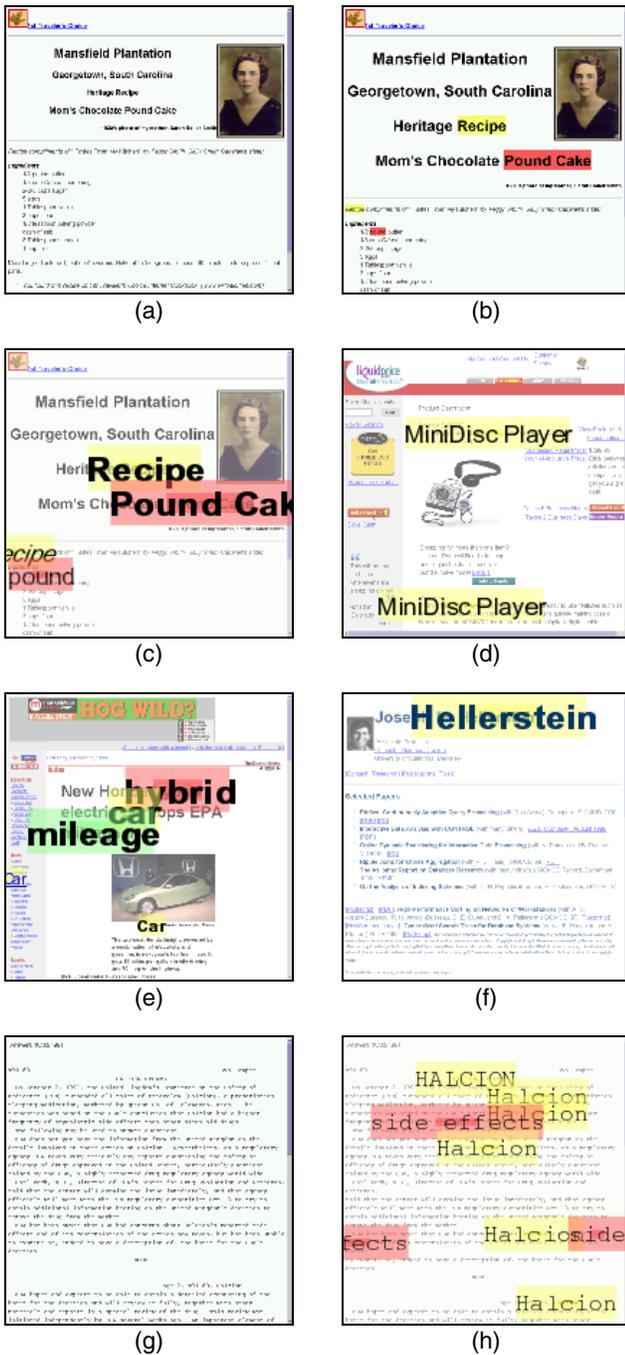


Figure 1. (a) Plain thumbnail. (b) Thumbnail enhanced with HTML modification. (c) Thumbnail enhanced with HTML and image modification. (d) E-commerce genre example. (e) News genre example. (f) Homepage genre example. (g) Plain thumbnail of textual page. (h) Enhanced thumbnail of textual page. (See color plate on page 552.)

might suggest a lack of utility for thumbnails, but the study may underestimate the importance of thumbnails, as users saw their layout with thumbnails present repeatedly throughout the study. Interestingly, users subjectively ranked the thumbnail images as the most helpful feature for retrieval.

SYSTEM

We implemented a system that generates both plain and enhanced thumbnails of HTML documents. The tool is written entirely in Java and utilizes a component Web browser, ICE Browser 5 [17]. The component browser provides access to the document as both an HTML document (source form) and a 2D graphics object (rendered form). As we will see, having convenient access to both interfaces greatly simplified the internal structure of the system.

Our system consists of three basic components. The preprocessor modifies the HTML in the original page, e.g., to change the color or size of certain elements. The renderer creates a scaled version of the modified HTML. The postprocessor modifies the image output by the renderer, e.g., to reduce its contrast or to add text callouts. This architectural separation is due to the fact that the various transformations are most easily applied to the document in different intermediate formats. The final thumbnails take on the order of a few seconds to generate from the raw HTML. In this section, we describe these components in turn. We then discuss some design issues that cut across the components.

HTML Modification

After retrieving the HTML document associated with a given URL, the preprocessor adjusts the appearance of the HTML elements. The user specifies the desired adjustments using an associative list of phrase/style pairs (or tag/style pairs). For example, the user might specify that each instance of the word “recipe” should be highlighted in “yellow.” Similarly, the user can specify that the text of each H1 header tag should be a certain size. Compare the plain thumbnail in Figure 1a with the modified thumbnail in Figure 1b. (Note that the examples of thumbnails presented in the paper and used in the experiment show only the top of the Web page, if it is a long document. We have also experimented with full-page thumbnails.)

This functionality is supported as follows. ICE Browser implements portions of the W3C Document Object Model (DOM) Level 1 Specification [19], a standard interface for programatically accessing and modifying HTML documents. The DOM presents the document as a hierarchy of HTML elements, with each element having an associated CSS style definition [18]. We can modify the HTML document’s appearance by manipulating each element’s CSS style.

One particularly useful modification is to adjust an element’s font size such that the text would still be “readable” in the thumbnail, where readability is specified as a given font size in the final rendered image. Compare the header text in Figure 1b with the header text in Figure 1a.

Rendering

This component delegates the rendering of the (modified) HTML to ICE Browser. Since ICE Browser uses the Java2D interface, the scaling factor for the entire document can be specified by a single operation on a graphics context object.

Image Modification

The postprocessor implements a variety of transformations that cannot be expressed in HTML. For the most part, these transformations require some amount of image processing. For example, a color wash may be applied, or additional graphical elements may be overlaid onto the thumbnail.

One useful modification is to render text phrases as *callouts* (enlarged text overlays) on top of the original thumbnail. The system accepts a phrase, a scale factor at which to re-render the phrase, and an alignment parameter for positioning the callout relative to the original position of the phrase within the document. Again, the resulting transformation can be applied easily using Java2D interfaces – but this time, it is applied only to a specified subset of elements. For example, in Figure 1c, the phrase “Pound Cake” was rendered center-aligned over its original position at four times its original size.

Design Issues

The discussion above provides an architectural view of the system and does not capture the many individual decisions involved in its design. These decisions often required significant attention to visual perception and attention management issues. In the remainder of this section, we give design details of a few of our thumbnail enhancement mechanisms.

Visual layering. In contrast with plain thumbnails, enhanced thumbnails incorporate textual elements. These textual elements have the potential to inhibit gist extraction either by distracting the viewer or by occluding other elements on the page. Through experimentation, we discovered that a good solution is to put textual elements in a separate visual layer, thereby allowing the user to quickly identify these elements or ignore them as desired.

We found that an effective way to create this visual layer is to make the textual elements appear as though they are not part of the original HTML document. We experimented with a number of ways of modifying HTML to try to achieve this effect, e.g., dramatically changing the font size, text color, or background color. Through observing a large number of thumbnails, we learned that because HTML documents have such diverse fonts and colors, our changes most often appear as though they are part of the original document. For example, a colored text header generally looks like a header that an HTML author specified for a document, not like an element that we have chosen to emphasize after the document was authored. Similarly, changing the font size of a certain keyword does not draw attention to it. When modifying font sizes of links in full-

size documents, Olston and Chi observed a related problem: participants had difficulty telling the degree to which the sizes had been modified [11].

After concluding that modifying the HTML is not an effective way of creating a separate visual layer, we experimented with a number of ways of modifying the image after it was rendered. We found that opaque overlays tend to occlude much of the thumbnail, making it difficult for the user to extract gist. Further, opaque background colors tend to give the illusion that the element is actually part of the original, unmodified page. Therefore, we alpha-blended the overlay with the original thumbnail, thereby creating transparent overlays. In our experiments, we found an alpha value of 0.5 to give good results. The resulting overlays have the appearance of being additions to the pages, as opposed to being mark-up included by the original author. They also occlude less of the page.

Color management. To further enhance the visibility of overlays, the system washes the original thumbnail with a given color. Saturated colors tend to draw attention more than desaturated colors. We recommend painting over the thumbnail with a white, transparent fill (we used an alpha of 0.4). This effectively desaturates the original thumbnail, reducing the attentional demands on the user from irrelevant items in the original page, and enhancing the attention-grabbing capabilities of items to which we want to draw attention. Notice the difference between Figure 1a and Figure 1c.

We were particularly interested in using overlays to create readable, attention-grabbing callouts of keywords. Sampling of a random collection of Web pages suggested that a large number of pages contain dark text on a light background. For such dark text, light, unsaturated colors yield highlights that most facilitate reading. However, as mentioned above, more saturated colors are better at drawing attention. We used a model of visual search [15] to select highlight colors that were just saturated enough to “pop out” against a typical thumbnail from our corpus. The resulting highlight colors greatly resemble those colors actually found in highlight pens.

By combining these image modification techniques, we are able to create callouts that can be easily detected while skimming, while simultaneously allowing the user to get the gist of the underlying thumbnails.

Resizing of text. The callouts are appropriate for elements to which we want to draw the user’s attention, e.g., keywords. While other textual elements are potentially useful, we do not want them to occupy much of the user’s conscious attention. Since HTML modifications appear to be part of the original document, and therefore do not draw the user’s attention, these modifications are highly appropriate for changes to text such as headers. We find that making the headings readable greatly increases the utility of the thumbnail, but the change is so subtle that

Category	Characteristics	Example Question	Approx. # Answers
Picture	Requires identification of a graphical element	“Find a picture of a giraffe in the wild.”	8/100
Homepage	Requires genre classification (correct pages somewhat textual, many incorrect pages entirely textual)	“Find Kern Holoman’s homepage.”	1/100
E-commerce	Requires genre classification (correct pages highly graphical; incorrect pages highly graphical, e.g., product reviews)	“Find an e-commerce site where you can buy a DVD player. Identify the price in dollars.”	15/100
Side-effects	Requires semantic information (word proximity and position in layout useful, genre useful)	“Find at least three side effects of Halcion.”	20/100

Table 1. Categories of questions performed by participants.

viewers often take advantage of the feature without being consciously aware that the text has been enhanced.

METHODS

We now discuss the study we conducted to compare the thumbnails generated by our system, plain thumbnails, and text summaries.

Participants

Data were collected from eighteen members of the Xerox PARC community, 6 women and 12 men. All were experienced Web users.

Question Categories

We chose four different question categories, and developed three questions for each category. First, participants were asked to locate a picture of a given entity. Second, participants were asked to locate the homepage of an individual they did not know. Third, participants were asked to locate a consumer electronics item for purchase. Fourth, participants were asked to locate three or more side-effects of a given drug. For an example question from each category, see Table 1. In addition to these four categories, we developed six practice questions, e.g., “Find the mileage of a hybrid car.”

These question categories are representative of tasks users commonly perform on the Web. Morrison et al. have developed a taxonomy based on an analysis of over 300 Web users’ responses to what Web activities significantly impacted their decisions and actions. The three most common types were e-commerce (21%), medical (13%), and finding people (9%) [10]. Morrison et al.’s data includes only information that led to a significant action or decision. We included the picture category because we believe it is representative of a common but less “significant” class of queries. The query results for our question categories yield Web pages that are both semantically and visually different. See the “Characteristics” column in Table 1.

Materials

We constructed our materials for the study in three phases:

Archiving Web Pages. Our corpus is based on URLs extracted from search results from Google. For example, for the E-commerce question on DVD players, we programmatically issued a query to Google using the terms “DVD” and “player” and extracted URLs from the result pages. Since the contents of Web pages often change, we downloaded the pages associated with these URLs so that we would have a consistent set of Web pages to show to our participants. Storing the pages locally provides the added advantage that network delays are avoided, so timing is more consistent.

Creating Summaries. After downloading the pages, we created three different summary materials for each page. First, we extracted the Google *text* summary associated with each URL. These summaries include the page’s title, excerpted text with search terms shown in bold, and the URL. Second, we created a *plain* thumbnail of the page (a scaled version of the page as in Figure 1a). Third, we created an *enhanced* thumbnail which had three primary differences from the plain thumbnail: (1) the fonts in H1 and H2 tags were modified so that they would be readable in the thumbnails; (2) highlighted callouts were included for keywords from the search query; and (3) the contrast level in the thumbnail was reduced so that the callouts would be more prominent (see Figure 1c).

Creating Pages Showing Collections of Summaries. For each of the 12 test questions and the 6 practice questions, we chose 100 result pages to present to the participants. While the choice of results and their ordering was largely random, we did modify the data set to remove pages that had errors (e.g., the page at the given URL could not be retrieved), to ensure that no answer appeared in the first 10 items of any collection so that the participants would need to examine at least 10 summaries for each question, and so that approximately the same number of correct answers appeared in each question associated with a given category (see Table 1).

For each question/type of summary (text, plain thumbnail, enhanced thumbnail) combination, we created a single HTML page that contained the summaries of the 100 result

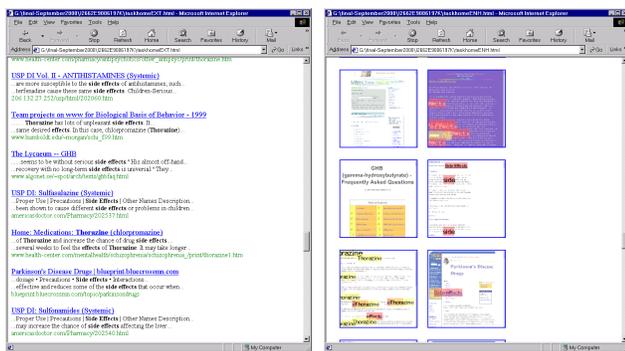


Figure 2. Browser containing text summaries (left) and browser containing enhanced thumbnails (right). (See color plate on page 552.)

pages, with hyperlinks to the actual pages as cached on the local workstation. For the text summaries, the title of the page was a hyperlink. For the thumbnails, the entire thumbnail was a hyperlink.

The text summaries were presented in a single column, using standard Google HTML formatting. The plain and enhanced thumbnails were presented in two columns. We sized the thumbnails to match the size of a typical text summary displayed at a normal font, so as to study the most efficient use of that space. The vertical spacing between the text summaries was the same size as the vertical and horizontal spacing between the plain and enhanced thumbnails. The browser was a consistent size during all experiments, so that approximately seven text summaries and approximately six thumbnails plus small portions of two additional thumbnails were visible on the screen at a given time (see Figure 2).

Procedure

Blocking. For each type of summary page (text, plain thumbnail, enhanced thumbnail) participants completed two practice questions and four test questions. The four test questions included one question of each type mentioned in the Question Category section above (i.e., Picture, Homepage, E-commerce, and Side-effects). The twelve test questions were distributed across all three types of link summary pages.

Process. The participant was introduced to one of the three types of summary page, which were presented in a counterbalanced order across participants. Participants completed two practice questions to familiarize themselves with that type of summary page, and then completed the four test questions. When the first set of test questions was finished, the participant repeated the procedure for the other two summary page types.

After the participant had answered all questions, the experimenter interviewed them about their experiences using the different summary pages. The experiment lasted approximately seventy-five minutes.

Instrumentation. Our instrumentation package consists of a program called WebLogger [13] that records user gestures (such as keystrokes or scrolling) and actions by the browser application (such as loading and rendering pages). We analyzed the data output by WebLogger to extract timing information and the number of page visits per question.

RESULTS

In this section we present our data on total search time and number of pages visited for each summary type. We performed ANOVAs with two within-subjects factors, summary type and question category. Overall, participants needed more time to answer some question categories than others; specifically, they were slower to answer Side-effects ($M=126\text{sec}$, $SD=83.2$) questions than E-commerce ($M=68\text{sec}$, $SD=46.6$), Homepage ($M=77\text{sec}$, $SD=79.6$), or Picture ($M=59\text{sec}$, $SD=58.1$) questions, $F(3,51)=20.28$, $p<.01$. There was no overall difference between participant times across the three summary types, $F(2,34)=2.17$, $p=.13$. Across the four question categories, Picture, Homepage, E-commerce, and Side-effects, participants performed differently depending on what type of summary they used, $F(6,102)=4.07$, $p<.01$.

Figure 3 shows the total search time for the three different summary types. Planned linear contrasts among the three types revealed differences. The contrasts, based on a two-tailed t distribution, are conservative tests of the differences between summary types. Participants answered questions more quickly with enhanced thumbnails ($M=67\text{sec}$, $SD=49.9$) than with text ($M=95\text{sec}$, $SD=78.1$; $t(34)=-2.05$, $p<.05$) or plain thumbnails ($M=86\text{sec}$, $SD=84.4$; $t(34)=-1.35$, $p=.09$), although there were no time differences between plain thumbnails and text, $t(34)=-0.69$, $p=.47$. Because enhanced thumbnails combine positive aspects of both plain thumbnails and text summaries, we used a linear contrast to determine that participants were faster to answer questions with enhanced thumbnails than the average time taken to answer questions with text and plain thumbnails, $t(34)=-1.96$, $p<.05$.

Most interesting was the interaction between summary type and question category. Figure 4 shows the total search time for the three different summary types and the four question categories. The data are averaged over participants, and within each question category we have averaged over the three questions for that category – thus each bar in the graph represents an average over 18 data points. Separate ANOVAs and linear contrasts were used to compare performance on summary type within each question category. Participants differed the most in the time needed to answer Picture questions across the summary types, $F(2,34)=23.28$, $p<.01$. Participants were equally fast to answer questions with the plain and enhanced thumbnails ($t(34)=-0.27$, $p=.39$), and thumbnails in general led to faster performance than text summaries ($t(34)=-3.48$, $p<.01$). For the Homepage category, we also saw differences in summary type, $F(2,34)=2.53$, $p=.10$. Time to complete the

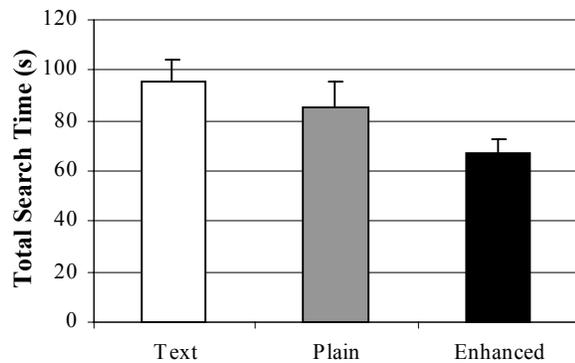


Figure 3. Total search time for the 3 summary types across all four question categories. Error bars show the standard error.

questions in the text and enhanced thumbnail conditions did not differ ($t(34)=-0.13$, $p=.44$), but these times when averaged were faster than the plain thumbnail condition ($t(34)=-2.13$, $p<.05$). For the E-commerce and Side-effects questions, there were no differences in search time across any of the summary types, $F_{E-commerce}(2,34)=1.02$, $p=.37$ and $F_{Side-effects}(2,34)=0.54$, $p=.59$.

The data for the average number of pages visited for the various summary types and question categories followed a similar pattern as the time data, with the exception of the Side-effect question where the differences between summary types were more pronounced. Number of visits did not differ across the question categories ($M_{Picture}=3.9$, $SD=3.7$, $M_{Home}=4.8$, $SD=6.7$, $M_{E-commerce}=4.4$, $SD=3.3$, $M_{Side-effects}=5.6$, $SD=5.0$), $F(3,51)=1.97$, $p=.13$. Participants answering questions with enhanced thumbnails ($M=3.8$, $SD=2.9$) and text ($M=4.4$, $SD=3.7$) visited fewer pages than those using plain thumbnails ($M=5.8$, $SD=6.9$), $F(2,34)=2.95$, $p=.07$. The pattern of visits differed for the three summary types, text, plain thumbnail, and enhanced thumbnail, across the four question categories, Picture, Homepage, E-commerce, and Side-effects, $F(6,102)=7.26$, $p<.01$. Overall, linear contrasts revealed that the visit data and the time data followed similar patterns.

Subject Response

At the conclusion of the experiment, participants were asked about their search strategies and opinions of the three types of summaries, text, plain thumbnails, and enhanced thumbnails. Several of the participants noted that using the enhanced thumbnails was intuitive and less work than using either the text or plain thumbnails. One participant commented that searching for information with text summaries did not seem difficult before he was exposed to searching with the thumbnails. Sixteen of the eighteen participants used the genre information present in the thumbnails. Fourteen participants used cues from the callouts, the relationship between search terms, the location of search terms, or how often the terms appeared, when

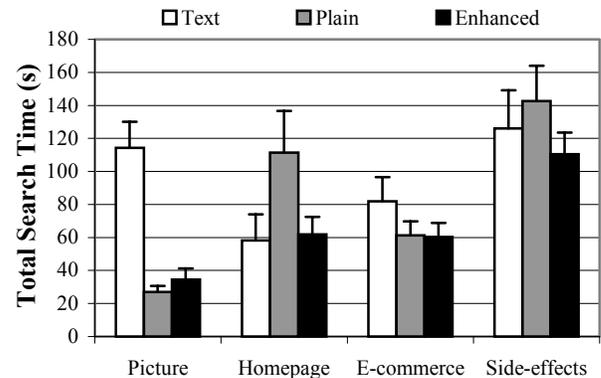


Figure 4. Total search time for the 3 summary types, grouped by question category. Error bars show the standard error.

searching for information with the enhanced thumbnails. Nine participants rated the enhanced thumbnails as their favorite summary type overall, while most others preferred the enhanced thumbnails for certain types of tasks.

DISCUSSION

As one might expect, the relative performance of text summaries, plain thumbnails, and enhanced thumbnails depends greatly upon the question category. For the Picture question, the text summaries required more search time and more pages visited than either type of thumbnail. It makes sense that thumbnails would be more informative for this question, as they allow a user to see the presence of a picture on a page. For the Homepage question, plain thumbnails tended to be worst, which again makes sense – the name of the person, either in a text summary or an enhanced thumbnail, aids in finding their homepage. Though one can perhaps classify a page as a homepage without this text information, such a classification is sometimes misleading, as search results often include homepages for people other than the target. The three types of summaries performed equally well for the E-commerce and Side-effects questions, perhaps because e-commerce and medical sites have strong visual genre cues and layout information as well as useful cues in the text summaries and URLs.

Overall, the relative performance of plain thumbnails and text was variable. These two summary types would sometimes yield the best performance (for tasks for which they were particularly well-suited) and sometimes the worst performance (for tasks for which they were a poor fit). Enhanced thumbnails (which combine the features of text summaries and plain thumbnails) were more consistent than either text summaries or plain thumbnails, having for all categories the best performance or performance that was statistically indistinguishable from the best. This effect is particularly interesting since study participants had developed strategies for using text summaries over a period

of years, and lacked corresponding experience with thumbnails.

CONCLUSIONS AND FUTURE WORK

We have presented enhanced thumbnails that work to combine the advantages of both text summaries and plain thumbnails. We have conducted a study to compare the performance of enhanced thumbnails with plain thumbnails and text summaries. Across the collection of question categories, we found that enhanced thumbnails yielded the best and most consistent performance.

In addition to conducting further studies, we are pursuing several extensions of this work. Items other than search keywords may be emphasized in the thumbnails, e.g., items returned by TFIDF or information scent computations as in [11], or representative images. Callouts on thumbnails might be positioned to minimize their occlusion of each other or of other useful information on the thumbnail, such as readable headers. Another direction we are pursuing is a browsing environment that integrates enhanced thumbnails with enhanced Web pages.

It would also be interesting to consider how one might build thumbnails into a production search engine. Doing so would introduce many significant issues, such as the bandwidth requirements to download the images and the time to generate thumbnails for a given query. Partial pre-computation of the thumbnails may address the latter, but would introduce storage requirements.

ACKNOWLEDGMENTS

We are very grateful to Rob Reeder for providing assistance with WebLogger, Paul Aoki for helpful discussions and comments, and Pam Schraedley for her contributions to the statistical analysis. This research was funded in part by Office of Naval Research Contract No. N00014-96-C-0097 to Peter Pirolli and Stuart K. Card.

REFERENCES

1. Ayers, E. and Stasko, J. Using Graphic History in Browsing the World Wide Web. In *Proc. 4th Intl. WWW Conf.*, December 1995.
2. Card, S. K., Robertson, G. G., and York, W. The WebBook and the Web Forager: An Information Workspace for the World-Wide Web. In *Proc. ACM SIGCHI'96*, April 1996, 111-117.
3. Chapman, A. (ed.). *Making Sense: Teaching Critical Reading Across the Curriculum*. The College Board: NY, 1993.
4. Cockburn, A., et al. WebView: A Graphical Aid for Revisiting Web Pages. In *Proc. OZCHI'99 Australian Conf. on HCI*, Nov. 1999.
5. Coltheart, V. (ed.). *Fleeting Memories: Cognition of Brief Visual Stimuli*. MIT Press: Cambridge, MA, 1999 (pp. 32-70).
6. Czerwinski, M. P., van Dantzich, M., Robertson, G., and Hoffman, H. The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D. In *Proc. INTERACT'99*, 1999, 163-170.
7. Hearst, M. TileBars: Visualization of Term Distribution Information in Full Text Information Access. In *Proc. ACM SIGCHI'95*, May 1995, 59-66.
8. Hightower, R., Ring, L., Helfman, J., Bederson, B., and Hollan, J. Graphical Multiscale Web Histories: A Study of PadPrints. In *Proc. Hypertext'98*, 1998, 58-65.
9. Kopetzky, T. and Mühlhäuser, M. Visual Preview for Link Traversal on the WWW. In *Proc. 8th Intl. WWW Conf.*, May 1999, 447-454.
10. Morrison, J. B., Pirolli, P., and Card, S. K. A Taxonomic Analysis of What World Wide Web Activities Significantly Impact People's Decisions and Actions. Xerox PARC Tech. Report UIR-R-2000-17.
11. Olston, C. and Chi, E. ScentTrails: Integrating Browsing and Searching on the World Wide Web. *In Review*.
12. Paivio, A. Pictures and Words in Visual Search. *Memory & Cognition* 2, 3, 515-521, 1974.
13. Reeder, R. W., Pirolli, P., and Card, S.K. WebLogger: A Data Collection Tool for Web-use Studies. Xerox PARC Tech. Report UIR-R-2000-06.
14. Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., and van Dantzich, M. Data Mountain: Using Spatial Memory for Document Management. In *Proc. ACM UIST'98*, November 1998, 153-162.
15. Rosenholtz, R. A Simple Saliency Model Predicts a Number of Motion Popout Phenomena. *Vision Research* 39, 3157-3163, 1999.
16. "Search Us, Says Google." *Technology Review*, November/December 2000.
<http://www.techreview.com/articles/nov00/qa.htm>
17. Wind River. ICE Browser 5.
<http://www.icesoft.no/icebrowser5/index.html>
18. W3C. CSS, Level 2 (CSS2) Specification.
<http://www.w3.org/TR/REC-CSS2/>.
19. W3C. DOM Level 1 Specification, Version 1.0.
<http://www.w3.org/TR/REC-DOM-Level-1/>.
20. Wynblatt, M. and Benson, D. Web Page Caricatures: Multimedia Summaries for WWW Documents. In *Proc. IEEE Intl. Conf. on Multimedia Computing and Systems*, June 1998, 194-199.