Image Representations for Accessing and Organizing Web Information

Jonathan I. Helfman[a] and James D. Hollan[b]

[a]AT&T Labs West, Menlo Park, CA 94025

[b]Distributed Cognition and HCI Laboratory
Department of Cognitive Science, UCSD, La Jolla, CA 92093

ABSTRACT

The web is enormous and constantly growing. User-interfaces for web-based applications need to make it easy for people to access relevant information without becoming overwhelmed or disoriented. Today's interfaces employ textual representations almost exclusively, typically organized in lists and hierarchies of web-page titles or URL taxonomies. Given the ability of images to assist memory and our frequent exploitation of space in everyday problem solving to simplify choice, perception, and mental computation, it is surprising that so little use is made of images and spatial organizations in accessing and organizing web information. The work we summarize in this paper suggests that spatial and temporal organization of selectable images may offer multiple advantages over textual lists of titles and URLs. We describe several image-based applications, detail basic image representation techniques, and discuss spatial and temporal strategies for organization.

Keywords: hierarchical cache visualization, image rating, image scaling, informational zoom, representative image, spiral layout, visual bookmarks, visual query results, visual session summaries, visual site indexes

1. INTRODUCTION

There is currently a shift in cognitive science away from the notion of cognition as a property of isolated individuals and toward a view of cognition as a property of larger social and technical systems.[1,2,3] This new approach extends the reach of cognition to encompass both interactions between people as well as interactions with resources and materials in the environment. This shift in perspective comes at a time of rapidly changing computational and networked infrastructure. Unfortunately, while the speed of interaction and the quantity and diversity of available information continue to grow, our methods for interacting with digital resources and materials remain virtually unchanged. We are locked within early metaphors of interaction, alienated from the familiar spatial organizing principles of the physical environment, and restricted from fully exploiting our exquisite abilities to use space to enhance communication, support collaboration, and assist in dealing effectively with the unceasingly expanding information collections we confront daily.

Kirsh classifies the functions of space into three main categories: spatial arrangements that simplify choice, spatial arrangements that simplify perception, and spatial dynamics that simplify internal computation.[4] In a series of studies involving several different task environments he found subjects used space to simplify choice by creating arrangements that served as heuristic cues. At other times subjects used spatial arrangements to change the memory requirements of tasks by positioning items needing immediate attention nearby. Similarly, Kirsh and Maglio found in laboratory studies of the computer game Tetris that players physically manipulate forms to save themselves computational effort, modified the environment to cue recall and speed identification, and made changes to the world to avoid costly and potentially error-prone computations.[5]

In addition to employing spatial layouts as an integral part of everyday problem solving, there is a voluminous literature documenting that people are particularly effective at processing images. For example, Shepard had subjects view 600 vacation slides.[6] Using a testing procedure in which subjects were required to decide which of two pictures they had seen before, he found that people were 98% correct immediately after seeing the original slides. In a similar study, subjects saw 2,560 slides for ten seconds each.[7] After several days, recognition memory was still above 90%. Paivio records numerous experiments in which associating pictures with words improves peoples' memories for the words.[8] While the similarity of images to each other and to distractor items can influence these results, the literature is clear that people have a remarkable ability to process images and their recognition memory for them is excellent.

Given the cognitive properties of images to assist memory, the important role space plays in everyday cognition, and our frequent exploitation of space to simplify choice, perception, and mental computation, it is surprising that so little use is made of images and spatial organizations in accessing and organizing web information.

1.1. Textual representations

People access and organize web information primarily by interacting with web-page titles and URLs. Each title and URL is an example of a textual representation — a short piece of text that represents a document in a user interface. Bookmarks and browsing histories are almost always presented as lists of textual representations. Search engine query results are also conveyed as lists of textual representations, often including portions of text from the web pages that match the query. Portals like Yahoo! are organized as large hierarchies of textual representations.

While interfaces that use textual representations may be adequate for helping people access and organize document collections, they are less effective on the web where people must process very large amounts of poorly-organized, constantly-changing information. When anyone can be an author or publisher, there is no single unifying organization, and finding information can be difficult. People rely on search engines to assist in finding information, but search engines deliver lists of thousands of textual representations, which can only be read one item at a time. Long lists of textual representations are not only time-consuming to read, but may also span multiple pages or require scroll bars, user interface techniques that hide information. When most of a list is hidden, it is difficult to see patterns in the results or to otherwise determine how best to reformulate a query. If what is being searched for is not near the top of the first page, it is unlikely to be found.

Despite the success of graphical user interfaces, which use icons to represent digital objects and commands, text is still the dominant form of user interface representation in almost all tools for accessing and organizing digital information. The use of textual representations for accessing and organizing web information is partially the result of concerns with efficiency. Text requires a very small amount of space in digital memory and can be rapidly transmitted over a digital network. Text is used for programming and for naming objects that need to be manipulated by programs, such as files, folders, devices, and other programs. Currently, even multi-media tools rely on menus of file names or URLs for accessing and organizing digital images, audio, and video.

1.2. Image representations

Images provide an alternative and complementary way to represent information. While text is an efficient for computers, images may be significantly more efficient representations for people. People seem to be more efficient at processing visual information than textual information for tasks related to access and organization. People can search and process multiple images in parallel, while text must be decoded serially. As described above, an extensive literature details the importance of images in memory and reasoning. Because they exploit characteristics of human memory and perception, interfaces that use images to represent information may be better for helping people access and organize very large poorly-structured information sources like the web.

In many cases, images from a web page provide good representations for the page’s content. Web pages about people almost always contain a picture of the person's face. Web pages about places commonly contain a picture of a landmark that clearly identifies the place. Web pages about objects, particularly objects for sale, almost always contain pictures of the objects. Even highly technical documents, such as chemistry abstracts, usually contain illustrations, charts, or diagrams, which may be even more descriptive of the content than the documents’ titles.[9] Web images are also commonly used for site differentiation. Images from the same web site often have a shared style that is indicative of the site's genre.

While the software engineering and computer science communities typically use the term *image representation* to refer to an abstract data type for representing an image to a computer, in this paper we use the term *image representation* to refer to an image that represents a web page in a user interface, and specifically to an image that, when selected, provides access to the represented web page in a regular web browser. We believe that selection is the most direct and intuitive way to provide access to represented web pages. Ensuring that image selection works in a consistent manner makes the user interfaces for image representation systems easier to learn and to use, and may also make them practical from a legal standpoint (see Section 3.1).

In Section 2 we introduce some promising applications of image representations. As we have gained experience designing and implementing these applications, we have identified several techniques that seem generally useful for systems that use images to represent information. Section 3 describes the most basic of these image representation techniques, such as image selection, rating, and scaling. Additional techniques for organizing multiple image representations in space and time are described in Section 4. Finally, to better understand the limitations of the techniques, in Section 5 we discuss them in the context of the applications, emphasizing the tradeoffs that occur when organizing images in both space and time.

2. APPLICATIONS OF IMAGE REPRESENTATIONS

2.1. Visual bookmarks and site indexes

Bookmarks for web pages are typically stored as web page titles, even though many titles don't convey enough information about the content of the bookmarked page to help people remember why they created it. In fact, it is not uncommon for web page titles to be blank. Instead of using web page titles, visual bookmarks use images from bookmarked pages to represent the pages. For most web pages, images from the page may convey more information about the content of the bookmarked page than the page's title. Because images have been shown to help people remember textual information, images from a bookmarked page may do a better job of helping people remember why they made a particular bookmark.

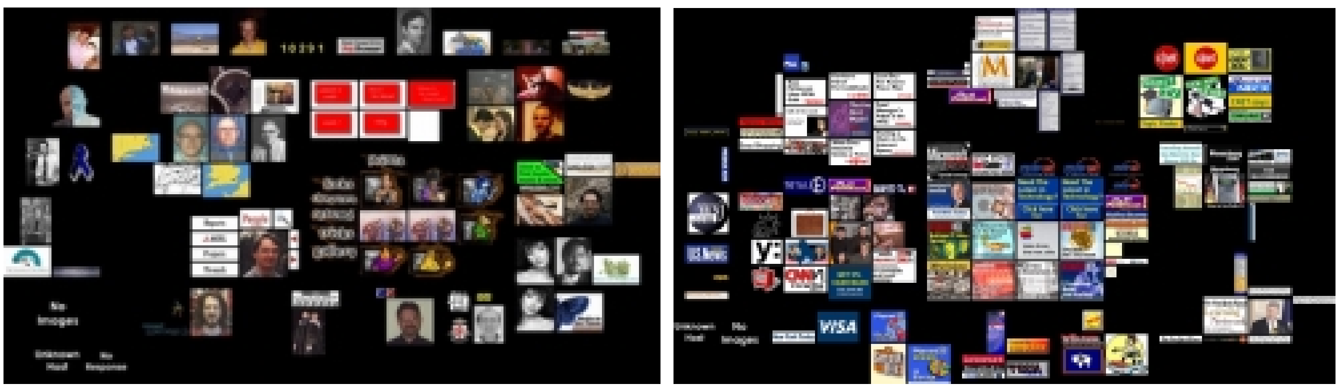

Figure 1. Visual bookmarks using iterative spiral layout

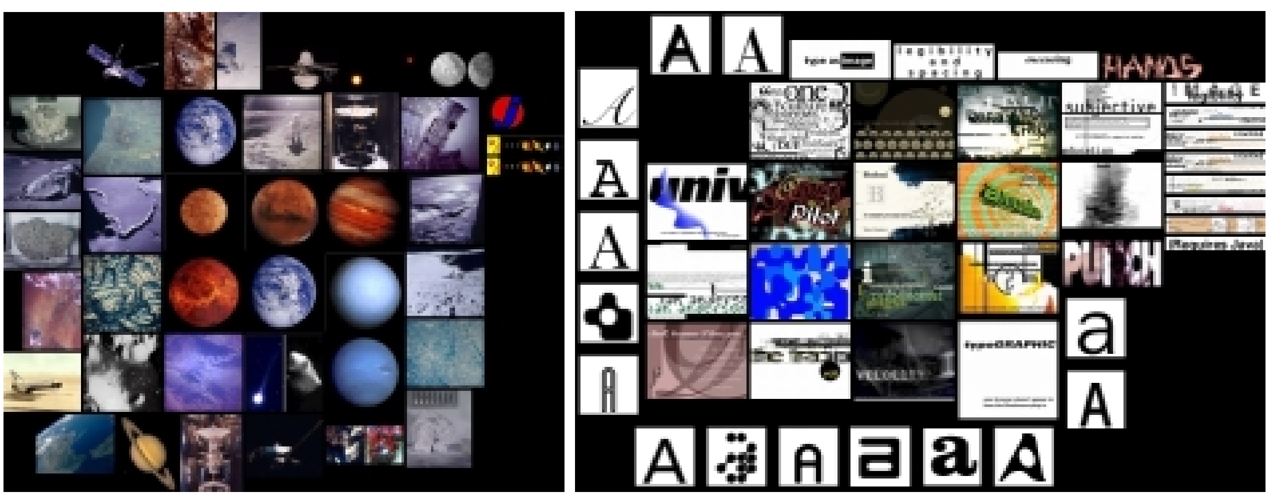

Figure 2. Visual site indexes using spiral layout

Figure 1 illustrates several bookmarks from two different categories. The images are scaled and positioned using the techniques described later in this paper. Each layout is stored as a client-side image map so the selectable regions within each image map correspond to small images (or thumbnails) within each layout.

A visual site index is a collection of images from several web pages at a common web site. Figure 2 shows visual indexes for two different sites. Unlike textual site maps, which can burden users with large amounts of text to decode, visual site indexes let people perceive multiple representations in a single glance, providing an immediate appreciation of what it would look like to visit a web site. Visual site indexes are useful for comparing multiple sites, illustrating site overviews, or introducing sites to others.

Rather than representing hypertext relationships between web pages, visual site indexes let people form their own associations between the images, which are more likely to be relevant to their tasks and interests. For example, a person interested in spacecraft and satellites could easily distinguish their images in Figure 2, left, even though the associated pages may be distributed throughout the web site.

2.2. Visual browsing session summaries

Figure 3. Visual browsing session summary using multi-scale layout

Visual browsing session summaries aid people in recovering from interruptions and can be employed to depict overviews of multiple browsing histories or tours of web sites (see Figure 3). Visual browsing session summaries are created by using an intermediary, such as an Internet Service Provider or a proxy server, to track a user's requests. Thumbnails of requested images or pages are then assembled into visual maps. Because people can remember many more images than words, and even use images to improve their memory for words, there is a better chance that visual session summaries will help people recover context when they need to revisit web pages. We have shown in a series of experiments that visual web browsing histories are more effective than traditional forms of histories.[10] We have also explored multi-scale layouts to more effectively use screen space and better reflect the structure of browsing sessions.

2.4. Visual query results

Image representations can also be used to augment familiar web tools. We are extending a textual search engine so that query results will include thumbnails of images from associated pages. We expect that when using visual query results it will be much faster and easier for people to identify a relevant image in a collection of images than a relevant title in a collection of web page titles.

Figure 4. A portion of the visual query results for the Google search: ``Taj Mahal''

To test the concept, we have written a proxy server, which adds thumbnails to the results of a Google web search (see Figure 4). The proxy server intercepts the results page before returning it to the web browser. It retrieves each web page that matched the query and determines if there are any images on the page. If there are, it retrieves and rates each image (see Section 3.2), creates a thumbnail of the highest rated image, and adds HTML code to the results page to reference the thumbnail. Finally, it returns the modified results page to the web browser, which displays the thumbnails adjacent to the titles of each web page that matched the query.

Visual query results seem to be particularly helpful for queries that use search terms with multiple meanings. In Figure 4, for example, it is possible to immediately distinguish between images that link to pages about the Taj Mahal monument and the Taj Mahal musician. Because most peoples' queries consist of a small number of search terms, techniques that help disambiguate query results quickly and easily have the potential to dramatically improve web searches.

One problem with our initial proxy server design is that it could be much faster if it pre-computed and stored thumbnails for each of the pages that are indexed by the search engine. We are currently redesigning our system in conjunction with the creators of a new search engine at AT&T Labs. To keep our system functionally independent of the search engine, our proxy server design is maintained, but simplified to work with pre-computed and stored thumbnails. We use multiple processes that read the search engine's logs and attempt to store a thumbnail for each of several million indexed pages. To minimize database and file system limitations, multiple thumbnails are stored together in a single file. The proxy server modifies the results of the search engine by inserting calls to a CGI script, which maps the URL of an indexed page to the appropriate thumbnail.

For all but the fastest connections, the visual query result application has bandwidth limitations. Because the query results page must load quickly, it must contain a minimum number of very small images. In fact, we have found that the most difficult aspect of implementing visual query results is that many of the indexed pages reside on slow web servers, which delays the task of determining a representative image. We finesse this problem by providing only thumbnails for pages on responsive servers. A side effect and potential benefit for users is that query results with thumbnails are indicative of pages on fast web servers, a piece of information not normally available yet particularly useful for users in a hurry.

2.5. Hierarchical cache visualization

As a final example, we are designing a hierarchical categorization portal to include image representations. One problem with taxonomies, such as Yahoo!, is that it is difficult to tell which links are useful or even which ones are regularly accessed by other users. Our design augments a taxonomy portal with separate proxy server caches for distinct sub-trees of the categorization hierarchy. For example, when users select links within the "science" category, the pages and images associated with the links will be cached by the "science" proxy server. Users will be able to view a hierarchical cache visualization displaying each of the images within the "science" cache. Hierarchical cache visualizations not only allow people to view representations of multiple pages in a single glance, but because the caches will only contain the most frequently and recently accessed pages, it will be immediately apparent which links are being regularly accessed by other users.

3. BASIC IMAGE REPRESENTATION TECHNIQUES

Each of the systems mentioned above makes use of a small set of basic image representation techniques. Here we describe techniques for providing access via selection, rating images, choosing representative images, and scaling images.

3.1. Image selection provides access

Perhaps the most fundamental image representation technique is the ability to select an image to access the web page that it represents. In order to implement image selection, associations must be maintained between images and web pages. In some cases, it is possible to store these associations together with the images. For example, we typically store a group of image representations as a client-side image map, which encodes URLs of image files along with URLs of the web pages that they represent. In the general case, however, there may be such a large number of associations that they must be stored and accessed through a dedicated server, as is the case when adding thumbnails to the results of a web search (described above in Section 2.4).

Implementing image selection can be difficult because most web browsers require platform-dependent code to direct them to retrieve and display web pages. Netscape's web browsers, for example, use X events on UNIX machines, Apple events on Macintosh machines, and either OLE2 (Object Linking and Embedding) or DDE (Dynamic Data Exchange) on Windows machines. Providing support for multiple platform-dependent APIs makes application code brittle and can compromise the portability of applications written in platform-independent languages such as Java.

Although implementing image selection may be difficult, it may also be crucial for the legality of image representation systems. Systems that use image representations might be accused of copyright infringement because they remove images from their original context. In many cases, the images are also reduced in size and, as a side effect of scaling, the colors of the original images are invariably changed. Using the derived images for personal applications, such as for illustrating personal bookmarks (see Section 2.1) would almost certainly constitute “fair use”. Using the derived images to generate a profit by anyone other than the creators of the original images, however, would likely be copyright infringement. But the legal issue is complex. Copyright law for digital media is still evolving. For example, consider that web search engines extract textual content from web pages and use the extracted content outside the context of the original web pages, on new web pages, which may indeed contain advertisements that generate a profit. A reason that this is not copyright infringement is that selecting a web page title or URL within a query result provides immediate access to the original web page that matched the query. The primary purpose of a search engine is, in fact, to promote access to the original source materials. The same logic should apply to image representation systems that enforce a similar policy in which selecting an image provides access to the original page from which the image was taken.

3.2. Image rating

Web pages often contain multiple images, some of which do a better job of representing the content on the page than others. Images that usually do a poor job of representing content are those that are used for decoration, navigation, and advertisement. Images that typically do a better job of representing content are photographs of people, places, or things, and diagrams, charts, and logos. While it is quite difficult to automatically determine what is depicted by any particular image, it is relatively easy to rate images by using heuristic measures of how likely they are to represent the content on the page. Image rating is a general technique that is useful for guiding image organization strategies (see Section 4) and for helping to select a single image to represent a group of images or a web page (see Section 3.3).

We have developed a very fast rating scheme, using only information available in image file headers, that tends to down-rate images used for decoration, navigation, and advertisement. Our rating scheme was derived empirically. After manually rating several hundred web page images as either “representative” or “non-representative,” we determined that most “representative” images are large, colorful, and have an almost square aspect ratio, while most “non-representative” images are small, have few colors, and non-square aspect ratios.

While our basic rating scheme uses only image dimensions and number of colors, we have also explored minor heuristic variations. For example, we exploit whether or not an image is used as a background texture or a link to a resource on a different web server (in many cases web advertisements are links to a different host).

3.3. Choosing a representative image

It is often useful to be able to choose a single *representative image* from a group of images.[11] One example is when a single image must be chosen to represent a web page containing multiple images. One might be motivated to do this when bandwidth limitations restrict the total number of image representations. A second example is when when building a visual category hierarchy where a single image must represent an entire category of documents and their associated images (see Section 2.5). In some cases, it may be possible to let a person choose which image best represents the group. The images can be displayed and the person can select the best one. When there are a small number of choices to be made, a manual solution may be preferable to an automatic solution, because it is much more likely that a person will make an effective choice. When there are a large number of choices to be made, human guidance can become impractical, and an automatic solution may be required. One automatic way to choose a representative image is to use the image rating heuristics discussed above, choosing the highest rated image to represent the group. We use this approach when selecting images to represent web pages in the results of a search engine query (see Section 2.4).

3.4. Image scaling

Unlike most textual representations, most images can be reduced by several scale factors and still be recognizable. When images are reduced in size, people can see more of them at once. Scaling images so that their maximum dimension is a fixed, small size can also facilitate organizing multiple images so that they don’t overlap (See Section 4.1). Scaling images is well understood in terms of sampling theory and to scale images without introducing aliasing artifacts usually requires some form of filtering. Scaling large images can be very slow, however, unless the required scale factor is known in advance and the filter weights can be preprocessed appropriately.[12] Unfortunately, when scaling web images so that their maximum dimension is a fixed size, the required scale factor cannot be predicted because web images can be any size.

To scale large web images quickly and accurately, we have developed a hybrid technique that combines a very fast algorithm for scaling images by a factor of two[13] with a slower algorithm for scaling images by arbitrary amounts.[14] Images are reduced by successive factors of two, until an additional reduction would make them too small. At this point, the resulting image is scaled to its final size by the slower algorithm, which is now guaranteed to be faster because it never needs to scale an image by more than a factor of two.

Our hybrid scaling technique is almost as fast as uniform point sampling with no filtering at all. We have implemented it within an *image server*, a C process with a socket interface that scales GIF and JPEG web images and returns the resulting image data or stores it on a web server.[15] We have found the modularity of the image server to be very convenient and have reused the server in each of our image representation applications as well as in other application not reported here. The image server also creates image maps using a variety of layout algorithms (see Section 4.1) and reads GIF and JPEG headers to identify information required for image rating.

4. ORGANIZATION OF IMAGE REPRESENTATIONS

We have introduced image representations, illustrated several of their applications, and described their most important associated techniques. In this section we discuss spatial and temporal organization in more detail. Since visual organizations need not be processed serially and can even enhance peoples' abilities to perceive multiple images at once, it seems clear that images may benefit from organizational strategies other than those that are appropriate for text.

4.1. Spiral Layout

The process of organizing objects on a two-dimensional surface is known as *2-D layout*. Layout algorithms have been developed for a variety of applications such as semiconductor chip design, circuit board design, and simplification of complex graph diagrams. We have found two-dimensional bin-packing algorithms, which attempt to minimize the total amount of surface area for a collection of non-overlapping rectangles, to be particularly useful for organizing image representations within a fixed rectangular area.[16] One problem with the bin-packing algorithms we have explored, however, is that they produce image clusters with the tallest images along one edge, and while this may indeed minimize the total surface area, the resulting effect seems visually arbitrary and imbalanced.

We have developed a *spiral layout algorithm*, which tends to produce more visually balanced configurations of image collections. The spiral algorithm positions the highest rated images near the center of a configuration and the lowest rated images on the periphery. The resulting configurations allow people to view the highest rated images with a minimum of eye movement by directing their gaze toward the center of the spiral. Although the algorithm is general enough to use any rating scheme, we have found that the image-rating scheme described above (see Section 3.2) to be adequate for the types of image collections we have explored.

For some image collections, particularly those containing a large number of images, we have found it useful to use two iterations of the spiral layout algorithm, each with a different rating scheme, to provide visual groupings of logically related images. In Figure 1, for example, images from each bookmarked page are first positioned in a spiral cluster using the image rating scheme described in Section 3.2. Then each cluster is positioned in a larger spiral of clusters using a rating scheme based on the number of images in each cluster, such that the largest clusters appear near the center. The resulting configurations appear as large clusters for each category, which each contain smaller sub-clusters for each bookmarked page.

4.2. Informational Zoom

One problem with spatial organizations of images is that they can require a lot of screen space. Shrinking images helps, but sacrifices high-frequency spatial details within the images. Animation techniques, in which images are organized in time as well as space, may offer additional opportunities for allowing people to access multiple image representations. One potential problem with animating images, however, is that a moving image may be difficult to select and thus diminish access to its associated web page. Another potential problem with animating images is that they may be disorienting, which could offset other benefits.

We have developed a method for temporarily scaling (or *zooming*) images, which does not make images difficult to select, and seems to minimize disorientation. Zooming-in is a common technique in cinematography, where it is used to provide emphasis and additional detail. Unlike a direct cut from a wide shot to a close-up, a continuous zoom is usually not disorienting because it helps viewers maintain context as the wide shot gradually transforms into the close-up.

While zooming is a natural technique for navigating through multi-scale virtual environments, one problem is that it can be disorienting. After changing their viewpoint, people can have difficulty interpreting their new location within the overall context. This is highly dependent upon feedback received during the change of viewpoint. For example, in Pad++ we employ animation techniques such as *slow-in and slow-out* to help maintain orientation within the larger context and support object constancy.[17]

Figure 5. Three frames of an informational zoom animation. Note the appearance of the textual label.

We call our new interactive technique for temporarily scaling images *informational zooming*. Informational zooming allows people to maintain their frame of reference while temporarily obtaining additional detailed information. Informational zooming starts with a collection of images in a fixed spatial organization. When a person rolls their mouse cursor over an image, a copy of the image gradually dissolves on while growing larger. As the image grows, additional details appear. When the image reaches a particular size it gradually dissolves off. An example is depicted in Figure 5. Because the zooming image dissolves on and off, it only obscures the original configuration of images temporarily. Unlike zooming that is used for navigation, the original configuration of images does not move or change, so a user's context and frame of reference are not altered, and there is little or no sense of disorientation. The image can be selected at any time, even while it is zooming, so access to its associated web page is never compromised. We implement informational zooms as Flash movies. They could also be created as Java applets.

We have also extended the informational zoom technique. Instead of just causing a copy of the original image to grow while dissolving on and off, multiple additional images can also grow while dissolving on and off. We use these additional images to provide a preview of what a user will see if they make the current selection. For example, in a visual bookmark application, the first image to zoom is the representative image for the bookmarked page, while the additional zooming images are the remaining images on the bookmarked page. Similarly, when images are used to represent a hierarchy of categories, the first image to zoom is the representative image for a category, while the additional zooming images are the remaining images within the category.

5. DISCUSSION

User interface design for web-based information systems requires careful attention to tradeoffs. To minimize time spent looking for information, whether browsing or searching, representations must be chosen that people can use efficiently. Even though textual representations help to minimize computer response time and storage requirements, image representations may be more helpful for people who need to access and organize information. In this section we discuss a number of underlying tradeoffs that must be considered.

In comparing the cognitive properties of images with words, we find images may be interpreted more quickly, remembered more efficiently, and even improve our memory for words. While perhaps the most obvious disadvantage of using images to represent information is their inefficient digital representation, advances in image compression, digital storage, and network capacities are converging to minimize the space consumed by images in digital memory and the time consumed transmitting images over digital networks.

All interface representations occupy computer screen space. Minimizing the amount of space occupied by representations allows more representations to be visible and increases the likelihood of seeing patterns. On the other hand, when more representations are visible it is also more likely that the display will appear cluttered, making it difficult for a person to find task-relevant representations. To some extent, the properties of the representations can be chosen to help minimize visual clutter, make it easier to identify groups, and assist search. As we have described above, visual clutter can be further minimized with spatial arrangements that cluster related representations.

Another disadvantage of using images to represent information is that, in general, images take up more space on a display than words. Scaling images by making them smaller while preserving their aspect ratio is an obvious technique to reduce the spatial requirements of images. Scaling images not only conserves screen space, but also reduces storage and transmission requirements. Scaling images such that their maximum dimension is a common size also increases the range of possible spatial layout algorithms for organizing collections of images.

Image scaling, however, also involves trade-offs. One issue is that scaling images removes information. Images with high-frequency content may become illegible at small scales (although this is even more of a disadvantage for scaling textual representations, which encode most of their content in the high-frequency corners and edges of letters). In addition, the original size of an image may be meaningful as may the relationship between sizes of images in a collection. We have developed an informational zooming animation technique to temporarily restore spatial details that are lost in scaling. We have also developed an image rating scheme that preserves information about the original sizes of images. Perhaps a more insidious disadvantage of using images to represent information is that scaling large images can be slow. To minimize this disadvantage we have developed a very fast image scaling algorithm.

Scaling also has implications for image selection, which we have described as a required interaction technique for letting people access the information represented by an image. In particular, small images may be difficult to select. The images emphasized by our rating technique tend to have a large thumbnail area, which corresponds to an approximately square aspect ratio. In fact, square images may be easier to select than most textual representations, which tend to be short and wide. Because it may be difficult to select images that overlap, we focus on layout algorithms that minimize overlap. Because it may be difficult to select moving images, we focus on animation techniques that make images larger and, therefore, even easier to select.

A key issue that must be confronted by any content summarization technique is the effectiveness of the summaries. The main purpose of our image rating scheme is to identify images with the most likelihood of representing the content of a web page and, simultaneously, identifying images such as advertisements and decorations with the least likelihood of representing the content of a web page. The spiral layout algorithms we have developed position highly rated images near the center of an arrangement so people can see them in a single glance with a minimum of eye movement. By combining image scaling with image rating our layout algorithms use spatial arrangements to simplify perception and choice.

Rating schemes are also useful for choosing representative images automatically. Although people are usually better at choosing representative images than computers, rating schemes produce reasonable results and by automating this task they make possible classes of applications such as visual query results, where a single image must represent a web page, as well as visual bookmarks, cache visualizations, and other applications in which a single image must represent a category in a hierarchy.

Just as images can be organized in space, they can also be organized in time. An animated GIF image, for example, is a collection of images with an associated policy for displaying them in succession. While temporal organizations of images can take less screen space than spatial arrangements and can minimize eye movements, they may require more time for users to watch and wait for a relevant representation to appear. Once a relevant representation appears, it may be difficult to select it before it is obscured by an irrelevant representation.

Although we have explored several applications that use temporal organizations of images, the most promising ones combine animation with some form of spatial organization, particularly because non-overlapping images are easier to select. It should come as no surprise, however, that there are significant tradeoffs to consider when combining spatial and temporal organizations. Not only are moving images difficult to select, they can also be distracting and disorienting. Our informational zooming technique is a combination of spatial and temporal organizations, which is designed to minimize distraction, because it occurs only when a person moves their cursor over an image, and disorientation, because it is temporary, uses transparency, and does not change the underlying spatial arrangement.

Visual session summaries differ from the other applications we have explored in the potential to use spatial arrangements to encode temporal information. We have found a variant of our spiral layout algorithm to be useful. It uses recency of access as its rating function so that images from the most recently accessed pages appear in the center of each spatial cluster. Sub-sessions are depicted as spatial sub-clusters. We are exploring similar visualizations of activity structures. We use the Unix *strace* facility to record and timestamp system call histories and then parse that low-level stream into higher-level summaries of activities that we represent visually. We are evaluating heuristics to help us identify activities by their temporal structure. A cluster of repetitive shifts between a fixed set of applications might provide evidence of a higher-level activity involving those applications. We also use the amount of active time spent in such clusters as an index of the importance of an activity. We are exploring the creation of multi-scale visual summaries of such activities that could allow one to easily return to them and the associated context.

When designing systems that humans are meant to use, we argue that designers should choose representations that humans can remember, interpret, identify, and reason with efficiently, even if these representations are less efficient in terms of digital memory, processing, and transmission. The work we summarize in this paper suggests that spatial and temporal organization of selectable images may offer multiple advantages over textual lists of titles and URLs. Image representation techniques and organizational strategies promise to improve access to web information and change the way people reference, locate, and organize web content.

REFERENCES

1. Edwin L. Hutchins, *Cognition in the Wild*, MIT Press, Cambridge, 1996.
2. James D. Hollan, Edwin L. Hutchins, and David Kirsh, ``Distributed cognition: A new theoretical foundation for human-computer interaction research,'' *ACM Transactions on Human-Computer Interaction*, in press.
3. Andy Clark, *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Cambridge, 1997.
4. David Kirsh, ``The Intelligent Use of Space,'' *Artificial Intelligence* **73**, pp. 31-68, 1995.
5. David Kirsh and P. Maglio, ``Some epistemic benefits of action: Tetris, a case study,'' *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, 1992.
6. Roger Shepard, ``Recognition memory for words, sentences, and pictures,'' *Journal of Verbal Learning & Verbal Behavior*, **6**, pp. 156-163, 1967.
7. Lionel Standing, Jerry Conezio, and Ralph Norman Haber, ``Perception and memory for pictures: Single-trial learning of 2500 visual stimuli,'' *Psychonomic Science*, **19(10)** pp. 73-74, 1970.
8. Allan Paivio, *Imagery and Verbal Processes*, Holt, Rinehart, & Winston, 1971.
9. Michael Lesk, Dennis Egan, Dan Ketchum, and Carol Lochbaum, ``Better things for better chemistry through multi-media,'' the Eighth Annual Conference of the University of Waterloo Centre for the New Oxford English Dictionary and Text Research, Waterloo, 1992.
10. Ron Hightower, Laura Ring, Jonathan Helfman, Ben Bederson, and James Hollan, ``Graphical multiscale web histories: A study of PadPrints,'' *Hypertext '98 Proceedings*, pp. 58-65, 1998.
11. Michael Wynblatt and Dan Benson, ``Web page caricatures: Visual summaries for WWW documents,'' *IEEE International Conference on Multimedia Computing and Systems*, 1998.
12. Ken Turkowski, ``Filters for common resampling tasks,'' in Andrew S. Glassner, editor, *Graphics Gems*, pp. 147-165, Academic Press, 1990.
13. Peter J. Burt, ``Fast filter transforms for image processing,'' *Computer Graphics and Image Processing*, **16(1**) pp. 20-51, 1981.
14. Dale Schumacher, ``General filtered image rescaling,'' in David Kirk, editor, *Graphics Gems III*, pp. 8-16, Academic Press, 1994.
15. Jonathan Helfman, *Image Representations for Access and Similarity-Based Organization of Web Information*, p.83, PhD Dissertation, University of New Mexico, Computer Science Department, July, 1999.
16. F. Chung, M. Garey, and D. Johnson, ``On packing two-dimensional bins,'' *SIAM Journal on Algebraic and Discrete Methods*, **3(1)** pp. 66-76, March, 1982.
17. B. Bederson, J. Hollan, K. Perlin, J. Meyer, D. Bacon, and G. Furnas, ``Pad++: a zoomable graphical sketchpad for exploring alternate interface physics,'' *Journal of Visual languages and Computing*, **7(3)**, pp. 3-31, 1996.