Wish I Hadn't Clicked That: Context Based Icons for Mobile Web Navigation and Directed Search Tasks

Vidya Setlur
Nokia Research Center
955 Page Mill Rd.
Palo Alto, CA 94304, USA

Samuel Rossoff
University of Victoria
Victoria, BC V8W 3P6
Canada

Bruce Gooch
University of Victoria
Victoria, BC V8W 3P6
Canada

ABSTRACT
Typical web navigation techniques tend to support undirected web browsing, a depth-first search of information pages. This search strategy often results in the unintentional behavior of 'web surfing', where a user starts in search of information, but is sidetracked by tangential links. A mobile user in particular, would prefer to extract the desired information quickly and with minimal mental effort. In this paper, we introduce 'SemantiLynx' to visually augment hyperlinks on web pages for better supporting the task of directed searches on small-screen ubiquitous platforms. Our algorithm comprises four parts: establishing the context of information related to a hyperlink, retrieving relevant imagery based on this context, applying image simplification, and finally compositing a visual icon for the given hyperlink. We evaluated our system by conducting user studies for directed web search tasks and comparing the results to using textual snippets and webpage thumbnails.

Author Keywords
web, mobile device, directed search, icons, hyperlinks

ACM Classification Keywords
H5.2 Information interfaces and presentation: User Interfaces

General Terms
Human Factors, Design, Algorithms

INTRODUCTION AND MOTIVATION
The fastest growing community of Internet users is made up of people who use various ubiquitous devices to access the Internet. Such web content is becoming more important for searching, sharing, expressing, and exchanging information on devices such as cell phones, hand-held PCs, PDAs, home-networked media appliances, and informational displays in automobiles and helmets. As the amount of information available on these small-screen, 'on-the-go' types of devices continues to grow, it is essential to prevent users from wading through a morass of irrelevant content to find a single piece of relevant information. Most people use the mobile Internet for *directed* searches, where the goal is to find information about a predetermined topic of interest [11]. Fact-finding, where the user searches for a particular piece of information, and subject-based exploration, where the user seeks basic understanding of a concept and related terms, are two examples of this kind of task. Typical web navigation techniques tend to support undirected searching, which often results in the unintentional behavior of 'web surfing.' This occurs when a user starts in search of some information, but is sidetracked by tangential links.

Despite the modifications and features introduced by web browsers, the primary method of web navigation is hyperlinks [3]. Current web navigation techniques are suited for depth-first traversal, where a user first selects a link on the current page, which in turn loads a new page. This process repeats until the user finds the needed information or the current search path is abandoned returning the user to the initial search. This process can be referred to as a hub-and-spoke strategy [3], where the user follows a path from the start page or *hub*, ending at a *spoke*, and then returning to the hub to start another spoke. This backtracking often becomes difficult as the context of the search is forgotten. Conkin [6] identifies this loss of context as a cognitive overload due to the fact that both the search task and navigation trail must be held in short term memory. This causes the web surfing behavior, as users inadvertently create secondary hubs while they traverse the search space. These secondary hubs may drift away from the original search topic, defeating the purpose of a directed search and leaving the user frustrated.

The three goals of SemantiLynx are:

- **Simplifying the decision making process during a directed search on a mobile device:** Decision making in a web based directed search is the cognitive process leading to the selection of a given web page to view. Navigating the web should be easier when the user is not forced to remember the context of the mobile search space [4]. By placing images representing contextual information, in the icons, we hope to leverage visual search to help users find a desired web page significantly faster.
- **Increasing the speed of such directed search tasks:** By reducing the number of user actions such as unintentional

IUI'11, February 13–16, 2011, Palo Alto, California, USA.

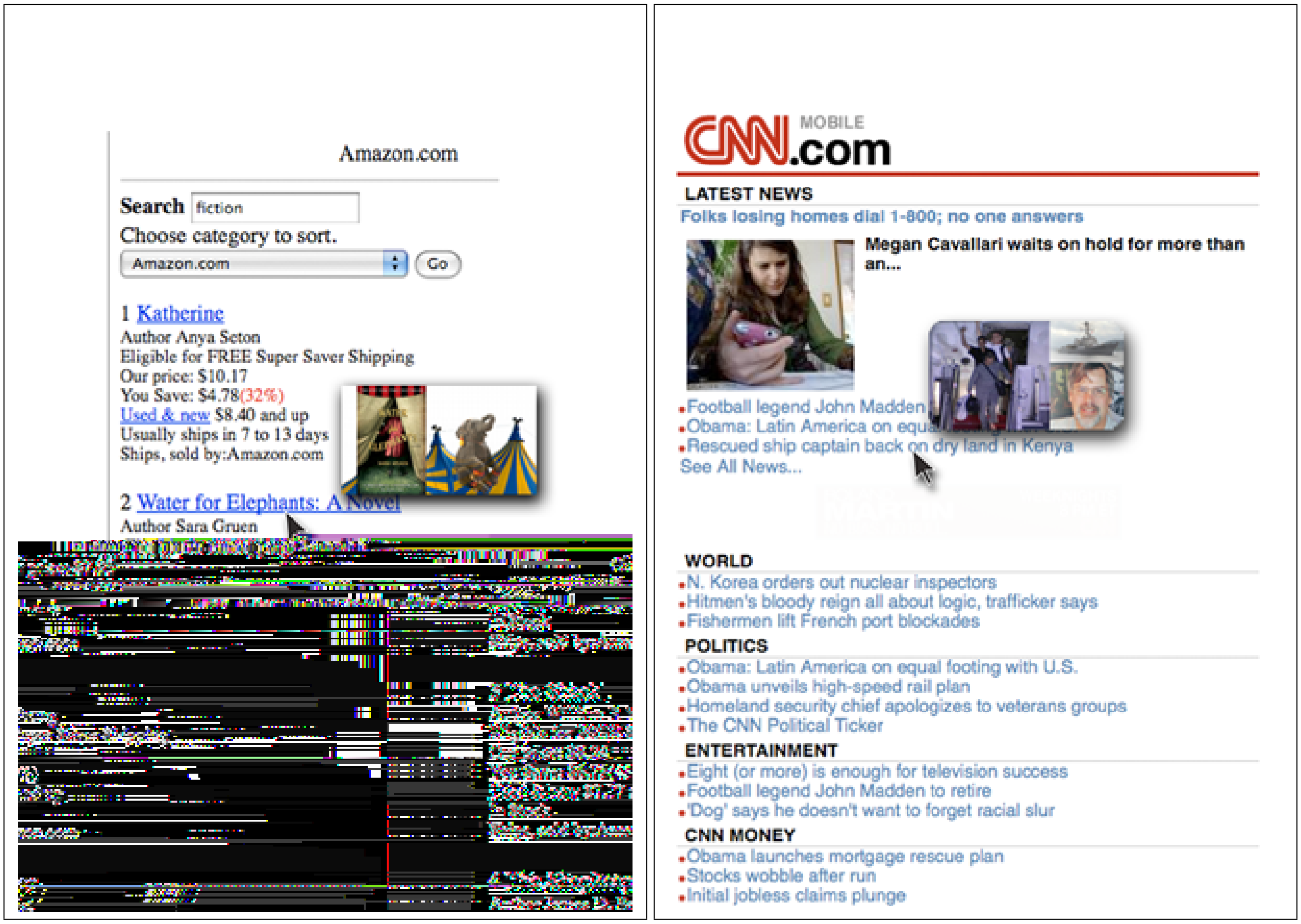

Figure 1. Examples of hyperlinks on two mobile web pages augmented with mouse-over SemantiLynx icons. Left: A SemantiLynx icon generated for the book 'Water for Elephants' composited with an image of the book cover obtained from the target web page, and augmented with additional images for keywords in the web page - 'circus elephant', 'circus,' obtained from a stock image database. Right: A SemantiLynx icon generated for the story on a rescued ship captain. In this case, the target web page has sufficient images to completely composite the icon.

clicking of hyperlinks and revisits of web pages, we hope to improve directed search speed.

- **Improving the memorability of web pages visited during a directed search task:** As users often revisit pages after a preliminary directed search, improved retention of these web pages may yield improvements in finding these web pages quicker.

Research Contribution

The main contribution of this work is an algorithm for automatically generating smart thumbnails of a target webpages using mere visual resources. Such thumbnails are optimized in terms of real-estate and layout, making them suitable for mobile browsing.

RELATED WORK

Standard web design practices that aid in user's navigation include bread crumbs and sitemaps [12]. While these design metaphors provide a user with a sense of where they are within a website, such metaphors do not provide a sense of where the user may want to go. Several websites (e.g. http://www.ask.com) have binoculars preview icons preceding some of the URLs on the search space. However, the preview is only a scaled down version of the target page, containing significant amount of unnecessary information further obfuscating the information.

Previous work has used a number of graphics techniques for various types of content search tasks. Woodruff *et al.* present designs for enhanced thumbnails that can be used to search the web [23]. Modifications include changing fonts for readability and overlaying keywords onto the reduced web page. While these thumbnails may benefit desktop users, the small size of mobile screens require a more efficient way of using the screen real-estate. 'WildThumb' helps users switch between opened Web pages by showing thumbnails of the pages in the extra space around the currently focused page [15]. 'Visual Snippets' explores how different representations are used in a variety of web related contexts and presents a representation that supports help users with revisitation of web pages [21]. While this work uses simple visual representations of a salient image, logo, and title to provide cues to a user, our algorithm utilizes additional semantics of the web pages based on text, images and layout to generate more concise icons with multiple images composited together based on relevance. These icons tend to further op-

Figure 2. Overview of SemantiLynx icon generation. The system takes as input, a web page that contains hyperlinks. Dominant images are extracted from the target page that the hyperlink points to. If more images are required, the system retrieves additional images from a stock database using the keywords of the web page as a query. The images are then composited based on the relevance weights of the individual images to form an icon.

timize their real-estate on the mobile screen with richer visual contexts for directed search tasks.

In the mobile domain, there have been attempts to adapt content for mobile web browsing [1, 22]. Summary thumbnails help users identify viewed material and distinguish between visually similar areas [13]. While these papers deal with the adaptation of web content prior to load, they do not address the greater issue of directed search tasks, and may still lead to the cognitive overload mentioned earlier.

SEMANTILYNX SYSTEM

SemantiLynx is a system that automatically generates icons, which reveal the information content of a web page with respect to a search context as shown in Figure 1. The icons have a resolution of 140×80 when generated, and we let the mobile browser proportionally scale the image in accordance to the given mobile real-estate in order to support a a variety of target display sizes. The system comprises a backend server that handles the icon generation and subsequent augmentation to a given web page, and a mobile phone (Figure 3) that has XHTML browser support and can view the augmented web page. The icons are displayed with the hyperlink using a mouse-over browser event that gets triggered based upon the input modality of the mobile device used. For example, on a touch-screen device, a mouse-over event is usually triggered by a light tap on the link, whereas on a non-touch-screen device, a mouse-over event is typically triggered when the hyperlink is brought to focus by using the navigation key. In addition, the SemantiLynx icons can be displayed by clicking on a 'preview' link next to the url.

Algorithm Overview

The algorithm for generating the SemantiLynx icons has the following steps and is illustrated in Figure 2:

1. *Establishing Hyperlink Context:* The semantic meaning of what the hyperlink represents is estimated by using a term weighting metric for parsing the content of the target web page that the hyperlink points to.
2. *Image Retrieval:* The metric is then used to extract dominant images and salient keywords in the target web page. The keywords can be used as a query to search for additional images in a stock image database, indexed by keywords.
3. *Saliency Based Region Extraction:* The images are simplified by segmenting them and computing an importance value for each segmented region.
4. *Image Composition:* These simplified images are composited together based on some simple design heuristics into the final SemantiLynx icons.

We now explain each step in greater detail.

ESTABLISHING HYPERLINK CONTEXT

A meaningful icon uses imagery that is semantically connected to its inherent content or purpose [19]. To find imagery that makes this connection, we need to establish a context for the hyperlink. We use a three-step process to determine the context for the hyperlink. The first step involves identifying dominant images present in the target web page. The second step involves determining the textual keywords that are salient to the target web page. The third step uses the surrounding text extracted from the dominant images, and the textual keywords to determine the relevance weights of these prominent images to the overall context of the web page. These relevance weights are used to determine the images and their associated compositing rules to generate the SemantiLynx icons.

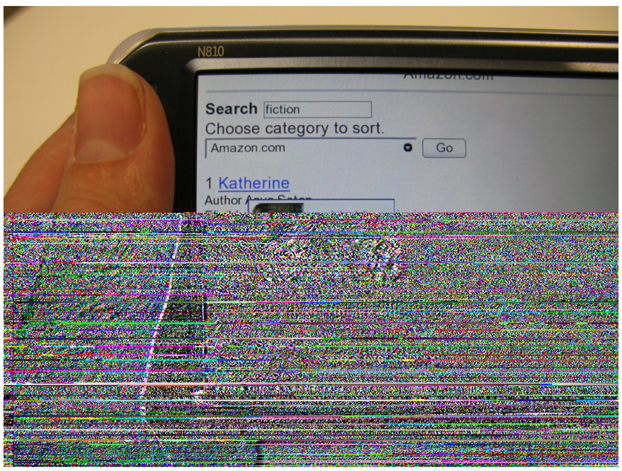

Figure 3. Viewing a SemantiLynx icon on a Nokia N810 Internet tablet.

Dominant Image Identification

Typically on a web page, there are several images, but many of them are often advertisements and logos. Dominant images are the most important and informative images that are often placed in the attentive areas of the web page. We use the following features for determining the dominant images [14]:

Aspect ratio. The aspect ratio of the image is computed as the ratio of its height and width. Dominant images have larger aspect ratios than non-dominant images.

Image format. We check if the image is an animation or not as a Boolean valued feature. Dominant images tend to be static. Advertisements, logos and banners tend to be animated.

Position. These are the x and y coordinates of the image. Dominant images tend to be located at the top or center of the web page.

Area ratio. The area ratio is the ratio of the image size to the web page size. Dominant images tend to have a higher area ration than the other images.

External source. This is a Boolean value to denote whether an image URL is provided by another website or not. Typically advertisements have different hosts than their hosting web pages.

For every image on the web page, we compute the individual feature values based on the above features. In order to compare the images with each other, we normalize the non-Boolean feature values of the images. Normalization is achieved by utilizing a linear function to map the minimum value of a feature to zero, and map the maximum value of the feature to one:

$$f(x) = \begin{cases} 0 & x = min \\ \frac{x-min}{max-min} & min < x < max \\ 1 & x = max \end{cases}$$

where x represents a feature, and min, max respectively represent the minimum and maximum values of this feature of al the images in the web page. The images are subsequently ranked such that the highest ranked image has the most number of maximum values of all the features.

Term Weighting

While it may be sufficient to merely utilize the most dominant images in the SemantiLynx icon, in practice we have found that often the highest ranked images may not completely represent the context of the web page. The image heuristics are intrinsically specific to the image properties themselves and not the textual content to which they are contextually related. Given the limited real-estate of the mobile browser and subsequently the real-estate of the icon that appear with the hyperlink, it is pertinent to maximize the relevance of the images used in the icon to help in the directed search task. We initially tried TF-IDF based term weighting with the assumption that frequently occuring words are representative of the webpage. However, in practice we observed that we needed to consider other characteristics inherent to the structure and layout of web content to identify more relevant keywords. We use the following heuristics in constructing an algorithm that extracts keywords in the web page [18]. The known semantics associated with the regularity of web pages makes it particularly amiable for our term weighting algorithm.

Remove stop words. Words such as ‘and’, ‘the’, ‘or’ are not good search terms.

Value frequently used words. Words used frequently are representative of the web page that the hyperlink points to.

Value emphasized words. Emphasized words used in titles, section headings and bold font draw more attention from the user, and are more representative of the web page’s content.

Value words that appear at the beginning of a web page more than words that occur at the end. Since reading is a linear process, words that occur earlier tend to be descriptive of the rest of the web page.

Figure 4. Saliency Based Region Extraction. From left to right: Original image, segmented image, .segmented image with a gray scale importance superimposed on it, extracted image region.

Punish words that appear to be intentionally de-emphasized. Words in small fonts (de-emphasized words) are exempt from the previous heuristic.

The pseudocode for the term weighting algorithm is as follows:

for word w collected **do**
  for (position p, style$s$) in $pos(w)$ **do**
    $weight = \frac{(numTerms^2 \delta)}{p^2}$
    if $weight > maxCount$ or s is the de-emphasized style **then**
      $weight = 1$
    end if
    if s is the emphasized style **then**
      $weight = 2 \times weight$
    end if
    $weight(w) = weight(w) + weight$
  end for
end for

Here, the first pass through all the terms in the document eliminates stop words. $maxCount$ is defined as the maximum number of times a given word has appeared in the web page. $numTerms$ is defined as the total number of words that are not stop words in the web page. $pos(w)$ contains a list of position-style pairs for a given word w. δ is a constant. In practice, a value of 0.2 tends to work well for δ.

Intuitively, the preliminary weight of a term varies inversely with the square of its position on the web page. This metric improves as the textual length increases, prompting the addition of the numerator that is proportional to a fraction of the square of the total number of terms, to reflect this fact. The term's final weight is the sum of the position weights that are less than $maxCount$, unless the term is de-emphasized. If a term is de-emphasized, its weight is 1. If a term is emphasized, its weight is doubled.

We run the term weighting algorithm on the main content portion of the web page, but separately run the term weighting algorithm on the associated text (image caption, alternate text, image file name) of the dominant images identified. The resulting term-weight pairs from the main content and the associated text of the dominant images are used to determine the relevance weights of the dominant images.

Determining Image Relevance

We use Vector Space Model (VSM) to compute the relevance weights of the dominant images to the given web page. The keywords in the main content of the web page is represented as an M-dimensional vector, where M is the number of distinct keywords in the document collection. Each element denoted by w_j is the weight of the jth word in the document. Similarly, associated text for each image is represented as a vector in the same term space. For each vector v of the web page term words, we use $cosine$ to evaluate its relevance with image vector i. Hence, relevance value between 0 and 1 is computed as:

$$r(v, i) = \frac{\vec{v} * \vec{i}}{|\vec{v}| * |\vec{i}|}$$

As the angle between the vectors shortens, the cosine angle approaches 1, meaning that the two vectors are getting closer, meaning that the relevance of the given dominant image to the overall context of the web page document, is high. The output at the end of this process, results in a scored set of dominant images based on weighted relevance. In practice, we have found that a threshold weight of greater than or equal to 0.3 is sufficient to have $3 - 5$ highest scored images for compositing into a SemantiLynx icon.

IMAGE RETRIEVAL

After determining the context of the hyperlink, there arise situations when there are either insufficient dominant images that are relevant enough or the target webpage contains only test, or highly weighted text in the web page for which there is no relevant image available on the page. In those cases, we leverage the term-weight pairs computed for the web page to automatically generate queries for retrieving images from a stock photography database
(`www.sxc.hu/`). This service provides a list of tightly coupled keywords associated with each image. The images typically are high-quality photographs that are subject-centric with uniform, neutral backgrounds, which simplifies our image extraction process. For example in Figure 2, 'free valet parking' is a phrase highly emphasized in the restaurant web page as parking is a premium in that area. By utilizing the highly weighted keywords to retrieve additional imagery, we enhance the context of the icon to include an image of a valet parking sign. In practice, we noticed that that users did not seem to get confused with the algorithm adding images to

the icon that did not come from the web page during a directed search task. We speculate that the reason being that the icons were primarily used to help users find web pages and were not used as web page bookmarks to indicate previously visited webpages.

SALIENCY BASED REGION EXTRACTION

To make the images as simple and recognizable as possible in the icon canvas size on the mobile screen, we want to remove unimportant visual information, particularly large backgrounds. Merely scaling a retrieved image and compositing it as an icon, may render it unrecognizable. We use a simple region extraction method based on image segmentation and image saliency (Figure 4). After segmenting the image into homogenous regions, we apply an importance map to identify important segmented regions in the image. This method tends to work well on images having few objects placed on a neutral background.

Image Segmentation

In order to identify important regions in the image, we must first segment the image. We use mean-shift image segmentation [5] to decompose the given image into homogeneous regions. The advantages of this approach include flexible modeling of the image and noise processes and consequent robustness in segmentation. The segmentation routine takes as input, the parameters: spatial radius h_s, color radius h_r, and the minimum number of pixels M that constitute a region. As with other segmentation methods, choosing optimal parameter values is often difficult. Therefore we over-segment the image using lower values of h_r and M and merge adjacent regions based on color and intensity distributions in a perceptually uniform color space, CIE-Luv. In practice, values of $h_s = 7, h_r = 6$, and $M = 50$, tends to work well for over-segmentation for most images.

We use a two step process to convert to CIE-Luv to account for the dependence of color appearance on spatial structure [7]. We first convert the input RGB image to the L$\alpha\beta$ opponent color space [16] that consists of three color planes, O_1, O_2, O_3, representing luminance, red-green, and the blue-yellow planes separately. Then each of the planes is smoothed directly by applying Gaussian kernels. These kernels are the sums of Gaussian functions with different values of standard deviation σ, computed according to:

$$\frac{1}{m}\Sigma_i \frac{w_i}{n_i} e^{\frac{x^2+y^2}{\sigma_i^2}} \quad (1)$$

We use the values of (w_i, a weighting term, and the standard deviation σ_i) from the work by Mirmehdi and Petrou [16]. The authors performed psychovisual measurements on human subjects to obtain these values. n_i is used to normalize the sum of the matrix elements of each Gaussian kernel, and m normalizes the sum of the elements of the final matrix to 1.

Once the kernels are applied to the image, we carry out color measurement in CIE-Luv, a perceptually uniform space. The pixels representing each segmented region are used to form 3D color histograms that contain a range of values for each of the color components. We then compute a color similarity measure called histogram intersection [20] to determine color similarity between regions, and perform region simplification by merging adjacent regions. Histogram intersection matches the image color histogram of a given segmented region with histograms of each of the adjacent regions. Given a pair of histograms, Q and T, each containing n buckets, the intersection of the histograms is defined to be:

$$\sum_{j=1}^{n} min(Q_j, T_j).$$

where j ranges over each color in the histograms.

Importance Map

To identify important regions, we first compute an importance map that assigns a scalar value to each pixel estimating the importance of that image location based on an attention model. Like previous methods [9] we use measures of visual salience (e.g. image regions likely to be interesting to the low-level vision system) and high-level detectors for specific objects that are likely to be important, such as faces and signs. Our implementation computes the importance map as a scaled sum of a visual saliency algorithm and a face detection algorithm.

The saliency and face detection algorithms take color images as input return gray-scale images whose pixel values represent the importance of the corresponding pixel in the input image. The importance map, which is the attention model for the image, is built up by combining a series of importance measures. We normalize pixel values from the attention model output images and sum them then re-normalize to create the importance map.

Image Attention Model

We use the saliency-based image attention model [10] to generate the first contribution to the importance map. The saliency model is used to extract attended locations in complex scenes based on a low level visual model that uses color, intensity and edge orientation as visual cues. The technique uses Gaussian pyramids to compute several ‘feature maps’ for three low level features: color C, intensity I, and orientation O, which represent the visual scene. Such feature extraction is achieved through linear filtering for the given feature type, followed by a center-surround operation which extracts local spatial discontinuities for each feature type. Spatial discontinuity locations are then combined into a unique ‘saliency map’ represented as:

$$S = \frac{1}{3}(N(I) + N(C) + N(O)) \quad (2)$$

where N denotes normalization.

The two-dimensional topographical saliency map is used to determine the importance values within the original image.

(a) Design rules used to composit the SemantiLynx.

(b) Images with their corresponding weights.

(c) Process of compositing the icon.

Figure 5. Bottom row, from left to right: Step 1: Place the most important image as large as possible onto the icon template with the center of the image placed in the first 1/3rd segment of the template. Step 2: Place the second most important image onto the icon template with a scaling factor and translucency proportional to its weight. The image is placed with center on the third 1/3rd segment of the template making sure it does not occlude the most important image. Steps 3 and 4: Place the next most important images on any remaining real-estate of the icon template or if there is no space, composited on the last 1/3rd segment of the template. The scaling factors and translucency are proportional to its weight. Perform blending along the seams of the image to create the final icon.

We binarize the saliency map to find the $ROIs$. The IV can be calculated as:

$$IV_{saliency} = \sum_{(i,\, j \in R)} B_{i,j} \cdot W^{i,j}{}_{saliency} \qquad (3)$$

$B_{i,j}$ denotes the gray-scale value of pixel (i, j) in the saliency map. Since people pay more attention to the region near the center of an image, a normalized Gaussian template centered at the image is used to assign the positional weight $W^{i,j}{}_{saliency}$.

Face Attention Model
Images of people are popular as well as important in many application areas. However, saliency map generation relies only on low-level features, and it might not be able to recognize faces correctly. The face is a highly important characteristic of human beings, and dominant faces in images certainly attract viewers' attention. Therefore, we use a face attention model in addition to the image attention model. By applying face detection [17], we obtain information about faces in the image such as position, region, and pose. The size and position of a face usually reflects its importance. Hence, the importance value in this model is calculated:

$$IV_{face} = \sqrt{Area_{face}} \cdot W^{i}_{pos} \qquad (4)$$

where $Area_{face}$ denotes the size of the detected face region, and W^{i}_{pos} is the weight of its position and $i \in [0, 8]$ is the index of the position.

The visual attention and face detection models are integrated together to get importance information. Currently, we adopt a linear combination to implement fusion scheme due to its effectiveness in detecting frontal and large profile view faces, and its simplicity. With such a scheme, each attention model should be normalized to $[0, 1]$. The IV of each IO is normalized to $[0, 1]$ and the final IV is computed according to the equation. We normalize importance values from both attention models and compute the importance value IV for each importance object IO as:

$$IV_i = w_k \cdot \overline{IV_i^k} \qquad (5)$$

where w_k is the weight of the model k, and $\overline{IV_i^k}$ is the normalized importance value of the IO_i detected in model k.

IMAGE COMPOSITION

A good visual composition for conveying the right information often relies on placing important elements, *i.e* the focal point of the design, within the visual center of a piece [8]. The commonly used graphics design rules we use in the system to achieve this goal are the rule of thirds, dominance and proportion. The *rule of thirds* states that that most visual compositions can be made more interesting by dividing the canvas into thirds vertically and/or horizontally and placing the most important element at one or more of the four intersections of those lines. *Dominance* is used to create a hierarchy of visual importance based on the relevance weights of the images. *Proportion* allows optimization of the placement of the images to maintain a sense of harmony and balance.

We implement these rules in our automatic compositing technique by initially placing the highest scoring image as large as possible on the icon canvas, and iteratively adding the other images in decreasing order of importance based on these design rules to produce the final SemantiLynx icons (Figure 5).

USER STUDY

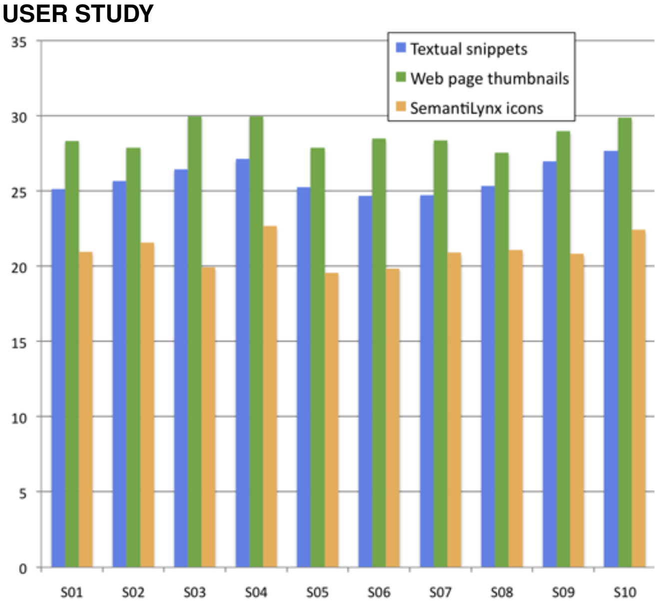

Figure 6. The average time over each of the directed search tasks is presented per participant. The X axis indicates the subject IDs and the Y axis indicates time in seconds. Average times are 20.97 seconds, 25.89 seconds and 28.72 seconds for SemantiLynx icons, text snippets and web page thumbnails respectively.

Studies have shown that search tasks in simplied environments may proceed faster than similar tasks executed in the full environment, and in particular, simplified images are often easier to understand than text [2]. Based on the results of prior studies, we hypothesize that SemantiLynx icons may decrease the amount of time needed to search for a specic web page among a collection of webpages on a mobile browser. This section presents the results of a user study to assess the performance of SemantiLynx icons as previews for directed search tasks on mobile devices with different screen resolutions and input modalities. In the study, the SemantiLynx icons are compared with two other preview representations - text snippets containing the keywords from the web pages, and scaled down thumbnails of the web pages similar to those created at `http://www.websnapr.com/` and `http://www.ask.com`.

Phase 1

The first phase of the user study involves a set of direct search tasks simulating a fact-finding exploration on two mobile devices. The first device is a Nokia N810, an Internet tablet equipped with a screen resolution of 800×400, and the display measuring 4.13" along the diagonal. The N810 has a touch screen interface where the icons appear over the hyperlinks when the users makes a light tap on the screen. The second device is a Nokia N95 phone with a screen resolution of 240×320, and the display measuring 2.6 " along the diagonal. The N95 is a non-touch screen device where the icons appear over the hyperlinks when the user selects the hyperlink with the joypad button.

There are a total of 80 trials across both devices, with 40 trials on each device using a within-subject design. 10 subjects (6 males, 4 females) participated in both sets of trials, including 2 practice trials. Subjects comprised 2 administrators, 1 person from a book store, 4 undergraduate students, 2 computer researchers and 1 doctor. Subjects had either normal or corrected-to-normal vision. Neither study shows gender being a significant factor.

The trials are randomly ordered across both the devices. In addition, each trial uses one of the three randomly chosen representations of SemantiLynx icons, textual snippets or web page thumbnails. The studies consist of 40 trials with 10 hyperlinks in each trial. The search terms used in the studies are the result of randomly chosen search terms from a compiled list of recent directed searches collected from the study volunteers that tend to be quite specific, such as 'pasta restaurant in downtown vancouver', and 'nobel prize winner physics 2008'. For each trial, there is only one correct hyperlink, while the other 9 are distracters. To mimic a search engine in the real environment, the distractors partially match the search terms. Each participant is asked to click on the hyperlink that best matches the search term provided on top of the web page in each trial. The participant moves to the next trial by clicking a 'Next' button located on the search page. Between trials, participants are presented a gray screen on the phone for 1.5 seconds.

Table 1. Average number of hub-and-spoke actions

Text Snippets	Web page Thumbnails	SemantiLynx
15.62	22.86	8.43

Table 2. Average time for finding Hyperlinks in Phase 2 (seconds)

Text Snippets	Web page Thumbnails	SemantiLynx
25.81	28.35	20.02

The average search time when SemantiLynx icons are used, is 20.97 seconds (Figure 6). This is 7.75 seconds faster than web page thumbnails and 4.92 seconds faster than text snippets. The accuracy of recognition is 99.9% for the text snippets and web page thumbnails, and 98% for SemantiLynx icons. Further, we observed a noticeably improved performance with SemantiLynx icons on the N95 with smaller display resolution. For this device, search times using SemantiLynx icons is 4.98 seconds faster than text snippets and 7.99 seconds faster than web page thumbnails. Using repeated measures ANOVA, there is a significant statistical difference ($p = 0.005$) in performance between the timing of the judgments in the first phase of the study. Table 1 shows the average number of hub-and-spoke actions, i.e. web page revisits for each user. An analysis of the average number of hub-and-spoke actions show a statistical significance for textual snippets and web page thumbnails versus SemantiLynx icons ($p = 0.007$), with fewer such actions for directed searches using SemantiLynx icons.

Phase 2

The second phase of the study is a duplicate of the first performed by the same subjects on the same data after a two-hour period. Such a priming effect can be beneficial as often times users seek the same page they have visited in the past. The average time for finding the given link in this phase using SemantiLynx is 20.02 seconds, an improvement of 0.95 seconds. Text snippets and web page thumbnails recorded performance improvements of 0.08 seconds and 0.37 seconds respectively (Table 2). For the second phase of the study, the accuracy of recognition is 100% for the text snippets, web page thumbnails and SemantiLynx icons. Using a repeated measures ANOVA, we find that there is a significant statistical difference ($p = 0.003$) in performance timing.

DISCUSSION

The SemantiLynx system creates semantically meaningful icons to better reveal the content pointed by hyperlinks. The results obtained from the user studies indicate that SemantiLynx tend to improve the experience of directed search tasks on mobile phones by minimizing the time spent on web page revisits and better recalling links visited during the search.

We conducted exit interviews with the participants to obtain a qualitative sense of the icons in the user studies. Participants found the images in the SemantiLynx icons to be helpful when trying to quickly visually discern the variation of the relevance of the hyperlinks to the task questions posed in the studies. Several participants indicated that it was often hard to remember all the keywords in the text snippets and the web page thumbnails were difficult to parse visually due to the limited screen real-estate of the mobile devices.

The effectiveness of the SemantiLynx icons however, does depend on various factors and has the following limitations:

1. The heuristics used in determining the relevance of imagery in the web pages are motivated based upon good web design practices. We have found that when web pages do not conform to these rules, the images and keywords retrieved from the web page are suboptimal leading to less relevant and confusing icons being generated. Given our dataset, the dominant image identification had a 89.7% accuracy. The failure cases include webpages that have lots of images and advertisements. In addition, the icon tends to be less informative for cases where the webpage is a search result page showing several search result items, where each item is associated with an image. In such a case, only a few of these images are used in the final image composition.

2. The region extraction method used to remove visual clutter in the images, depends on the segmentability of the images. We found that images that are subject-centric with neutral backgrounds are more favorable for this process.

3. If all the hyperlinks on a search page point to nearly the same content, the SemantiLynx icons generated for these hyperlinks may not be very visually distinguishable leading to more hub-and-spoke actions.

We ran the icon generation algorithm on a Dell Optiplex GX280 CPU 3.06 GHz processor with 2 GB RAM. The average time taken to generate each SemantiLynx icon is 2.7 seconds when there is minimal network latency. The icons are generated once and associated with the web page at creation time, for subsequent viewing on a variety of small-screen web browsers.

The ubiquitous web is a reality and more desktop portals now also have mobile versions. The majority of web sites often take existing web content tailored for desktop viewing, and make minimal optimizations for mobile viewing. We are currently exploring integrating the SemantiLynx software into mobile web creation tools, and investigate ways in which the quality of the icons created, could in turn influence the design of the web pages themselves as a feedback loop. Further, allowing web designers to have control over the imagery and the compositing rules during the SemantiLynx icon creation, could help facilitate new directions in web design and browsing for a variety of small-screen ubiquitous devices.

CONCLUSION

Computing infrastructure is constantly evolving to deliver rich amounts of web content to home theater screens, cellular phone displays, networked PDAs, and even displays embedded in refrigerators, elevators and airplane seats. Our vision is that increasingly ubiquitous displays can provide people with information when and where they need it, provide more effective channels of inter-personal communication, deliver educational media, and provide entertainment. Achieving this vision requires providing web content amenable for a variety of display devices. The SemantiLynx work is a step in that direction.

We have demonstrated an automatic *visual* technique to improve directed searching using hyperlink navigation on a small screen device. The system provides a tool for automatically augmenting hyperlinks with semantically meaningful icons before the web pages get published to be viewed on the device. In areas where screen real-estate is limited, SemantiLynx provides an increase more than proportionate to the space it takes. The real strength of our approach lies in leveraging the semantic information available on the web to automatically compute visualizations that improve human performance. We report the results of two user studies carried out using SemantiLynx as stimulus. Results show that SemantiLynx may yield quicker response times and improved retention in directed search tasks on various small-screen web browsers.

REFERENCES

1. S. Björk, L. E. Holmquist, J. Redström, I. Bretan, R. Danielsson, J. Karlgren, and K. Franzén. West: a web browser for small terminals. In *UIST '99: Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 187–196, New York, NY, USA, 1999. ACM.
2. M. D. Byrne. Using icons to find documents: simplicity is critical. In *INTERCHI '93: Proceedings of the INTERCHI '93 conference on Human factors in computing systems*, pages 446–453, Amsterdam, The Netherlands, The Netherlands, 1993. IOS Press.
3. L. D. Catledge and J. E. Pitkow. Characterizing browsing strategies in the world-wide web. In *Proceedings of the Third International World-Wide Web conference on Technology, tools and applications*, pages 1065–1073, New York, NY, USA, 1995. Elsevier North-Holland, Inc.
4. E. H. Chi, P. Pirolli, K. Chen, and J. Pitkow. Using information scent to model user information needs and actions and the web. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 490–497, New York, NY, USA, 2001. ACM.
5. D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, May 2002.
6. J. Conklin. Hypertext: an introduction and survey. *Computer*, 20(9):17–41, 1987.
7. D. J. Fleet and D. J. Heeger. Embedding invisible information in color images. In *IEEE Signal Processing Society 1997 International Conference on Image Processing (ICIP'97)*, pages 532–535, October 1997.
8. J. A. Gatto. *Exploring Visual Design: The Elements and Principles*. Davis Publications, 2000. 282 pages.
9. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.
10. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, 1998.
11. M. Kellar, C. Watters, and M. Shepherd. The impact of task on the usage of web browser navigation mechanisms. In *GI '06: Proceedings of Graphics Interface 2006*, pages 235–242, Toronto, Ont., Canada, Canada, 2006. Canadian Information Processing Society.
12. S. J. Koyani, R. W. Bailey, and J. R. Nall. *Research Based Web Design and Usability Guidelines*. Computer Psychology, 2004. 232 pages.
13. H. Lam and P. Baudisch. Summary thumbnails: readable overviews for small screen web browsers. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 681–690, New York, NY, USA, 2005. ACM.
14. Z. Li, S. Shi, and L. Zhang. Improving relevance judgment of web search results with image excerpts. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 21–30, New York, NY, USA, 2008. ACM.
15. S. Liu and K. Tajima. Wildthumb: a web browser supporting efficient task management on wide displays. In *IUI '10: Proceeding of the 14th international conference on Intelligent user interfaces*, pages 159–168, New York, NY, USA, 2010. ACM.
16. M. Mirmehdi and M. Petrou. Segmentation of color textures. In *Transactions on Pattern Analysis and Machine Intelligence*, pages 142–159. IEEE, 2000.
17. C. M. U. Robotics Institute. Face detection demonstration.
18. G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Commun. ACM*, 18(11):613–620, 1975.
19. V. Setlur, C. Albrecht-Buehler, A. Gooch, S. Rossoff, and B. Gooch. Semanticons: Visual metaphors as file icons. *Computer Graphics Forum*, 2005.
20. M. Swain and D. Ballard. Color indexing. *International Journal on Computer Vision*, 7(1):11–32, 1991.
21. J. Teevan, E. Cutrell, D. Fisher, S. M. Drucker, G. Ramos, P. Andre, and C. Hu. Visual snippets: Summarizing web pages for search and revisitation. 2009.
22. J. O. Wobbrock, J. Forlizzi, S. E. Hudson, and B. A. Myers. Webthumb: interaction techniques for small-screen browsers. In *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 205–208, New York, NY, USA, 2002. ACM.
23. A. Woodruff, A. Faulring, R. Rosenholtz, J. Morrsion, and P. Pirolli. Using thumbnails to search the web. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 198–205, New York, NY, USA, 2001. ACM.