Visual Snippets: Summarizing Web Pages for Search and Revisitation

Jaime Teevan[1], Edward Cutrell[1], Danyel Fisher[1], Steven M. Drucker[1], Gonzalo Ramos[1], Paul André[2], Chang Hu[3]

[1] Microsoft Corporation
Redmond, WA 98052 USA

[2] Electronics & Comp Science
University of Southampton
SO17 1BJ, United Kingdom

[3] University of Maryland
College Park, MD 20742

[1]{teevan, cutrell, danyelf, sdrucker, gonzalor}@microsoft.com; [2]pa2@ecs.soton.ac.uk; [3]changhu@cs.umd.edu

ABSTRACT

People regularly interact with different representations of Web pages. A person looking for new information may initially find a Web page represented as a short snippet rendered by a search engine. When he wants to return to the same page the next day, the page may instead be represented by a link in his browser history. Previous research has explored how to best represent Web pages in support of specific task types, but, as we find in this paper, consistency in representation across tasks is also important. We explore how different representations are used in a variety of contexts and present a compact representation that supports both the identification of new, relevant Web pages and the re-finding of previously viewed pages.

Author Keywords

Thumbnails, Web search, Web browsing, revisitation, re-finding, semantic zoom, visual snippets.

ACM Classification Keywords

H5.2: Information interfaces and presentation: User Interfaces – Graphical user interfaces. H3.3. Information storage and retrieval: Information search and retrieval.

INTRODUCTION

Search and re-finding tasks are among the most common activities on the internet. A Pew Internet and American Life report showed that Web searches were second only to email [15], and studies of revisitation [1] have found that anywhere from 50% [7, 19] to 80% [5] of all Web surfing behavior involves visiting previously visited Web pages. People use search engines [20], bookmarks, browser history mechanisms, and their memory to find and return to Web-based information [4].

In order to accomplish search and re-finding tasks, a user must interact with different representations of Web pages. Search engines typically represent the pages in their result lists as textual snippets, with a title, a query-based page summary, and a URL. Previously viewed Web pages are represented in many ways, including as thumbnails, titles in a user's history, captions within search results, URLs in the address bar, or colored hyperlinks. These different representations are intended to support different tasks.

There are several drawbacks to the existing representations. For one, while individual representations may be well suited to particular navigational tasks, people often navigate to the same Web page in many different contexts. Users may not recognize the thumbnail they see now as the same page as the search snippet they saw before. The success of a representation needs to be considered in the context of a person's entire Web interaction. Additionally, those representations that effectively help people accomplish their task often require valuable screen real estate to do so [10]. This limits a user's ability to see many different Web pages in a search result list or browsing history in one view.

In this paper, we report on a study of 197 people's interactions with compact Web page representations. We analyze the success of each representation in supporting fast navigation to both new and previously viewed content and explore the importance of consistency of representation across different navigational task types.

We find that text snippets are effective for finding new Web pages that have never been seen before. Thumbnails, in contrast, are good for supporting re-finding, but primarily when the page's thumbnail has been seen before. This means that in order for a thumbnail to be useful for re-finding, it needs to be seen initially in a context where it is not particularly useful. A representation we call a *visual snippet* captures the best of these two representations: it supports finding new information comparable to text snippets, and re-finding in a comparable manner to thumbnails – even when it has not been seen before. Visual snippets are designed to maintain the size and visually distinct advantages of thumbnails while containing the same essential elements as text snippets.

Following a review of relevant literature, we discuss how the visual snippets were designed and generated. We then describe the study we conducted to test the effectiveness of visual snippets for both finding and re-finding tasks, compared with thumbnails and text snippets. We conclude

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.
Copyright 2009 ACM 978-1-60558-246-7/09/04…$5.00

with an exploration of how the visual snippets can be further improved to create Web page representations that we believe will markedly improve people's overall Web search and revisitation experiences.

EXISTING WEB PAGE REPRESENTATIONS

The ideal Web page representation for different tasks has been the target of much research. As noted above, some representations are best suited for finding new information, while others are designed to optimize re-finding.

Representations designed for finding new information seek to surface a page's relevant content. The most widespread search-based representations are summary *text snippets* that accompany Web search results. Text snippets are capable of encoding significant amounts of information but suffer from two problems. First, they take up a great deal of space; we estimate a three-line snippet uses around 80x800 pixels. Second, text snippets do not capture visual information about Web pages and therefore lose spatial structure and visual features that may help determine relevance.

The second most common Web page representation for search is a scaled-down bitmap, or *thumbnail*, displaying a snapshot of a particular Web page as rendered in the browser. Some Web search services such as RedZee [15] and searchme [18], as well as some browser extensions [3], display search results as collections of thumbnails or as mash-ups that present both thumbnails and snippets. While visually compelling, in practice the thumbnails are either too large (thus precluding the display of other information) or too small (thus failing to effectively surface a page's relevant content). Kaasten el al. [10] explored this tension by attempting to identify an optimal thumbnail size for Web histories. Their work provides insight into many of the factors that influence recognition when using thumbnails, Web page titles and URLs. Relevant for our research, they find that for thumbnails above 208x208 pixels, users could recognize 80% of pages.

To achieve smaller representations for search, several research efforts have proposed the use of image representations of Web pages that call out relevant text. For example, Woodruff et al. [22, 23] explored the use of enhanced thumbnails, which are thumbnails with relevant text content (e.g., matching search terms) highlighted and superimposed at a larger scale. Baudisch et al. [2] also propose surfacing relevant text on reduced versions of Web pages. Their Fishnet Web browser collapses Web pages via a fisheye viewport that compresses the text above and below the center of the screen, while surfacing relevant keywords (e.g., from an in-document search) with pop-outs. Other work by Lam and Baudisch [11] focused on enabling navigation within a particular Web page on a small device. Their strategy took advantage of the document object model of the Web page to selectively collapse sections of the page.

While the above Web page representations are intended to support the finding of new information, other representations have been explicitly designed to support revisitation. Such representations often do so by surfacing metadata about the page. For example, Cockburn and McKenzie [5] built thumbnails that show a person's interaction with the page by marking pages that are frequently visited. The Data Mountain from Robertson et al. [15] used regular scaled-down thumbnails coupled with a 2½D spatial layout surface to leverage people's ability to associate content with location. They showed an improvement over standard bookmarking mechanisms for re-finding saved pages. Similarly, PadPrints [8] used thumbnail representations to show past Web pages; they found that showing browser history helped users move through backtracking tasks more rapidly.

Figure 1. Hand-generated thumbnails created by a designer. Below each is a thumbnail of the page made to scale.

Previous studies of Web page representations have looked only at how the representation performs in a single context. In this paper we explore how different representations perform *across* contexts and how seeing a given representation in one context may affect how that or a different representation of the same page is used in another. In addition to studying well known representations like text snippets and thumbnails, we develop and test a representation intended to support both finding and recognition tasks while using as few pixels as possible. For this representation we borrow the idea of calling out important regions from a Web page, but rather than focusing on elements relating to the specific navigational task (e.g., query terms in the case of Woodruff et al. [22, 23] or visitation data in the case of Cockburn and McKenzie [5]), we emphasize three important constant components. In the next section we describe how we identified and used these components and in later sections we discuss how our representations may be further augmented to include task-specific information.

VISUAL SNIPPET DESIGN

Design Motivation

To get an idea of how best to represent Web pages independent of task, we began by considering a number of high-quality human-generated representations. We gave a graphic designer 20 Web pages and asked him to design small, 120x120 pixel thumbnails for each page. Figure 1 shows several of the thumbnails the designer created.

On inspection, we observed a consistent pattern across the hand-generated thumbnails. The majority of each designer thumbnail contained three elements:

1. Some salient text from the page (e.g., "Ketzel Levine's Talking Plants" in the left thumbnail in Figure 1).
2. A salient image, cropped so as to leave some low contrast space on which to place the text (e.g., the temple image in the center thumbnail in Figure 1).
3. A watermarked logo to brand the thumbnail (e.g., the dpreview.com logo in the right thumbnail in Figure 1).

Interestingly, these three components are similar to components typically captured by textual Web search result snippets. The salient text in the designer's thumbnail can be seen as analogous to the page's title in a search result snippet, and in many cases the salient text actually was the same as the page's title. The image the designer selected could be seen as a type of summary of the page's content, similar to a search result's text summary of the page. And the thumbnail's logo provides branding information, in the same way that the URL in a search result often does. The consistent pattern suggested it might be possible to automatically create high quality *visual snippets*.

Given these insights, we interviewed two graphic designers and one usability engineer to gather additional impressions regarding important features of a Web page for creating small-scale page representations. All three confirmed the value of emphasizing a page's logo, title, and salient image. Branding information and the page's title were viewed as central for distinguishing visually similar pages. The logo was also cited as an indicator of the trustworthiness of the source, with the page's banner or URL being a suitable substitute. A useful insight for automatically extracting these components from a page was that items "above the fold" (or visible in the browser window when a page is first loaded) were highlighted as particularly significant.

In addition to confirming our observations of the designer thumbnails, interviewees also mentioned that preserving the color or layout of the page could be valuable. For example, one interviewee said the diagrammatic composition of the different HTML elements would likely play an important role in revisitation tasks. Although the visual snippets we studied here do not take advantage of page structure to compose essential elements, we present an extension that does so, particularly for mid-sized representations.

Visual Snippet Generation

This section describes the visual snippet generator. It uses the three components identified through design analysis to build a small representation for an arbitrary Web page.

Identifying the Component Pieces

First, we must identify the components (title, salient image, and logo). We extract the title from the page's HTML, and we can use machine learning to extract the salient image and logo if they are present. Previous research suggests that logo classification can be done with 85% accuracy based on features of the image alone [21]. Additional features, such as the image's location within a Web page, size, name, link structure, and surrounding text can improve the accuracy of logo detection [13]. For large Web sites, looking at many pages within the same site may be useful, as the logo is often consistent across pages. Maekawa et al. [13] found that the identification of content images can be done with even greater accuracy than logo detection.

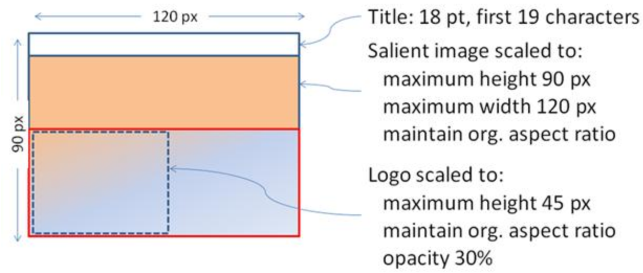


Figure 2. The visual snippet template for automatically generating the snippet given a salient image, logo, and title.

In our experiments, we treat logo and salient image extractions as black boxes that we initially implemented in "Wizard of Oz" style. Two authors viewed the Web pages used in our experiments and quickly identified a logo and a salient image for each by hand, focusing on above-the-fold content as suggested by our design analysis. Hu and Bagga [9] found that manual image categorization can have a high error rate (19.1%), so it is likely that the number of errors introduced through manual classification corresponds to what would be found through automatic classification, although the errors may be somewhat different in quality. Later we present a fully automated implementation that successfully mimics the manual extraction in quality.

Compiling the Component Pieces

Following extraction, we automatically compile the component pieces into a visual snippet. Figure 2 shows the template we used to automatically generate a visual snippet given a salient image, logo, and title. Figure 3 shows three Web pages and the visual snippets we derived from them.

The visual snippet generation process involves four steps:

1. Cropping and scaling the salient image. The image is cropped manually along one dimension to an aspect ratio of 4x3 and scaled to 120x90. If no salient image is identified, a snapshot of the page is used instead, appropriately scaled.

2. Scaling the logo. The logo is scaled to fit within a 120x45 rectangle while preserving its original aspect ratio. The logo's scale is chosen so that it either fills half of the height or the full width of the visual snippet. If no logo is available, it is omitted.

3. Cropping the title. Kaasten et al. [10] found 30-39 letters to be necessary to provide medium-quality

Figure 3. Several example Web pages (bottom) and the visual snippets built from those pages (above).

recognition of a specific Web page. Strings of this length are possible in text snippets but are infeasible for smaller representations. Because the leftmost 15-20 letters of a page's title [10] yield reasonable recognition of the page's site, we use the first 19 characters of the title. If no title is available, it is omitted from the final snippet representation.

4. Composing the pieces. The three processed pieces are then composed as shown in Figure 2. The logo is made semi-transparent and overlaid on top of the salient image, and the salient text is placed above both images on a white background for readability. We place the logo in the lower left-hand corner of the visual snippet because that is where the URL appears in a typical text snippet. We hypothesize such a placement is consistent with existing expectations.

Note that all component processing is done without consideration of how the pieces will compose. It is likely that allowing for interactions will lead to better visual snippets. For example, the salient text is currently placed above the image for readability and consistency, but it would be simple to automatically identify low contrast areas in the salient image on which to place the text instead, much as the designer did in his original thumbnail creations (see Figure 1). Similarly, it may be beneficial to crop the image so as to leave a low contrast area in the lower left hand corner for the logo. Many extracted logos are not rendered on transparent backgrounds. Because logos with transparent backgrounds appear to compose better, it may be valuable to try to identify the background and make it transparent.

STUDY OF SEARCH AND RE-FINDING

To explore how well visual snippets support search and revisitation tasks, we conducted a study to compare how participants used different representations types to find and re-find content. Our goal was, first, to understand how different renderings of snippets support *finding* tasks; second, to explore how different renderings support *re-finding* tasks; and, third, to investigate whether consistency in representation makes any difference across tasks.

Snippet Representation

The three representations explored in the study were:

1. **Text snippets** (555x78). The title, a one line summary, and URL for a Web page were captured from a popular search engine. The text display was generic and not tailored to a particular query (e.g., there was no hit-highlighting).

2. **Visual snippets** (120x90). Created as described in the previous section. Note that visual snippets are less than a quarter of the size of text snippets.

3. **Thumbnails** (120x90). For comparison, we also created thumbnails of the page that were the same size as the visual snippets.

Figure 4 shows some examples. Note that text snippets are significantly larger (roughly four times the pixel area) than the other representations, and this is true even in the absence of the additional white space required for effectively rendering a list of text results.

Study Design

Participants completed a two phase study. The first phase involved searching for new information among a set of Web pages, and the second involved revisiting the information found during the first.

Phase I: Search

In Phase I of the study, participants were asked to perform 12 search tasks. For each task they were given a task description and a set of 20 search results associated with the task. Each participant completed four of the 12 tasks with each type of Web page representation so that we could perform a within-subjects comparison of representation. Web search performance is associated with very large inter-person variability; we hoped to minimize this by comparing performance between representations within a single user.

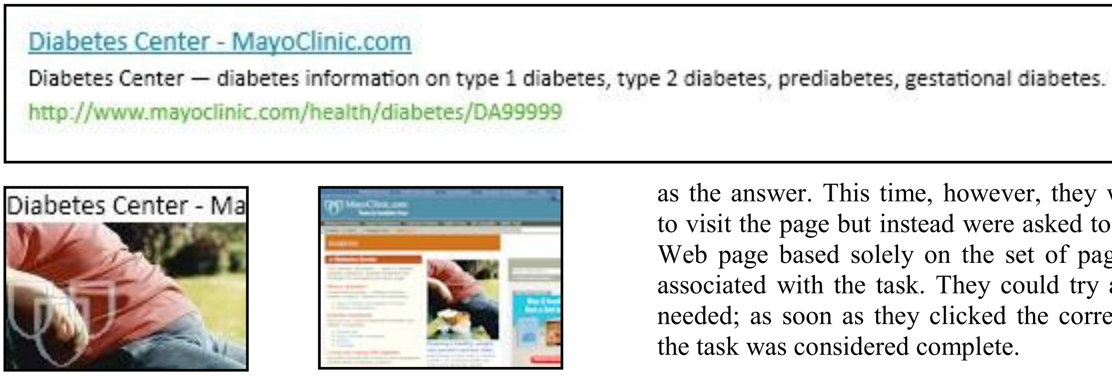


Figure 4. An example of the three snippet types explored in our study for a single page: text snippets (top), visual snippets (bottom left), and thumbnails (bottom right).

Because all users saw all three representations, we could collect more reliable qualitative preference measures by asking participants to provide relative preferences. The type of Web page representation for each task was counterbalanced between participants and the order of presentation was pseudo-randomized to avoid order effects.

In each search task, participants were asked to find some information that was guaranteed to be available on at least one of the Web pages in each result set. Broadly, there were three main types of tasks (four of each type): homepage finding (e.g., "*On Dave Barry's blog, find his presidential campaign icon.*"); shopping ("*Where can you buy a 8GB iPod Nano for under $230?*"); and medical ("*About what percent of school age children are affected by ADHD?*). The answers to the homepage finding tasks were on only one of the twenty results. The answers to the medical and shopping tasks could be found on two to five of the results.

During the search phase, participants could click on a Web page representation to see the full Web page and click back to return to the result list. When participants found a result containing what they determined to be an answer to the question, they were instructed to click on the answer within the target Web page. The selected page was recorded for use in Phase II, and that task was considered complete. We did not require participants to find a "correct" page but rather allowed them to decide for themselves when their information need was satisfied.

At the end of Phase I, participants filled out a survey including demographic information as well as impressions of their experience in performing the task.

Phase II: Revisitation

In addition to exploring how the different representations support search, Phase I also served as a priming phase for a follow up study of how people recognize previously viewed pages. One day after participants completed Phase I, they were asked to complete a second phase of the experiment.

In Phase II, participants were given the same task descriptions they saw during Phase I and were asked to identify the Web page that they had selected the day before as the answer. This time, however, they were not required to visit the page but instead were asked to re-find the target Web page based solely on the set of page representations associated with the task. They could try as many times as needed; as soon as they clicked the correct representation, the task was considered complete.

In Phase II, we were interested in knowing whether the type of representation of search results in Phase I would affect the recall of those same pages the next day. For example, if a participant used thumbnails during the search task, would that participant be better able to remember the correct pages when using thumbnails during the revisitation task?

We showed the same set of pages in Phase II as in Phase I. However, participants saw only a single representation type (text snippet, visual snippet, or thumbnail) in Phase II; representation was a between-subjects variable. By requiring each individual to interact with only a single representation during Phase II, we were able to assess the effect of the representation type on the recall of Web pages that participants had seen the day before as well as look at the effect of congruency of the representation.

Participants

Participants were recruited from across the entire employee population of a large software company. Phase I was completed by 276 people; of those, 197 went on to complete Phase II. Participants came from a range of job roles, including executive, design, engineering, and sales. They ranged in age from 18 to 65 years old (more than half reported they were 26-35), and 86% were male. All were heavy users of Web search, with most reporting that they searched the Web several times a day or more.

RESULTS

We explored the data we collected to understand how different renderings of snippets supported search (Phase I) and revisitation (Phase II) tasks as well as to investigate if consistency in representation across tasks was important.

Search (Phase I)

For the search phase we were interested in two quantitative measures of performance: task completion time and the number of page views for each task. We performed two 3 (Representation) x 4 (Repetition) within-subjects repeated measures ANOVAs (RMANOVA), looking first at task completion time and next at number of page clicks. We also explored several qualitative measures of representation quality, including preference judgments and free form comments.

Task Completion Time

For task completion time, there was a main effect only for Repetition ($F(3,579)=3.515$, $p<.015$). Unsurprisingly, as the

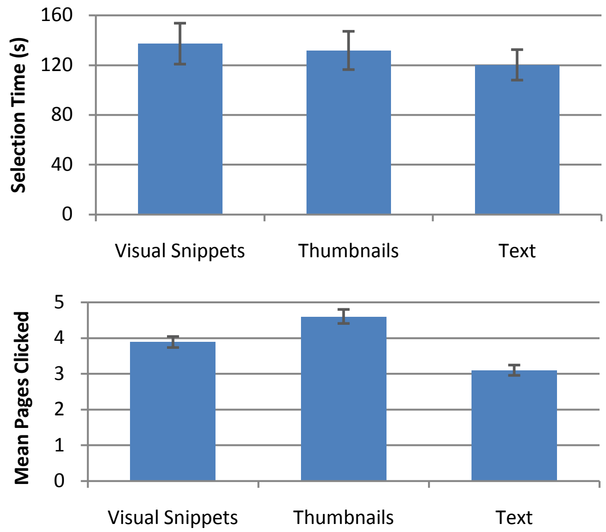


Figure 5. Mean selection times and pages clicked (±SEM) for each Web page representation in the Search (Phase I) task.

experiment progressed, participants got faster at searching. Completion times averaged 166 seconds to complete the first task and decreased to 100 seconds to complete the last.

There was no effect on task time for Representation and no significant interaction. As Figure 5 shows, while the average time to complete the task was smallest for text snippets, this was not significantly different from either visual snippets or page thumbnails. This suggests our participants were able to find new information quickly, independent of how the pages were represented.

Number of Page Clicks

Even though there was little difference in selection time, it does appear that people explored the results in different ways depending on how the results were represented. When we analyzed the number of search results clicked prior to completing the search task, we found significant main effects for both Representation ($F(2,390)=26.2$; $p<.001$) and Repetition ($F(3,585)=5.51$; $p<.001$) with no interaction.

As was observed for completion time, as participants performed more searches, they also got a bit more efficient at searching: they looked at an average of about 4 pages initially, and this dropped to 3.3 pages by the last task.

More interestingly, participants clicked on the fewest number of results when searching using text snippets, and the largest number when using thumbnail representations, with visual snippets falling in between (see Figure 5). Post-hoc pair-wise comparisons (Bonferonni-adjusted) show significant differences between all representations.

Qualitative Measures

We also looked at the participants' subjective experience with the three different representations. For the search tasks, visual snippets and text snippets were judged to be equally easy to use and well-liked, and both scored significantly better than thumbnails.

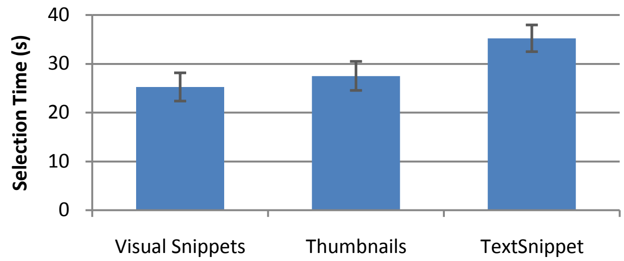


Figure 6. Mean selection times (±SEM) for each type of Web page representation in the Revisitation (Phase II) task.

Participants judged ease of use on a seven-point Likert scale, with 1 being very hard and 7 being very easy; text snippets received a mean rating of 3.96, visual snippets 3.97, and thumbnails 3.24. Because Likert scale responses are not normally distributed, standard t-test comparisons cannot be used. Pairwise comparisons between ranks using the Mann-Whitney U test showed significant differences between text snippets and thumbnails ($z=5.07$, $p<0.001$) and between visual snippets and thumbnails ($z=5.46$, $p<0.001$) but no difference in ease of use between text and visual snippets. Similarly, when participants were asked if they liked a particular representation on a seven-point scale (1=no, 7=yes), text received a mean rating of 4.51, visual snippets 4.28, and thumbnails 3.75. Again, text and visual snippets were each liked significantly more than thumbnails ($z=5.80$, $p<0.001$ and $z=4.28$, $p<0.001$ respectively) but did not differ significantly from each other.

We also explored the comments participants made about their experiences with the three different representations. A number of people mentioned using branding information to find what they were looking for, referring specifically to the URL in the text snippet or the logo in the visual snippets as a source of that information. As suggested by the designers during design analysis, these two components appear to have served similar functions. For example, one participant said, "When I see a Web site's name in a visual snippet, I get the same information from the URL and I generally weight that heavily." Only one participant mentioned using the page layout in the thumbnail representation to identify brand. Visual representations of pages from unknown domains may have been less valuable, as suggested by a participant who reported, "The usefulness of thumbnailing pages that I've never been to is limited."

A number of subjects mentioned that the value of the different representations varied by task, with the visual snippets being particularly useful for shopping tasks. This may be because people prefer to shop at trusted sites and are familiar with the shopping site logos highlighted in the visual snippets. As one participant said, "The nice thing with the [visual snippets] was when I was looking for the cheap price I knew Amazon was usually the cheapest so I just had to look for the Amazon logo. When looking for information the images were not helpful."

A common complaint with the thumbnail representations was that the size was too small (e.g., "Thumbnails were

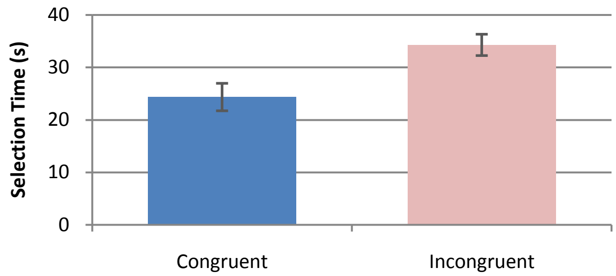


Figure 7. Mean selection times (±SEM) for Web page representations in the Revisitation task that were (in)congruent with those seen in the Search task (Phase I).

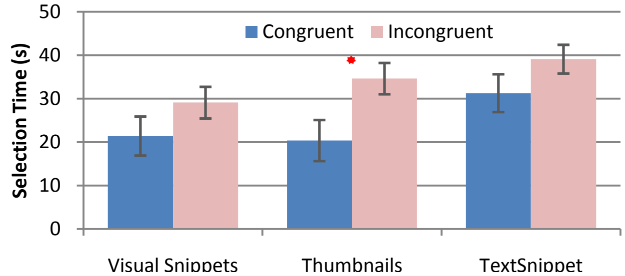


Figure 8. Mean selection times (±SEM) for each type of Web page representation in the Revisitation task broken down by congruency with the Search task (Phase I).

generally too small to be helpful"). These comments were not surprising given that we know from previous research [10] that thumbnails of the size used in the study are too small to support recognition even at the site level. Many subjects suggested combining visual and text representations either by creating a single composite or through the use of hover.

Discussion for the Search Task

Overall, we observed no significant difference in time to task completion for any representation. However, visual snippets required fewer clicks to complete the task than thumbnails, and visual snippets were subjectively preferred over thumbnails. We believe consideration of all of these observations is necessary to understand how the different representations were employed for search.

It is interesting to observe that participants clicked more often on thumbnail representations than text and visual snippets, while taking about the same amount of time to complete the task overall. Timing differences can be difficult to assess in tasks like those studied, and the number of clicks may be a reasonable proxy for effort involved in the task, especially for systems like the Web with significant latency following clicks. The pages in our test loaded almost instantaneously. In systems with more latency for loading Web pages, the increased number of clicks for thumbnails could translate into longer overall task time due to waiting for page loads. Text and visual snippets would presumably be less affected by this.

One way to understand the observed difference is that participants spent more time looking at the text and visual snippet representations and deciding what to click than they did with the thumbnail representations. However, the different processing times allowed participants to find what they were looking for just as quickly because they used different click strategies.

Revisitation (Phase II)

During the second phase participants were asked to re-find the correct results that they had identified during the initial search phase the day before. In general, the task completion times were considerably faster for revisitation than search, suggesting participants did indeed use their memory of the results from their initial search to help them revisit the correct result. On average, participants completed each revisitation task a full minute and a half faster than they did the search task (29.3 seconds v. 129.5 seconds).

Phase II was largely a between-subjects design. Participants interacted with the same representation type throughout the second phase. Because they interacted with all three types during the initial phase, this meant that for one-third of the tasks in Phase II the representation type used was congruent with the representation type used in Phase I, and for two-thirds of the tasks the representation type was different.

We performed a between-subjects 3 (Representation) x 2 (Congruence) ANOVA. There were significant main effects for both Representation ($F(5,1526)=3.39$; $p<.005$) and Congruence ($F(1,1526)=313.60$; $p<.001$). There was no significant interaction.

Effect of Representation on Completion Time

Figure 6 displays the mean amount of time it took to re-find the correct result found during Phase I, broken down by representation type. Visual snippets were the fastest for re-finding, followed by thumbnails. Text snippets were the slowest. The trend suggests visual representations of previously viewed pages may support faster revisitation. Follow up pair-wise comparisons showed that only the difference between text and visual snippets was significant.

Effect of Congruence on Completion time

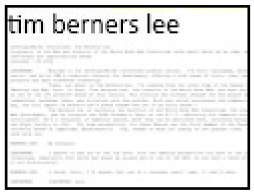
We also looked at the effect of congruency on revisitation time. When the representation type was congruent across both the search and revisitation phases, we saw a significant decrease in task completion time compared to when the representations were different (see Figure 7). Previous interactions with a given type of representation appear to improve performance for re-finding later; familiarity helps.

Deeper analysis of the data shows that this effect is stronger for thumbnails than for either visual or text snippets. Figure 8 shows the difference in task completion time for each representation type broken down by congruency. There was a significant difference for congruency for thumbnail representations ($t(475)=2.54$; $p<.01$), but the differences were not significant for text and visual snippets.

Discussion of the Revisitation Task

During the second phase we observed a tendency for visual snippets to be fastest, significantly better than text snippets, and that thumbnails were significantly less likely to be

(No salient image)



(No salient image or logo)

Figure 9. Examples where the visual snippet creation fails because of failure to extract a salient image or logo.

recognized if they were not seen during the initial priming phase. Interacting with an actual Web page was not enough to recognize or use a thumbnail of the page for re-finding. In contrast, text and visual snippets seem to have captured some of what the participants internalized about the pages during their initial interactions, making them better representations for revisiting previously seen pages. The ability of the visual snippets to perform better on incongruent tasks is important because in many cases where Web page representations are useful (e.g., histories), we cannot assume a user will have had prior exposure to the exact same representation. In real-world situations, the expectation of congruency across tasks is likely to be hard to enforce.

Overall, we found that for finding tasks text snippets were easy to use, well liked, and required relatively few clicks to find the information target. In contrast, for re-finding tasks the visual representations were the fastest. Visual snippets appeared to capture the best of text and thumbnails; they were as easy to use and well liked as text snippets for finding and as fast as thumbnails for re-finding without requiring congruency.

IMPROVING THE VISUAL SNIPPETS

Encouraged by these results, we implemented a fully automated visual snippet generator. This allowed us to confirm that the extraction of important components from a Web page could indeed be done automatically, and to explore several avenues for improving the generation algorithm. In this section, we first show that automatically generated visual snippets were as high quality as the ones created via manual component extraction. Then we discuss some problems with the design as it stands and present improvements to the system that correct for these problems.

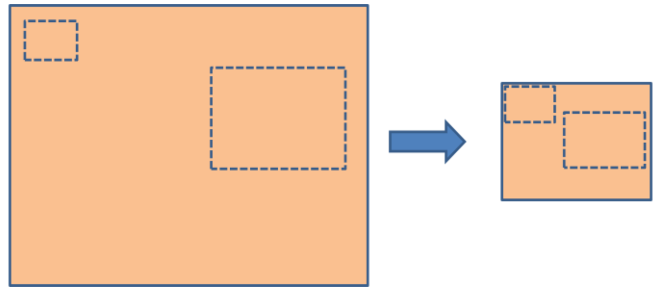


Figure 10. Images and text are scaled at a differential rate from the rest of the page.

Confirming Automatic Generation

As described earlier, in our study we manually extracted the logo and salient image from each Web page in our collection and then automatically composed the pieces to create the visual snippet. Given the success of visual snippets described above, we implemented fully automatic extractors. The salient image was simply the largest image on the page, and logo was selected using machine learning over several features, including the image's name, aspect ratio, surrounding link structure, and location.

To confirm that the fully automated visual snippets are of similar quality to the partially automated visual snippets, we conducted a study in which we asked people to tell us which representation they preferred. In the study, 128 participants viewed an average of six Web pages each. The pages were selected from the set used in the initial study. After five seconds, participants were presented with the two visual snippets and asked to select the representation that better matched the page they just saw. Of the 723 comparisons, we found that people preferred the snippets used in our study 362 times, and the fully automated visual snippets 361 times. There was no statistical difference between the two. However, because the automatic generation was not tested in our experiment of representation use, there may be observable differences in how they are used compared with manual generation.

Given that it appears we have identified a successful way to generate visual snippets in a fully automated fashion, we can now explore the problems with the existing design and easily implement improvements.

Problems with Existing Visual Snippet Design

Problem 1: Snippets Visually Distinct from Parent Page

One problem is that while visual snippets convey an overall impression of the Web pages they represent, they can be quite visually distinct from their parent pages. In our design analysis, several designers suggested that a correlation between the page layout and page color would be useful for revisitation. We also hypothesize that representations that are similar to the target may help users better orient themselves within the target when they choose to visit it.

Problem 2: Visual Snippets Do Not Scale Well

Another problem is that visual snippets do not appear to scale well. Although their small size is beneficial for many

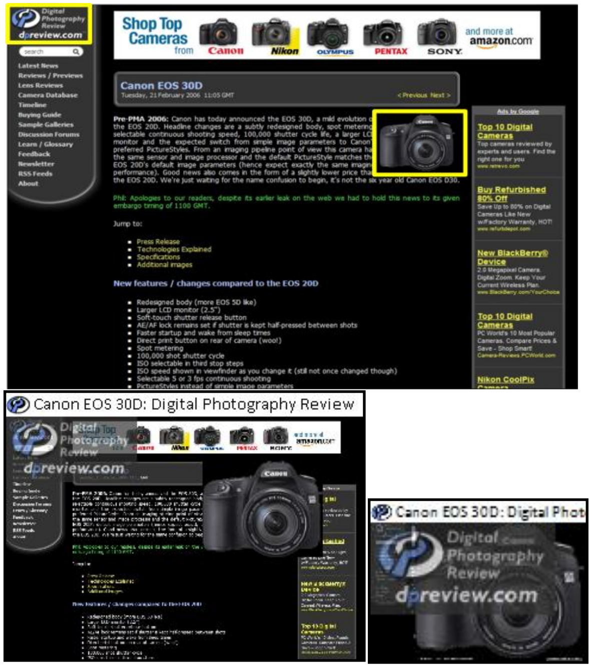


Figure 11. An improved visual snippet shown at different scales. The location of the image, logo, and text are preserved.

Figure 12. Another example of an improved visual snippet. Salient text is highlighted differently at different scales.

tasks, increasing the size does not provide any additional information. In contrast, thumbnails are easier to recognize as they get larger [10]. Optimally, we would provide as much information as possible within the space available while still providing important semantic information.

Problem 3: Bad Defaults when Extraction Fails

We also observed that the visual snippets were not very effective (and look like regular thumbnails) when a salient image or logo was not available. Figure 9 shows two examples of this failure. Here we use the full thumbnail in such cases, but we believe better defaults would help.

Improved Visual Snippet Generation

To create a visual snippet that is better connected visually to its parent page and that scales better to different sizes, rather than extracting the salient components from a page and using them to create a new representation by composing them, our improved visual snippet generator resizes the selected images and text and overlays them directly onto a scaled version of the Web page. As illustrated in Figure 10, the key idea is to scale the selected images and text differently from the overall Web page. The exact placement of the salient aspects corresponds to their original position on the page, offset as necessary to prevent them from overflowing the borders of the resized page.

An example of this improved visual snippet design can be seen in Figure 11. The top of the figure shows the original Web page with the page's logo and salient image highlighted in yellow. When the page is scaled to the size of the original visual snippets, as shown in the lower right-hand corner of Figure 11, it looks very similar. However, as it is scaled to larger dimensions, such as is shown in the lower left-hand corner of Figure 11, additional page-level information can be shown.

With this improved design, it is possible to highlight additional aspects of a page as the page is represented at different sizes. For example, Figure 12 shows a Web page where the dominant image on the page, the logo, and an article title are identified as salient. The component pieces can be scaled differently so that some salient pieces are emphasized when there is enough room, while the snippet still reduces to the original design at small sizes.

This design provides users with some orientation within the target Web page should they click through and to enables semantically meaningful thumbnails to be represented at different sizes. Further, by identifying additional page elements, we can create visual snippets that fail gracefully when a salient image or logo is not identified.

These additional page elements could also enable us to create thumbnails that are consistent across navigational tasks at small sizes, but tailored to best support the task when there is room. For example, we could create query specific representations by selecting the query text that appears on a page, as was done by Woodruff et al. [22, 23].

The improved visual snippet generator shares some aspects with the one proposed by Woodruff et al. [22, 23], but our emphasis is on creating a visual summary as opposed to enhancing the presence or absence of a particular textual term. As a result, our representations are context independent. The importance we observed of congruency across tasks suggests a consistent representation across many different uses may be valuable for users. In future work it will be interesting to explore how our technique can be combined with Woodruff et al.'s to create representations that appear consistent across many different task types (e.g., that consistently highlight the title, logo, and salient image) but also call out task-relevant information when appropriate (e.g., query terms for a search task).

CONCLUSION AND FUTURE WORK

In this paper, we looked at how different representations of Web pages affected people's ability to recognize new relevant Web content and return to previously viewed Web pages. We found that our novel visual snippets support search while being significantly smaller than text snippets, and are particularly valuable for revisitation. We believe our findings can be used to significantly improve people's search and browse experiences.

Small representations like the visual snippets allow a greater number of results to be viewed at once. This is particularly important on mobile devices, where screen real estate is limited, but also important for history functionality where a large number of pages must be viewed together. Further, small visual snippets could be used to complement text snippets in search result pages. With only a small reduction in the amount of text, a hybrid snippet could occupy the same amount of space as current text snippets.

We believe it may be possible to construct even smaller visual snippets that are consistent with the snippets we have explored using just the logo and image. These micro-representations could be used in a bookmark or history list the way favicons currently are.

One area alluded to in our discussion of the improved visual snippets that we plan to explore further is the transition between a Web page's representation and the full page. Representations can serve an important role not just in identifying a target page, but also in orienting a person within the target. This can be done by making the representation consistent with the target or by animating a transition between the representation and the target, both of which are supported by the improved visual snippets. Understanding the value of these features is particularly interesting as complex animation on the Web becomes more technologically feasible.

ACKNOWLEDGMENTS

Our thanks to Joey Pitt and the designers who worked with us, and to Merrie Morris, Susan Dumais, Mary Czerwinski, Dan Robbins, Greg Smith, Desney Tan, and Dan Morris.

REFERENCES

1. Adar, E., Teevan, J., and Dumais, S. T. (2008). Large scale analysis of Web revisitation patterns. In *Proceedings of CHI '08,* pp. 1197-1206.
2. Baudisch, P., Lee, B. and Hanna, L. (2004). Fishnet, a fisheye Web browser with search term popouts: A comparative evaluation with overview and linear view. In *Proceedings of AVI '04*, pp. 133-140.
3. Better Search Firefox Extension. From https://addons.mozilla.org/en-US/firefox/addon/211
4. Bruce, H., Jones, W. and Dumais, S. (2004). Keeping and re-finding information on the Web: What do people do and what do they need? In *Proceedings of ASIST '04*.
5. Cockburn, A. and McKenzie, B. (2001). What do Web users do? An empirical analysis of Web use. *International Journal of Human-Computer Studies*, 54(6): 903-922.
6. Cockburn, A. and Greenberg, S. (2000). Issues of page representation and organisation in Web browser-revisitation tools. *Australian Journal of Information Systems*, 7(2):120-127.
7. Herder, E. (2005). Characterizations of user Web revisit behavior. In *Proceedings of Workshop on Adaptivity and User Modeling in Interactive Systems*.
8. Hightower, R., Ring, L., Helfman, J., Bederson, B., Hollan, J. (1998). Graphical Multiscale Web Histories: A Study of PadPrints. *Hypertext 1998*, 58-65.
9. Hu, J. and Bagga, A. (2004). Categorizing images in Web documents. *IEEE MultiMedia*, 11(1): 22-30.
10. Kaasten, S., Greenberg, S. and Edwards, C. (2002). How people recognize previously seen Web pages from titles, URLs and thumbnails. In *Proceedings of HCI '02*, pp. 247-265.
11. Lam, H. and Baudisch, P. (2005). Summary Thumbnails: Readable overviews for small screen Web browsers. In *Proceedings of CHI '05*, pp. 681-690.
12. Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19(1): 458-465.
13. Maekawa, T., Hara, T. and Nishio, S. (2006). Image classification for mobile Web browsing. In *Proceedings of WWW 2006*, 43-52
14. Obendorf, H., Weinreich, W., Herder, E. and Mayer, M. (2007). Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. In *Proceedings of CHI '07*, pp. 597-606.
15. Ranie, L. and Shermak, J. (2005). Pew Internet and American Life Project: Data memo on search engine use. Retrieved from http://pewinternet.org/pdfs/PIP_SearchData_1105.pdf.
16. RedZee. http://www.redzee.com
17. Robertson, G., Czerwinski M., Larson, K., Robbins, D. Thiel, D. van Dantzich, M. (1998) Data mountain: using spatial memory for document management. In *Proceedings of UIST '98*, pp. 153-162.
18. Searchme. http://www.searchme.com
19. Tauscher, L. and Greenberg, S. (1997). How people revisit Web pages: Empirical findings and implications for the design of history systems. *Int. J. of Human-Computer Studies*, 47(1):97-137.
20. Teevan, J., Adar, E., Jones, R. and Potts, M. A. (2007). Information re-retrieval: Repeat queries in Yahoo's logs. In *Proceedings of SIGIR '07*, pp. 151-158.
21. Voutsakis, E., Petrakis, E. G. M. and Milios, E. (2005). Weighted link analysis for logo and trademark image retrieval on the Web. In *Proceedings of WI '05*, 581-585.
22. Woodruff, A., Faulring, A., Rosenhotlz, R., Morrison, J. and Pirolli, P. (2001). Using thumbnails to search the Web. In *Proceedings of CHI '01*, pp. 198-205.
23. Woodruff, A., Rosenholtz, R., Morrison, J., Faulring, A. and Pirolli, P. (2002). A comparison on the use of text summaries, plain thumbnails, and enhanced thumbnails for Web search tasks. *JASIST*, 53(2), 172-185.