

Exploring Web Search Results Using Coordinated Views

Orland Hoeber and Xue Dong Yang
Department of Computer Science
University of Regina
Regina, Saskatchewan, Canada S4S 0A2
{hoeber, yang}@cs.uregina.ca

Abstract

Most web search engines use a list-based representation of web search results, promoting a sequential evaluation of the document surrogates. Commonly, these search engines only display ten document surrogates per page, limiting the users' ability to explore the search results. We have developed two systems to support the visual exploration of web search results: HotMap and Concept Highlighter. In both of these systems, the search results are provided at two levels of detail: an overview map that provides a compact and abstract representation of the top 100 documents returned by the underlying search engine; and a detail window that shows 20 to 25 documents at a time. In this paper, we will discuss how these coordinated views support the visual exploration of web search results.

1 Introduction

A recent survey reported that over 85% of web searches are conducted using one of four primary web search engines [13]. All four of these search engines (Google [4], Yahoo [20], MSN [10], and AOL [2]) use a very similar web search interface, wherein the search results are presented to the user in a list-based representation containing ten document surrogates per page. Commonly, the document surrogates consist of the title of the document, a snippet showing the context of the query terms, and the URL of the document.

Two independent studies have shown that users often view only one to three pages of search results [14, 16]. While it is possible that many users are able to satisfy their information needs by viewing only 10 to 30 document surrogates, we believe this is evidence that users are unwilling to view many non-relevant documents before abandoning the search results.

One of the reasons users do not delve deeply into the set of search results is because the list-based representation does not readily support the users' relevance decision mak-

ing tasks. Document surrogates must be considered one-by-one in blocks of ten at a time, resulting in an inability to explore the search results easily. Commonly, users are required to consider many non-relevant documents in the search results before finding enough relevant documents to fulfill their information needs. To address this shortcoming of the list-based representations, we have developed two interfaces to support the visual exploration of web search results.

In HotMap [8], the frequency of each of the query terms from the users' queries are depicted visually using colour coding. This allows the users to easily identify "hot" documents based on the frequent appearance of the query terms within the document surrogates. In addition to this visual representation, the search results can be dynamically re-sorted based on the query term frequencies, supporting an interactive exploration of the search results.

In Concept Highlighter [9], a set of relevant concepts is generated from a concept knowledge base using the query terms; interactive concept-based fuzzy clustering is used to cluster the search results with respect to these concepts. As the users select the concepts that are relevant to their information need, the search results are re-sorted based on the fuzzy membership score of each document surrogate with respect to the selected concepts. Colour coding is used to visually represent the fuzzy membership scores, allowing the users to easily determine the degree to which each document surrogate belongs to the set of concepts that have been selected.

Both of these systems use the same framework for providing coordinated views of the search results at two levels of detail. The overview map depicts the top 100 document surrogates returned by the underlying search engine in a compact and abstract representation. The detail window shows 20 to 25 document surrogates at a time, allowing the users to determine the relevance of individual document surrogates to their information needs. These coordinated views provide the users with the ability to interactively explore the search results both from an overview perspective and a de-

tailed perspective. Within the same information display, the users can determine the general properties of the top 100 document surrogates, as well as the specific properties of individual document surrogates.

The remainder of this paper is organized as follows: An overview of search results representations is provided in Section 2. In Section 3, the two tools for the visual exploration of web search results are described. Section 4 discussed the coordinated views used in our two systems to represent web search results. Section 5 provides an overview of the subjective evaluations that were performed on these systems; Section 6 presents the findings from these studies. A discussion is provided in Section 7. Conclusions and future work are presented in Section 8.

2 Background

According to North & Shneiderman's taxonomy, the coordination between multiple views can be classified by the action performed as well as the data presented in the views [11]. Along the action dimension, selecting items in one view can cause the selection of items in other views; navigating the data in one view can cause the navigation of data in other views; and selecting items in one view can cause changes to the navigation of data in other views. Along the data dimension, the multiple views can present data from the same collection, or from different collections. We use this taxonomy to classify the literature relevant to the display and exploration of web search results.

2.1 Textual Representations of Web Search Results

While the supplemental information provided by the four primary web search engines (Google [4], Yahoo [20], MSN [10], and AOL [2]) differs slightly, all use a textual representation of the web search results, displaying a list of ten document surrogates per page. Within these lists, each document surrogate primarily consists of the title of the document, a snippet showing the context of the query terms, and the URL of the document.

Since the documents themselves are primarily textual, this textual representation of web search results is a logical method for displaying this information. However, since these lists are static, they provide a limited ability to manipulate the search results; they only lend themselves to a sequential evaluation of ten document surrogates at a time. This textual interface style requires the users to read the information (which can be time consuming), or scan the information (which can be error prone). Further, since text can consume a large amount of space in an interface, the list based representations are commonly limited to ten document surrogates per page.

None of the four primary web search engines provide additional views of the web search results. The only information provided to the users about the search results set are the ten document surrogates listed on the page currently being viewed. These systems do not allow the users to generate an overview of the search results, view the search results at different levels of detail, or provide support for an exploration of the search results. Since multiple views are not present, the list-based representations of web search results cannot be classified according to the multiple view coordination taxonomy.

2.2 Representing Clustered Search Results

Recently, a number of publicly available web search systems have been developed that cluster the web search results, allowing the users to browse the clusters as they explore the search results. Many of these systems use hierarchical clustering algorithms, and primarily differ in the representation of the clusters and the methods for interaction with the clusters. Two such systems are Vivisimo [17] and Grokker [6].

In Vivisimo, the hierarchical clusters are represented as a tree. The nodes in the tree can be expanded and collapsed in a manner similar to file directory navigation. When a tree node is selected, the document surrogates contained within that cluster are displayed in a separate frame using a list-based representation similar to that of the four primary search engines.

In addition to providing a tree-like navigation scheme similar to Vivisimo, Grokker also provides a visual representation of the hierarchical cluster structure. This visual representation uses nested circles to represent the clusters and their children, and provides the ability for the users to see the sizes of the clusters and whether they contain additional children clusters or document surrogates. Like Vivisimo, when a cluster is selected, the document surrogates that are contained within that cluster are displayed in a separate frame using a list-based representation.

In both of these systems, two views of the search results are provided: a cluster view, and a textual list-based view. As the users make selections within the cluster view, the list-based view is updated to display the list of search results contained within the selected cluster. While these coordinated views do support an exploration of the search results with respect to cluster selection, this exploration is limited to a narrowing down of the search results based on the selected clusters.

According to the action dimension in the multiple view coordination taxonomy, both of these systems use item selections in one view (i.e., cluster selection) to cause changes in the navigation of data in another view (i.e., the corre-

sponding list of document surrogates). Along the data dimension, these multiple views present data from different collections (i.e., the cluster tree, and the document surrogate list).

2.3 Visual Representations of Web Search Results

While there are many systems that generate visual representations of information retrieval results, there are relatively few that support the visual exploration of web search results. Two systems that provide visual representations of web search results and multiple views of the data are xFind [1] and the work by Roberts et al. [12].

xFind provides three different interfaces to a custom web document indexing system: a simple list-based representation; a customizable scatterplot representation; and a vector-based spatial clustering representation. While these representations of the search results take advantage of the extra information that is available through their indexing system, this information is generally not available with other search engines.

In the scatterplot representation, both an overview and a zoomed-in view of the plot are provided, allowing the user to simultaneously view context and detail, and remain oriented within this visual representation of the search results. In the spatial clustering representation, a secondary frame provides a text-based list of clusters which allow the users to highlight the relevant documents within the visual cluster representation.

Although the multiple views in these two visual representations support the exploration of the web search results, the spatial layout of the two visual representations result in document surrogates being mapped to points in the two-dimensional display. This makes it difficult to view the additional information present in the document surrogates, or to make direct comparisons among multiple document surrogates.

Within the action dimension in the multiple view coordination taxonomy, the scatterplot and the spatial clustering representations are classified differently. The scatterplot representation uses navigation in one view to cause changes in the navigation of data in another view; both views present data from the same collection. The spatial clustering representation uses item selection in one view to cause the selection of items in another view; the views present data from different collections.

In the work by Roberts et al. [12], web search results are presented in three coordinated representations. In a scatterplot representation, the domains of the document URLs are represented using glyphs, with the number of external and internal links mapped to the x and y position in the plot. This representation allows the users to identify where the

document is coming from (i.e., .com, .net, .ca, etc.) and the different types of links in the document.

A second glyph-based representation uses a “quartile” consisting of concentric boxes divided into quarters. This is used to represent the external (inside box) and internal (outside box) links. These glyphs are provided in a horizontal list that wraps like text.

The third representation provides a simple list of the URLs of the documents in the search result. The coordination among these views is provided such that a selection made in one representation highlights the same items in the other representation.

The primary information represented in this interface are the number of external and internal links on each of the documents, and information about the URL of the document. While this information may be of value to a searcher seeking documents with a high number of internal and external links (i.e., a “jumping off list”), there is little indication that this is of value to general web searchers. Not even the document titles or snippets are readily available.

According to the action dimension in the multiple view coordination taxonomy, this system uses item selections in one view to cause the selection of items in another view (i.e., the corresponding glyphs or URLs in the other views). Along the data dimension, this system presents data from the same collection in the multiple views.

3 Visual Exploration of Web Search Results

Our motivation for investigating visual representations of web search results came from Wise et al., who noted that “the need to read and assess large amounts of text that is retrieved through even the most efficient means puts a severe upper limit on the amount of text information that can be processed by any analyst for any purpose” [19]. We have found that this upper limit can be addressed by applying information visualization techniques to the presentation of the search results, and providing interaction features that promote the exploration of the search results.

Information visualization takes advantage of the human visual information processing systems by generating graphical representations of data or concepts [18]. The cognitive activity involved in viewing and processing a visual representation allows users to gain understanding or insight into the underlying data. With respect to the visualization of search results, the ultimate goal is to allow users to *see* the information without having to *read* the information.

Even with a visual representation of the information, Spence noted that a “mere rearrangement of how the data is displayed can lead to a surprising degree of additional insight into that data” [15]. This rearrangement of the data can be especially valuable for web search results, allowing

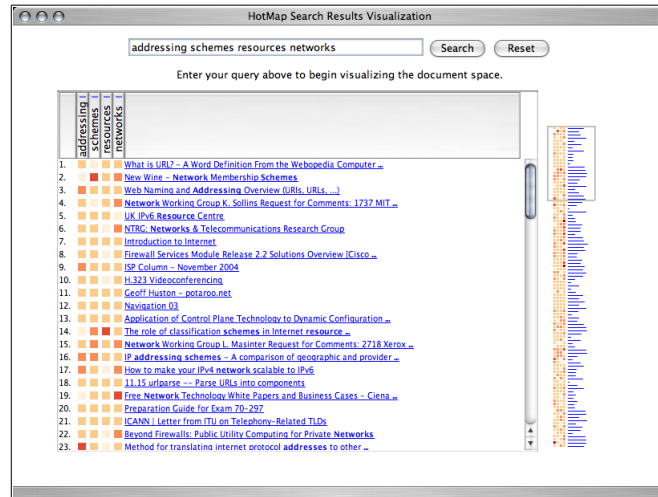


Figure 1. A sample web search using HotMap. Note the search results include the query term frequencies displayed at two levels of detail in the overview map and the detail window.

the users to re-sort the search results based on the properties they think are important to their information needs.

In the remainder of this section, we describe the two systems we have developed to support the visual and interactive exploration of web search results.

3.1 HotMap

HotMap is a meta-search system that retrieves the top search results returned by the Google API [5] for a given user query, and presents these results in a compact visual manner that supports both visual information processing and user-directed exploration. Our inspiration for this work is based on a common method for evaluating the relevance of a document surrogate: identifying which of the query terms appear in the title and snippet. HotMap represents the query term frequencies in the form of a colour-code on a heat scale. Multiple occurrences of a query term result in a dark red colour; fewer occurrences are represented by progressively lighter shades of red and orange. As shown in Figure 1, this colour coding allows the users to see the “hot” documents easily (and provides the inspiration for naming our system *HotMap*).

In HotMap, the search results are presented in a grid-based layout at two levels of detail: an overview map provides a compact representation of the top 100 search results, and a detail window provides a focused view of approximately 20 documents at a time. The details of these coordinated views of the web search results are provided in Section 4.

In order to promote the exploration of the search results,

the users can interactively sort and re-sort the search results based on the query term frequencies. In addition, nested sorting allows the users to sort based on an ordered selection of their query terms. These sorting features are easily activated within the visual interface by clicking on the column headers; the results of sorting operations are instantly reflected in both the overview map and the detail window.

The visual representation of the query term frequencies in the detail window is visually similar to the TileBars work by Hearst [7]. HotMap, however, extends this tile mosaic metaphor to represent web search results in near real-time. It also adds features for viewing the search results at different levels of detail, and supports the users as they interactively explore the search results.

According to the action dimension in the multiple view coordination taxonomy, HotMap uses navigation of the data in one view to cause an equivalent navigation of the data in the other view. Along the data dimension, this system presents data from the same collection in the multiple views.

The specific details of the design and implementation of the HotMap system, along with user study results, are provided in [8]. A more complete study measuring the information retrieval performance, user performance, and efficiency is currently underway.

3.2 Concept Highlighter

Concept Highlighter is a meta-search system that presents the search results returned by the Google API [5] in an interactive fuzzy clustering interface. The system automatically matches the users’ queries to a concept knowledge

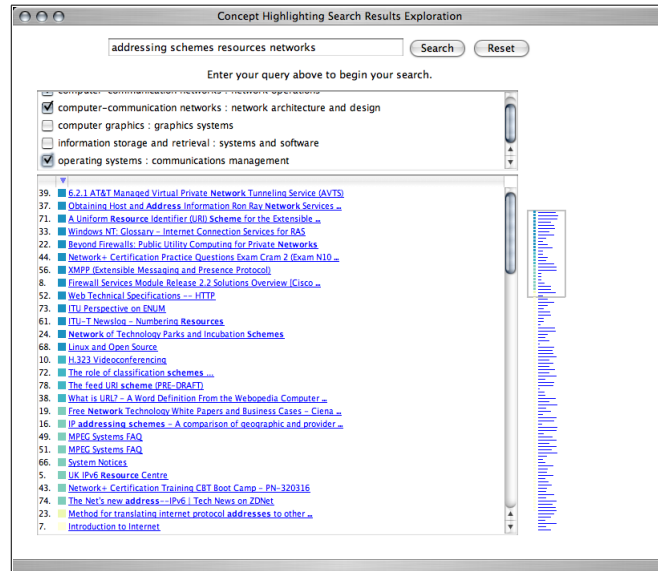


Figure 2. A sample web search using Concept Highlighter. Note the search results include fuzzy membership scores displayed at two levels of detail in the overview map and the detail window.

base, resulting in a set of potentially relevant concepts. The concept knowledge base is used to select additional terms that have been used to describe these concepts. All the concepts as well as the top 100 document surrogates in the search results are represented using a vector-space model; a single-pass fuzzy clustering algorithm assigns fuzzy membership scores to each concept-document surrogate pair.

The Concept Highlighter interface resembles that of HotMap, with the addition of a list concepts provided at the top of the interface (see Figure 2). The detail window provides a focused view of approximately 25 document surrogates at a time, while the overview map provides a compact and abstract representation of the first 100 document surrogates returned by the Google API.

As the users select the concepts that are relevant to their information needs, the search results are automatically re-sorted based on the fuzzy membership scores. The selection of multiple concepts results in a merging of the clusters and the addition of the fuzzy membership scores. As a result, the document surrogates that are most similar to the selected concepts are moved to a more prominent location in the search results.

In addition to the re-sorting of the search results, colour coding is used to indicate the fuzzy clustering score with respect to the selected concepts. This allows the users to interpret the relative similarity of the document surrogates to the selected concepts.

This system can be classified in two different ways according to the multiple view coordination taxonomy. With

respect to the concept selection, selecting items in one view causes changes to the navigation of the data in the other view (i.e., the re-sorting of the data in the detail window and the overview map). These views use data from different collections. With respect to the coordination between the detail window and overview map, the classification is the same as for HotMap: navigating the data in one view causes an equivalent navigation of the data in the other view. Along the data dimension, these views present data from the same collection.

The specific details of the design and implementation of Concept Highlighter are provided in our previous work [9]. Studies measuring the information retrieval performance, user performance, and efficiency are in progress.

4 Coordinated Views of Web Search Results

Both HotMap and Concept Highlighter use a similar framework for providing multiple views of web search results. Commonly, there is a design tradeoff between the utility provided by the multiple views and the cognitive overhead in learning and using the system [3]. In designing these coordinated views, we have attempted to minimize this cognitive overhead while providing meaningful information for the exploration of the web search results at both a overview level and a zoomed-in level. The details of these coordinated views, as well as the methods for maintaining coordination between the views is provided in the remainder of this section.

4.1 Detail Window

In the detail window, a compact representation of 20 to 25 document surrogates is provided. Of the information that is present in the document surrogate, only the title of the document is displayed persistently; tool tips are used to hide and dynamically show the additional details available in the document surrogate, such as the snippet and the URL.

The users can click the mouse button on any title in the detail window to access the corresponding document. Doing so opens a new browser window and loads the selected document. In order to indicate to the user that this is an available option, the document titles are underlined and coloured blue. After a title has been clicked and the document has been opened, the colour of this title is turned to purple to indicate that the document has been viewed. The underlining and colour selection follow the defacto standard for new and visited links in a web page.

In addition to representing the document surrogate information provided by the Google API, the detail window also displays the additional information provided by HotMap and Concept Highlighter in a visual manner. In terms of initiating an exploration of the web search results, the detail window provides access to the primary interaction features of these systems.

In HotMap, the detail window provides a column for each of the terms in the users' queries. A colour code is provided in the grid to represent the frequency of the corresponding query term (column) within the corresponding document surrogate (row). Clicking on the column header for a term will re-sort the search results according to the frequency of the corresponding term. Holding down the control key while clicking the column headers generates a nested sorting of the search results according to the ordered selection of the query terms.

In Concept Highlighter, the detail window provides a single column to indicate the fuzzy membership score of each document surrogate with respect to the selected concepts. As the users select concepts that are relevant to their information needs, the colour code in this column is updated to represent the fuzzy membership scores. In addition, the search results are automatically re-sorted based on these scores, bringing the document surrogates that are most similar to the selected concepts to the top of the list. While the default is to re-sort the search results, the users can turn this sorting off and return the search results to the order provided by the Google API. However, the colour code representing the fuzzy membership score remains, allowing the users to visually determine which document surrogates belong to the selected cluster.

4.2 Overview Map

The primary goal in the design of the overview map was to provide an indication of the features of ten pages worth of search results (i.e., 100 document surrogates) in a single compact representation. This was challenging since even for a high resolution display, there is a relatively small number of pixels available to represent this information. Therefore, an abstract representation of the web search results is necessary.

While many of the standard display resolutions have more horizontal space than vertical space, we chose to display the overview map in a vertical orientation. Since the search results are represented in a vertical layout in the detail window, coordinating the orientation allows the users to more easily make visual connections between the two representations.

Given the small amount of vertical pixels at our disposal, and the large amount of search results to represent, we chose to use only three vertical pixels to represent each document surrogate, with a one pixel gap between adjacent objects. The result is a compact representation that consumes only 400 vertical pixels of screen space in the representation of 100 web search results.

Even though a small number of pixels are used to represent each document surrogate, it is still important to convey information about the features of the search results. For HotMap, the overview map provides a three by three pixel block to visually represent the query term frequency for each term in the users' query. For Concept Highlighter, the overview map provides a single three by three pixel block to represent the fuzzy membership score for each document surrogate with respect to the selected concepts.

In order to make it easier for the users to make a visual connection between the overview map and the detail window, lines were drawn in the overview map whose lengths are relative to the lengths of the document surrogate titles. These lines use same colour as the titles in the detail window. When the link colours change in the detail window as a result of the users selecting to view specific documents, a corresponding colour change is made in the overview map for the selected document surrogates.

The end result is that the overview map provides a "zoomed out" view of the search results. The colour coding of the frequency or clustering information, the length of the title line, and the link colour of the title line all provide landmarks allowing the user to recognize the features of the search results and remain oriented within these visual representations. This consistency between the overview map and the detail window makes comparisons between the views easier to perform [3]. See Figure 3 for a magnified view of the overview maps used in HotMap and Concept Highlighter.

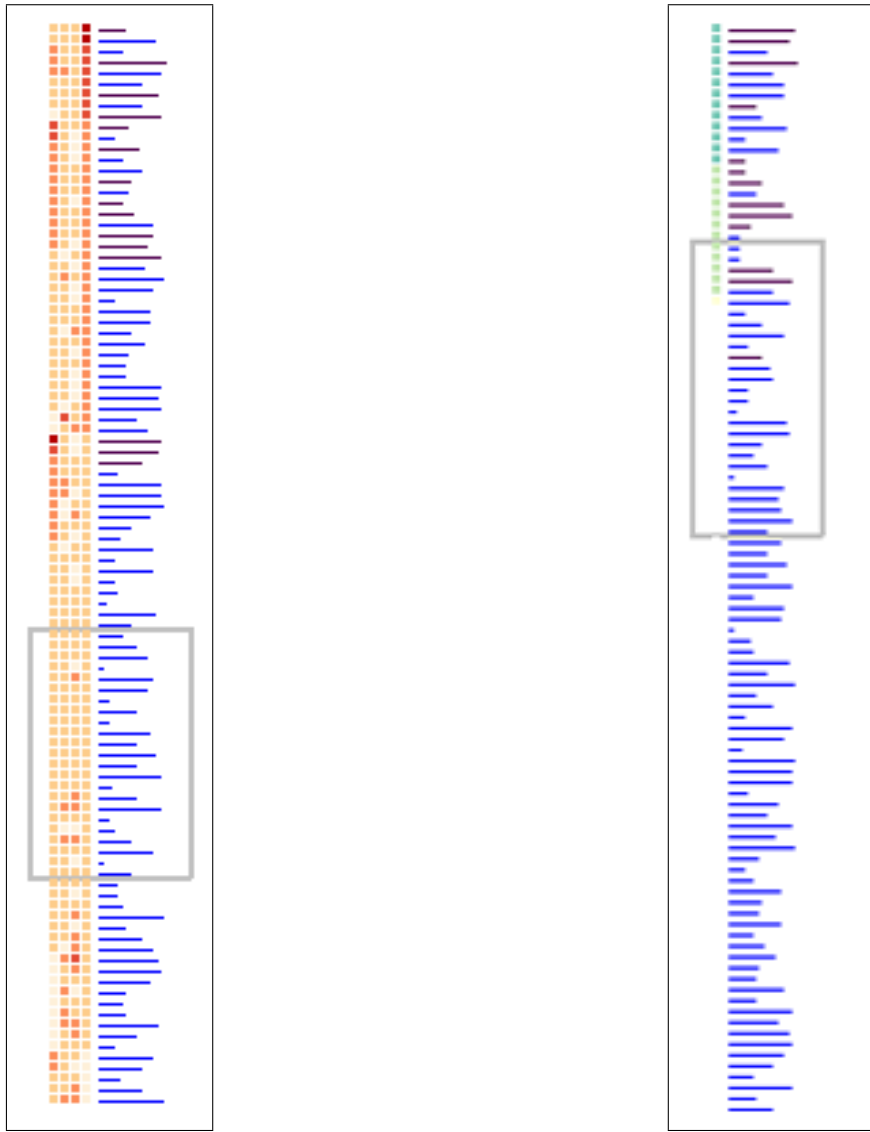


Figure 3. Magnified images of the overview maps used in HotMap (left) and Concept Highlighter (right). In HotMap, the colour coding is used to signify the frequency of the query terms within the search results. In Concept Highlighter, the colour coding is used to signify the fuzzy membership score of the document surrogates with respect to the selected concepts. In both overview maps, the blue lines represent documents that have not yet been viewed; purple lines represent documents that have been viewed. The scroll box provides a perceptual cue as to the location of the document surrogates that are currently being displayed in the detail window. These abstract representations of the first 100 documents returned by the search engine support the visual exploration of the web search results.

4.3 Coordination of Views

When multiple views of the same data at different levels of detail are provided, as in these systems, it is very important to ensure that these views remain coordinated, and that this coordination is self-evident. Synchronized scrolling provides the primary means for coordinating the detail window and the overview map. In addition, perceptual cues are provided to make the relationships between these views apparent to the user [3].

In order to aid the user in remaining aware of the coordination between the detail window and the overview map, a scroll box is placed in the overview map to indicate which document surrogates are currently being viewed in the detail window (see Figure 3). As the users scroll in the detail window, the location of the scroll box is automatically updated.

To support the users' task of visually exploring the search results within the overview map, clicking or dragging the mouse to any location in the overview map automatically relocates or "jumps" the scroll box to that location and scrolls the detail window to display the corresponding document surrogates. This ability to "jump" to locations of interest is especially valuable in HotMap, where "hot" documents can be easily identified deep within the search results. It is also useful in Concept Highlighter, especially when the users wish to see the search results in the order provided by Google at the same time as the visual representation of the fuzzy membership scores.

Having the line length and colour in the overview map coordinated with the document title length and colour in the detail window provides visual landmarks to the user. This, together with the scroll box, allow the users to see the relationship between the document surrogates they are viewing in the detail window and the top 100 document surrogates returned by the Google API, and remain oriented within the visual representations as they explore the web search results.

5 Subjective Evaluations

A user evaluation was conducted using 10 computer science graduate students over two task sessions. In the first task session, the participants were assigned specific information seeking tasks to perform using both a Google-like interface and the original prototype of HotMap (which included some of the features of Concept Highlighter). In the second task session, the participants were asked to perform a search on a topic in which they are knowledgeable using a Google-like interface, a refined version of HotMap (see Figure 1), and Concept Highlighter (see Figure 2).

All the participants identified themselves as frequent computer users (i.e., more than 10 times per week), and

highly experienced with computers. 70% of the participants indicated that they perform more than 10 web searches per week; 20% indicated that they perform 5-10 web searches per week; and 10% indicated that they perform 1-5 web searches per week. 70% of the participants identified themselves as having a high level of experience with web searching; 30% identified themselves as having a moderate level of experience with web searching.

Clearly, the participants in this study are much more experienced with computers and with web searching than the general public. While the conclusions we can make from this study may not generalize to all web searchers, they will give us insight into how experts can benefit from the visual exploration of web search results. These insights will provide a basis for evaluating our system with a more diverse user population.

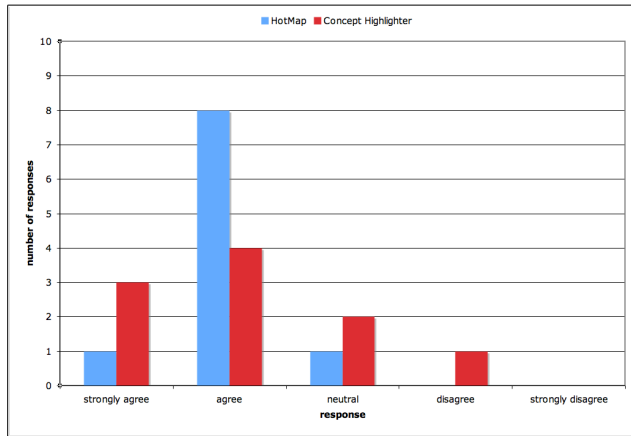
6 Results

After the first set of tasks were performed, the participants were asked to rate their preferences when using a web search tool. 70% of the participants indicated that they had a preference for a compact representation of the web search results; 10% were neutral; and only 10% indicated that they did not prefer a compact representation. This supported our hypothesis that many users would prefer a compact representation of the search results from which they can access the snippet and URL of the document via a tool tip. As such, when the original prototype of HotMap was split into two separate systems (HotMap and Concept Highlighter), this compact representation was retained for both systems.

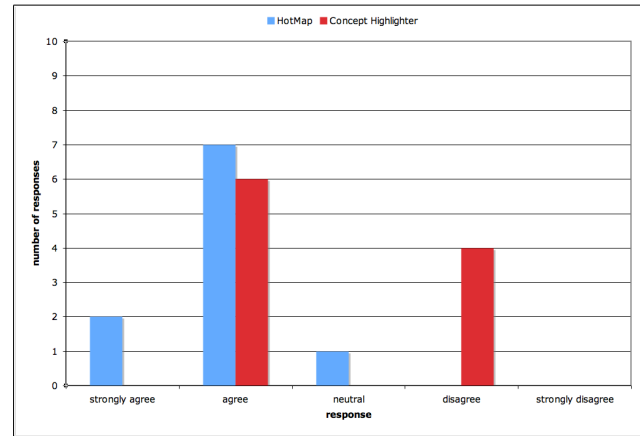
After completing the search tasks in the second usage session, a survey was administered to measure the subjective reactions of the participants to the features of the systems. The four measurements relevant to the coordinated views in these systems were the ease of use of the overview map, the usefulness of the methods for allowing the users to maintain coordination between the two views, the usefulness of the ability to jump to specific locations in the overview map, and the usefulness of the colour coding.

In terms of the ease of use of the overview map, all the participants agreed or strongly agreed that this feature was easy to use in HotMap. In Concept Highlighter, the responses were more varied, but skewed towards agreement (see Figure 4a). The differences in these responses may be a result of the differences in the amount of information displayed in the overview map between these systems. In HotMap, all of the query term frequencies are displayed in the overview map as soon as the search results are loaded, resulting in a very colourful representation of the search results which draws the users attention to this feature.

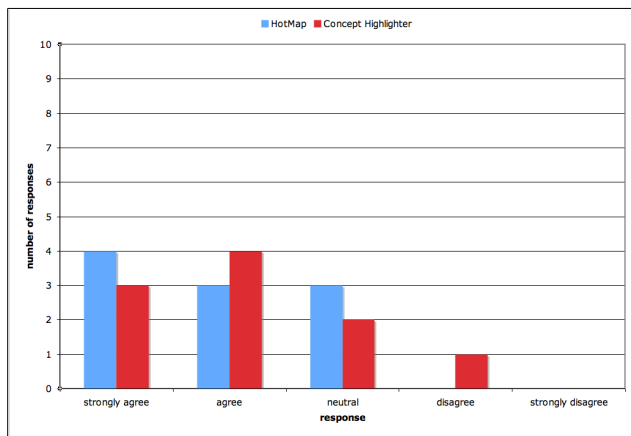
Most participants agreed or strongly agreed that the methods for indicating the coordination between the



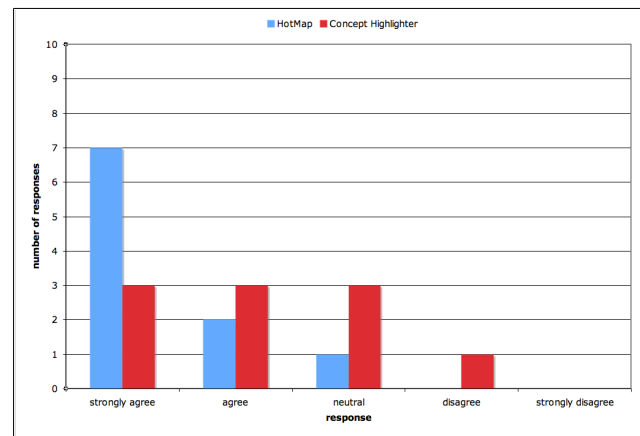
(a) ease of use of overview map



(b) usefulness of coordination between views



(c) usefulness of jump feature



(d) usefulness of colour-coding of data

Figure 4. These histograms show the responses to the subjective evaluation questions regarding the ease of use and usefulness of the coordinated views in HotMap and Concept Highlighter. Clearly, the coordinated views are more useful in HotMap than in Concept Highlighter.

overview map and the detail window in HotMap were useful. However, there was a near even split between agreeing and disagreeing that this feature was useful in Concept Highlighter (see Figure 4b). The disagreements with this statement among the Concept Highlighter users can be attributed to the fact that a number of these users did not consider the overview map during their task; they simply selected the relevant concepts and evaluated the re-ordered search results one-by-one.

All the participants were either neutral, agreed, or strongly agreed that the ability to jump to regions of interest in the search results was a useful feature in HotMap. While there was a tendency to agree that this feature was useful in Concept Highlighter, some participants did not find this feature useful at all (see Figure 4c). The more positive responses in HotMap were a result of being able to visually identify documents of interest easily, and jump to those lo-

cations. Within Concept Highlighter, this feature was less useful since by default, the search results were re-ordered to bring the more relevant documents to the top.

While most participants strongly agreed that the colour coding in HotMap was useful, the responses were over a much broader range for Concept Highlighter (see Figure 4d). This can be attributed to the more subtle colours used in Concept Highlighter, as well as the fact that no colour is displayed until a sufficient number of concepts have been selected.

In general, these responses were very positive for HotMap; for Concept Highlighter, the results were somewhat varied but generally positive. We conclude that the coordinated views which allow the user to view the search results at both an abstract level and a detailed level is a very valuable feature in support of the visual exploration of web search results in HotMap. Since the Concept Highlighter

system re-sorts the search results based on the fuzzy membership score by default, the usefulness of these coordinated views proved to be valuable, but to a lesser extent.

7 Discussion

We believe the ability to view the results of a web search both at a detailed level and an overview level can be very valuable when exploring the results of a web search. The overview map allows the users to determine the characteristics of a large number of documents in a single compact and abstract representation. At the same time, the detail window supports the primary task of deciding the relevance of specific document surrogates, and selecting to view a document.

Support for the visual exploration of the web search results is provided both via the tools for interactively re-sorting the search results based on the query term frequency (in HotMap) and the concept selection (in Concept Highlighter), and via the ability to “jump” to locations of interest in the overview map. This ability to quickly navigate the search results via a mouse click in the overview map is especially powerful in HotMap, where users can easily identify “hot” documents in the search results.

Whenever two views of the data are provided at different levels of detail, care must be taken to ensure the users remain oriented within these representations. Early versions of HotMap used a single vertical bar in the overview map to indicate which documents were being displayed in the detail window. Preliminary studies showed that this indicator was too subtle to ensure proper awareness of the search results space. As such, the scroll box was added to make the coordination between these views more apparent.

We believe that these systems can be very effective when the results of a web search are somewhat ambiguous, or when a user is not very familiar with the topic of their web search. In these cases, the ability to explore the web search results is very valuable. Multiple views of the web search results support this exploration by providing both an overview of a large set of search results, and a detailed display that supports the primary task of document surrogate evaluation and document selection. As the exploration takes place, these multiple views indicate to the user the effects of re-sorting the search results both at a local level and at a global level.

8 Conclusions and Future Work

The list-based representation provided by the four primary web search engines supports the linear evaluation of the web search results, but not an exploration of the web search results. To address this shortcoming, we have devel-

oped two systems to support the visual exploration of web search results: HotMap and Concept Highlighter.

These systems use a common framework for coordinated views to represent the search results both at a detailed level, displaying 20 to 25 document surrogates at a time, and at an overview level, displaying 100 document surrogates in an abstract representation. Support for a visual exploration of the web search results is provided through the visual inspection of the overview map, re-sorting the search results based on query term frequency (in HotMap), and re-sorting the search results based on concept selection (in Concept Highlighter).

Subjective evaluations collected during a user study with expert web searchers indicate that the overview map is an easy to use representation of the web search results at an abstract level. In addition, the coordination between the views, the ability to jump to locations of interest, and the colour coding of the data were all indicated as useful features by most participants. The responses were more positive for HotMap than Concept Highlighter, leading us to conclude that these coordinated views are of more value when there is sufficient useful data to display in both views.

A more complete user study with a larger and more diverse participant population is currently underway. Future work includes the integration of these tools into a single interface for the visual exploration of web search results, and the integration of this work into our larger research project of developing a complete framework for a visual and interactive web information retrieval support system.

References

- [1] K. Andrews, C. Gutl, J. Moser, and V. Sabol. Search result visualization with xFind. In *Proceedings of the Second International Workshop on User Interfaces to Data Intensive Systems*, 2001.
- [2] AOL. AOL search engine. <http://www.aol.com/>.
- [3] M. Q. W. Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the ACM Advanced Visual Interfaces*, 2000.
- [4] Google. Google search engine. <http://www.google.com/>.
- [5] Google. Google web API. www.google.com/apis/, 2005.
- [6] Grokker. Grokker search engine. <http://www.grokker.com/>.
- [7] M. Hearst. TileBars: Visualization of term distribution information in full text information access. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1995.
- [8] O. Hoerber and X.-D. Yang. The visual exploration of web search results using HotMap. In *Proceedings of the International Conference on Information Visualization*, 2006.
- [9] O. Hoerber and X.-D. Yang. Visually exploring concept-based fuzzy clusters in web search results. In *Proceedings of the Atlantic Web Intelligence Conference*, 2006.
- [10] MSN. MSN search engine. <http://www.msn.com/>.

- [11] C. North and B. Shneiderman. A taxonomy of multiple window coordinations. Technical Report Dept of Computer Science Technical Report CS-TR-3854, University of Maryland, College Park, 1997.
- [12] J. Roberts, N. Boukhelifa, and P. Rodgers. Multiform glyph based web search results visualization. In *Proceedings of the International Conference on Information Visualization*, 2002.
- [13] SearchEngineWatch. Nielsen netratings search engine ratings. <http://searchenginewatch.com/reports/article.php/2156451>, January 2006.
- [14] C. Silverstein, M. Henzinger, H. Marais, and M. Moricz. Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1), 1999.
- [15] R. Spence. *Information Visualization*. ACM Press, 2001.
- [16] A. Spink, D. Wolfram, B. J. Jansen, and T. Saracevic. Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 2001.
- [17] Vivisimo. Vivisimo search engine. <http://www.vivisimo.com/>.
- [18] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2004.
- [19] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In *Proceedings of IEEE Information Visualization*, 1995.
- [20] Yahoo. Yahoo! search engine. <http://www.yahoo.com/>.