Opportunistic Search with Semantic Fisheye Views

EPFL Technical Report: IC/2004/42

Paul Janecek*

Pearl Pu†

School of Computer and Communication Sciences
Swiss Federal Institute of Technology, Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

Figure 1: Reducing visual complexity with semantic fisheye view techniques.

ABSTRACT

Search goals are often too complex or poorly defined to be solved in a single query. While refining their search goals, users are likely to apply a variety of strategies, such as searching for more general or more specific concepts in reaction to the information and structures they encounter in the results. This is called opportunistic search. In this paper we describe how semantic fisheye views (SFEV) can be designed to effectively support this search process by enabling rapid, interactive exploration of the multiple contexts that are useful for different opportunistic search strategies. Similar to other focus + context techniques, SFEVs visually emphasize and increase the detail of information related to the focus and de-emphasize or filter less important information. The contribution of the SFEV approach is the flexible definition of context as a combination of interest metrics, which can be reconfigured and combined to support a wide range of information visualizations and leading to the discovery of diverse new search goals.

To further characterize the effectiveness of this technique for opportunistic search, we have developed a visual information retrieval interface for a large collection of annotated images that implements two distinctly different SFEVs; the first one uses similarity metrics to guide exploration over the images and keywords in the collection, and the second uses metrics that derive conceptual distance from an external general semantic model, WordNet. The results of a formal user experiment suggest that semantic-guided search is significantly more effective than similarity based search for complex opportunistic search and sensemaking tasks.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: User Interfaces—Theory and methods, graphical user interfaces (GUI), screen design;

Keywords: opportunistic search, semantic fisheye views, focus + context, information visualization

1 INTRODUCTION

In physical collections of documents and images, the high cost in time and effort to access information limits the amount of information that a user can process. The exponential growth of digital information and technology over the last few decades has significantly changed the cost-structure of interacting with information. People can now rapidly access vast collections of documents and multimedia with online search tools and digital libraries. However, when search goals are complex or uncertain, accessing information is only a small part of the search problem. In many cases, search is iterative, interactive, and opportunistic; analyzing the results of a query often leads to the discovery of unfamiliar vocabulary and relationships that guide the future direction of search. Opportunistic search is still a time-consuming and poorly-supported process in modern Information Retrieval (IR) interfaces.

Visual IR Interfaces (VIRI) use information visualization techniques to reveal relationships between documents and allow a user to rapidly shift between search and browsing tasks [15]. The tight coupling between visualization and interaction make VIRIs powerful tools for discovering global relationships between documents. However, there are several properties of visual representations that limit their effectiveness for opportunistic search. First, as the amount and complexity of information accessible in a VIRI grows, it is increasingly difficult to represent all of it in the limited space of a display. A visualization is often a compromise between showing a small amount of information in detail, or a large amount of information abstractly. Second, within a dense display of information, it is often difficult for users to see or visually distinguish the subset that is immediately useful for their current task. Selecting or navigating to each item to view detailed information incurs a cost in time and cognitive effort. Third, any single visual representation optimally supports only a small set of tasks [9, 34]. Supporting a diverse set of tasks, such as comparing content, structural and tem-

*e-mail: paul.janecek@epfl.ch

†e-mail: pearl.pu@epfl.ch

poral relationships between documents, will often require multiple representations.

Semantic fisheye views (SFEVs) are interactive focus + context visualization techniques that are designed to address these limitations by monitoring the relative importance of information for the user's current task [14]. SFEVs emphasize or increase the detail of the most important information, and de-emphasize or filter less important information [21]. This is useful for both selecting the information to represent in the constraints of a single display, and reducing the visual complexity within a dense visualization. Furthermore, by monitoring relative importance in multiple contexts, these techniques can interactively reveal semantic or structural relationships that may not be visible in the original representation.

In this paper, we describe a prototype that uses SFEV techniques to support a number of different opportunistic search strategies within a large collection of professionally annotated images. The first technique emphasizes images and keywords that are similar in content to the focus, and the second technique emphasizes information that is conceptually related to the focus based on WordNet, a general lexical ontology of the English language. These different SFEVs correspond to classic search strategies used for opportunistic search over heterogeneous collections of information [1, 2]. The prototype demonstrates how SFEVs can be used to support a variety of search strategies, and the results of an initial experiment suggests that semantic-guided search is more effective than similarity-guided search for sensemaking tasks.

In the following section we describe opportunistic search strategies in greater detail, and discuss how WordNet can be used to guide query generalization and specialization strategies. We then discuss our general framework for designing semantic fisheye views. This framework is independent of any particular representation, and coordinates the multiple interest metrics and emphasis techniques that create the fisheye view effect. We then describe in greater detail the implementation of similarity and semantic-guided search using semantic fisheye views.

2 RELATED WORK

This work is based primarily on three overlapping areas of research: effective search behavior, VIRI, and focus + context interactive visualization techniques.

2.1 Opportunistic Search

The search process is traditionally modeled in IR as a matching function that calculates similarity between a representation of the information in the collection and a representation of the users information need [18]. This simplistic view of search does not take into account the multiple contexts in which information can be categorized and used. Bates, on the other hand, describes search as an interactive process that evolves in response to the information found. Furthermore, the results of a search are not limited to documents, but also include the knowledge accumulated during the search process [2]. This model of search is different from similarity-based search in both the diversity of the information collected, and its evolving, opportunistic nature.

Bates identifies a number of different strategies that people use during this process, such as following the relationships between documents (e.g., footnotes or citations), or browsing over the structure of a collection (e.g., journal run or area scanning) [1, 2]. Bates suggests that IR interfaces would be more effective for this type of search if they more directly support these search strategies at a higher level [3]. Whereas Bates derives her recommendations from a behavioral approach to IR, Ingwersen comes to similar conclusions from a cognitive approach. He proposes representing information and measuring relevance in a variety of overlapping, inten-

tionally redundant ways to support the different needs and cognitive structures of the user [19].

Analyzing and making sense of information is an integral part of the search process; in some cases, such as research, sensemaking is the goal. Sensemaking is the process of developing mental or externally modeled schemas that allow a user to compare, categorize, and efficiently access information [32]. Building and manipulating these schemas is a combination of integrating structural and semantic relationships available in the information or the user's existing knowledge (top-down), and constructing new schemas from disconnected information encountered in the collection (bottom-up). Once constructed, these schemas are used to "... prescribe what to look for in the data, what questions to ask, and how the answers are to be organized [32]."

There are several implications of this previous research for interfaces to support opportunistic search. First, search is a much more complex activity than simply describing and finding a document. Search evolves as a user encounters information, analyzes it, and gathers the concepts and structures that are useful to them for their tasks and for future search. This implies that an interface should be highly interactive, allowing search to rapidly branch as a result of the different types of information discovered. Second, effective sensemaking and search strategies require a variety of different structures to explain the information discovered, and guide search to conceptually related information. These structures may exist a priori, be derived from the results, or be constructed by the user, and the interface should make them available for rapidly accessing related information. Third, these structures are more useful when they correspond to the knowledge and goals of the user. This final point implies that the interface should have some method of adapting the presentation of information to the changing focus and needs of the user.

2.2 Visual IR Interfaces (VIRI)

The desiderata for VIRI proposed by Furnas and Rauch, as well as their discussion of the synergy between search and browsing in NaviQue, further emphasize the importance of multiple navigable structures in interactive search [15]. The NaviQue prototype is a Zoomable User Interface (ZUI) based on Pad++ [4] that allows a user to access more detailed information through zooming. A limitation of ZUIs is that the information required for different search strategies may be distributed too far apart to be viewed in detail at one time. Magic lenses [6] and spatial distortion techniques [23] show detailed information in the vicinity of the focus while also showing global context, however they are also limited by the usefulness of local information for different search strategies.

An alternative approach is to combine visualizations that are designed to support different search strategies. These views may be independent, such as ThemeView/Themescape for overviews and Tumbleweed for more detailed analysis in the SPIRE suite of tools [17], or tightly coordinated, such as the maps and parallel coordinates of City'O'Scope [7] and the document and keyword maps of IDM [16]. Although coordinated views are able to show information in multiple contexts simultaneously, there is a cost in aligning information between views. Brushing techniques are useful for identifying objects between views, but they do not project relational information between views.

2.3 Focus + Context Techniques

Furnas first described fisheye views as a technique for selectively reducing the information in a display to show the most interesting items, where interest was calculated as a tradeoff between *a priori* importance (global context), and relevance to the user's current task [14]. Furnas suggested that this general technique could be

Figure 2: A model showing that the fundamental difference between distortion, graphical fisheye view, and semantic fisheye view techniques is where *degree of interest* is calculated.

used to create compact views in a variety of different domains by redefining the function that calculates *degree of interest*.

Researchers have developed a wide range of fisheye view or focus + context techniques. We use a modified version of the Data State Model [11], as shown in Fig. 2, to compare the different types of focus + context techniques. Distortion techniques [23] use geometric transforms to magnify the area near the focus. Graphical fisheye views [33], increase the size or detail of information related to the focus within the structure and constraints of a graphical model. The effectiveness of distortion techniques and graphical fisheye views for opportunistic search depends largely on whether distance within the view or graphical model corresponds to the needs of the user. Semantic fisheye views, on the other hand, are independent of a particular visual representation, and calculate *conceptual distance* from the focus within one or more data models [21]. For example, ScentTrails [27] can be considered a recent example of a semantic fisheye view where the *degree of interest* of objects in a Web page is calculated using a model of Information Scent [28, 10] and relevant information is then highlighted by modifying the underlying HTML.

An important property of semantic fisheye views is that since *degree of interest* is calculated independent of a view they can be reused with different visualizations as suggested by [8], and extend a single visualization to show multiple contexts as demonstrated by the InfoLens prototype [26]. For example, the ZoomIllustrator adjusts the detail of objects in a graphical anatomical model based on user interaction within a related hypertext [31]. The resulting view may reveal information that is distributed throughout the entire model, such as adjusting the transparency of skin and muscles to show the components of the circulatory system, as well as filtering or adding information within the related hypertext. In this case, detailed information is shown based on an underlying semantic models (e.g., the structure of the circulatory system and the hypertext documentation) rather than proximity in either graphical model.

3 THE CORBIS IMAGE COLLECTION

To evaluate the effectiveness of SFEVs for opportunistic search, we have developed a prototype that allows a user to interactively explore a large collection of images and keywords using two different types of search strategies. The first SFEV is based on the similarity and direct relationships between objects in the collection, which corresponds to common search and browsing behavior. The second SFEV uses WordNet to find concepts related to the focus, which corresponds to several search strategies identified by Bates [1]. When the user's focus changes, the prototype calculates the *degree of interest* of related objects using a combination of different metrics, and then updates their representation using emphasis

techniques.

The prototype is built on a diverse collection of over 56,500 “royalty-free” image thumbnails donated to us by Corbis for this research. The images in the collection cover a wide range of subjects, and are annotated with a rich vocabulary of over 28,000 unique keywords. Each image in the collection is professionally annotated with keywords describing various aspects of the image, such as people, animals, objects, actions, the mood, and the location. The annotations are unstructured lists of words, but are generally ordered from most to least important. Each image has, on average, 23 keywords, which is sufficient to use traditional information retrieval techniques with few modifications.

The annotations are provided with the images, and were produced with the aid of a thesaurus for several reasons. First, the thesaurus aids the experts annotating the images in attributing the most specific concepts relevant to the image. Second, the thesaurus assures a controlled vocabulary by expanding each concept with a limited number of common synonyms. Third, the thesaurus facilitates search by adding the flattened hierarchy of keywords from a specific concept upward to progressively more general concepts.

The thesaurus used for annotating the images is proprietary and is not distributed with the images. Furthermore, there is no explicit indication within the annotations that indicates the structure they are derived from. However, we have found that there is sufficient overlap between this thesaurus and WordNet to use the semantic relationships in WordNet to effectively search and browse over the image collection.

4 VISUAL LAYOUT

We use a spring layout to position the images and keywords in the workspace by modeling both types of objects as nodes in a graph, and the relationships between them as edges. The importance of a keyword in an image is assigned to the weight of the edge connecting them. A single keyword may be connected to multiple images, and there are no direct connections between images or between keywords.

Our layout algorithm is adapted from the implementation of the spring layout in the graphical model of the KAON library [25]. In general, a spring algorithm assigns a repulsive force to the nodes in a graph, and models the edges between nodes as springs with a tension based on their weight. To reduce empty space in the layout, we separated the repulsive force of each node into x and y components that varied according to the image or keyword dimensions. Furthermore, to improve legibility we exaggerated the y-component of the repulsive force for keywords. The algorithm initially places nodes randomly, and then iteratively attempts to minimize tension. The resulting layout tends to place highly connected sets of nodes into visual clusters, and minimally connected nodes drift towards the outside of the representation.

5 SEMANTIC FISHEYE VIEWS

The images and keywords that are found as the result of lexical and semantic queries are loaded into the local workspace of the prototype. The cost of accessing information in the local workspace is much lower than query-based interaction, which encourages opportunistic search and exploration. This workspace is designed to allow users to rapidly browse over the collection by simply brushing over images and keywords. As the focus changes, the interface calculates the relative interest of all every objects in the workspace and smoothly animates changes in their representation.

The semantic fisheye view framework describes two components that influence the degree of interest assigned to each object in the interface. The first is a priori interest (API), which models the static landmarks in the information collection. The second component is

Figure 3: The API of the images reflect their relevance to a query for the “mare, female horse” concept.

the result of one or more interest metrics, and models the importance of objects with respect to the current focus. A wide range of semantic fisheye views can result from changing the balance and composition of these two components. We use the following general equation to calculate the degree of interest within a context:

$$DOI_{context}(x|fp = y) = API(x) + \sum_{i=1}^n w_i dist_i(x, y). \quad (1)$$

Each distance metric calculates relative interest in a different way, and potentially within different semantic models. For example, we use separate metrics to calculate the weight of images and keywords.

We define a focus as a tuple of one or more objects from the different domains of information in the workspace:

$$f_i = \langle Q_i^*, K_i^*, I_i^*, L_i^*, C_i^*, F_{n < i}^* \rangle. \quad (2)$$

The focus may include objects from the history of queries (Q), the keywords (K) and images (I) extracted from the Corbis annotated image collection, the lemmas (L) and concepts (C) extracted from WordNet, and the history of previous foci (F).

In the following section we discuss how API is used to establish global context in the workspace. Then we describe the interest metrics we have developed to support two distinctly different approaches to opportunistic search, one based on similarity and the other on semantic structure.

5.1 A Priori Interest (API)

Conceptually, the API establishes the global context in which the user searches. As the user moves the focus over the images and keywords in the collection the system will continuously recalculate the DOI of objects. However, when there is no current focus, the DOI will always return to the API value. In this way, objects with a high API will remain prominent and serve as visual landmarks.

The prototype allows the user to set the API in two different ways. First, the API may be defined by the DOI of a previous

Figure 4: Calculating the similarity between keywords and images.

query. In this case, we model the user’s focus as a lexical or semantic query, and the degree of interest of the images in the workspace reflect their relevance to the query. When the user moves their attention into the workspace, this relevance becomes the API. Fig. 3 shows the default view with the API set from the results of a semantic query for the mare, female horse concept. The size of the images reflects their API. The screenshot on the right adds keywords to the workspace. The size and brightness of the keywords reflects their frequency in the results of the query. Alternatively, the API of the keywords can be set to emphasize infrequent keywords by using idf. This is effective for highlighting unique words, such as names.

The user may also define the API from the values of a previous focus. In this case, the API is used to accumulate important objects in the workspace. In a fluid workspace, the relative sizes of information will rapidly follow changes of the focus. The API is the component of the fisheye used to model the information that should remain stable. In this case, we use API to model information that the user would like to remain persistent, such as a selection. For example, this would allow a user to compare multiple foci by selecting one object and then brushing over another.

5.2 Similarity-guided Browsing

The first set of interest metrics are designed to reveal information in the collection that is similar to the current focus based on content. We model the similarity between keywords and images as symmetric, and derive the value from the order keywords are used to annotate an image. This heuristic depends on how the images are annotated, and generally works well in this collection.

Fig. 4 shows a data state model of similarity-guided search when a keyword is the focus. After the links from a keyword to related images are expanded, the collection is limited to an ordered set of the most important based on a threshold. The DOI distribution is distorted and scaled to increase contrast, and then passed to a function that animates changes in the DOI of objects in the collection. Emphasis techniques depend on the DOI value to determine the visual weight to render objects, so any number of visual properties can be automatically coordinated in this way. Fig. 5 shows the similarity value calculated using this metric for an image to all of its keywords. The degree of interest is shown using size and saturation in the representation of the images, keywords, and edges.

In addition to the direct links between images and keywords, we calculate the similarity between images based on the keywords they have in common, as shown in Fig. 6 and Fig. 7. The prototype also displays a limited number of the keywords the images have in common. For example, all images are connected to the keyword “horses”, the cluster on the bottom right is highly connected to “Foals” and “Baby animals”, and a group on the top left is highly connected to the keyword “Females.”

The metrics discussed above reveal similarities between objects,

Figure 5: Representing the similarity between an image and its related keywords.

Figure 7: Representing the similarity between images.

Figure 6: Calculating the similarity between images.

Figure 8: A portion of WordNet showing the overlapping generalization and composition hierarchies.

and create visual clusters of tightly connected nodes. When the user changes their focus, the interface recalculates the degree of interest of every object to reflect their similarity to the new focus. This supports opportunistic discovery of images that are similarly annotated, and the keywords that are used in the vocabulary of the collection. In the following section, we describe metrics for finding information that is conceptually related, but not similarly annotated.

5.3 Semantic-guided Browsing

The second type of SFEV implemented in the prototype allows a user to search and browse over the image collection using the semantic relationships modeled in WordNet. In the following section we give a brief overview of WordNet, and then describe our implementation.

5.3.1 WordNet

WordNet is a general lexical ontology originally developed at Princeton in 1985 to test psycholinguistic theories of the English language [24]. WordNet has continued to evolve since then, and the version this research is based on (1.7.1, August 2002) contains approximately 140,000 unique word forms with 111,000 different senses.

Only the base forms of words are stored in the WordNet database, and are referred to as *lemmas*. The different senses of words (i.e., concepts) are modeled as sets of synonyms, and are referred to in WordNet's documentation as synsets. These concepts and lemmas are organized into a network using a set of semantic and lexical relationships. For example, Fig. 8 shows a portion of the overlapping generalization and composition hierarchies for the concept "horse".

The overlap between the thesaurus used to annotate images in the

Figure 9: The correspondence between the hierarchical annotation in Corbis, and the hyponym generalization hierarchy in WordNet [22]

Figure 10: A portion of the semantic neighborhood of the “horse” concept based on the SUPER, SUB and SIBLING strategies.

Corbis collection and the hierarchies in WordNet enables us to derive relationships between the keywords in Corbis and the concepts in WordNet with an accuracy of over 90% for the most frequently occurring keywords [22]. Fig. 9 shows an example of the overlap between the annotations of an image of an Arabian horse and the generalization hierarchy in WordNet.

5.3.2 Semantic Interest Metrics

Fig. 10 shows examples of the semantic neighborhood that could be found using several common strategies people use to expand queries [1]. For example, if a search for “horse” returned too many images, a person may use a more specific query, such as “wild horse” or “foal” to find a smaller, more manageable set of images. However, in order to apply these strategies without access to a semantic model such as WordNet, a person would have to know the existence of more general, more specific, or related concepts and add them to the query. Implementing these strategies using SFEV techniques allows a user to simply brush over a keyword and see the related concepts that exist in the image collection.

We use a single complex composite metric to support three of the basic search strategies described by Bates: SUPER, SUB, and SIBLING. Each of these strategies is modeled as a directed search in WordNet along a particular type of relationship. The composite metric coordinates information common to these three strategies and aggregates their results, but each strategy operates in a separate thread to optimize performance. The output of this composite metric is a limited sample of the most interesting results from each strategy. These results are carefully limited to maintain a relatively constant amount of visual complexity in the interface and to avoid having the results of one strategy dominate the others.

We decided to support all three strategies simultaneously for several reasons. First and foremost, this allows users to opportunistically discover concepts in the neighborhood of their focus without having to explicitly search for them. Second, we repeatedly found that having this semantic neighborhood visible allowed users to adjust more easily when the formal semantic structure modeled in WordNet did not match their expectations. Third, the combination

Figure 11: Calculating interest with the SUPER, SUB and SIBLING strategies using a composite interest metric.

Figure 12: Representing a sample of concepts most closely related to “horses” calculated with multiple strategies. The legend on the top right indicates that more general concepts are shown in red, more specific concepts are shown in cyan, and siblings are shown in magenta.

of strategies allows a user to continuously build their knowledge of the domain and the vocabulary of the collection.

Fig. 11 shows a data state model tracing the flow of information in a composite semantic metric where the user is focusing on a keyword. This model is divided into four vertical regions: WordNet, the Image Collection, the Graphical Model, and the View. Conceptually, data flows from the left towards the view, and the user interacts from the right. The leftmost vertical region is the subgraph of WordNet containing all the concepts directly and indirectly related to the keywords loaded into the workspace. The vertical region in the center of the model is the set of images and their keywords loaded into the workspace as a result of one or more queries. WordNet and the image collection are separate models connected through a mapping function.

The SUPER, SUB, and SIBLING strategies are implemented using a complex composite interest metric, shown in the gray box labeled Interest on the top left of Fig. 11. When a user brushes over

a keyword in the graphical model, the composite metric is passed the associated keyword object as a focus, and the metric calculates the DOI for related images and keywords using the different strategies in parallel. The results from each strategy are limited so that the final view has a representative sample from each strategy. The resulting image is shown in Fig. 12.

5.4 Emphasis Techniques

In the SFEV framework, emphasis techniques modify the visual representation of information to reveal changes in degree of interest. The goal of the emphasis techniques is to align the visual weight of objects with their importance in a particular context so that the most interesting objects are immediately apparent (i.e., pop out), and less interesting objects fade to the background. The relative contrast creates a visual ordering that allows a user to rapidly and opportunistically access contextual information. Although this approach is conceptually straightforward, in practice there are complex interactions between the visual scales used to encode interest, the distribution of degree of interest values, and the density of information in the display.

The most effective visual scale for presenting quantitative information, such as degree of interest, is position. For this reason, position is almost always used by visualization techniques to encode the relationships between the primary data values. For example, we use position in this prototype to show global relationships between images and keywords using the spring layout. Graphical fisheye views often distort position and scale to increase the visibility and detail of information spatially near the focus. Because the interest metrics used by semantic fisheye views are independent of spatial properties, distorting position to emphasize relative interest would be too disorienting to support rapid exploration. For this reason, we do not distort position in this prototype. Instead, we use a combination of size, color saturation, and density to encode degree of interest. These visual scales are less effective for precise quantitative comparisons of interest, but are reasonably effective for showing categorical and ordinal information (e.g., membership in the set of interesting objects, and relative order).

Each emphasis technique is scaled to a separate range based on several factors. First, according to Webers Law and Stevens Law the sensitivity and accuracy of the human eye for comparing magnitude of visual stimuli is different for each type of emphasis technique [12, 13]. In general, non-spatial emphasis techniques are not effective for quantitative comparisons, but they are useful for ordinal comparison. For example, on a complex display, it is difficult to perceive keywords that are below 50 percent saturation. Second, we create overlapping bands of DOI for background, selected and highlighted information. Background information has high contrast and distortion so that only a small representative portion of the entire collection is shown in detail. Selected information has a higher minimum bound and narrower range than background information so that it is easier to see the entire set of selected objects. Highlighted objects have higher minimum and maximum values to make them more visually apparent than most other information.

5.5 Interaction

The user may select among different types of SFEVs from a pull-down menu, and precisely refine the parameters of the interest metrics and emphasis techniques using a floating control panel. However, contextual control would be more effective, such as the contextual menu proposed by [29] for controlling multiple contexts in ZUIs.

Responsiveness is critical in maintaining a coherent visualization and allowing rapid opportunistic search, therefore the interface uses brushing rather than selection to activate the focus of the fisheye view. This makes the interface highly sensitive to changes in the

focus, but also potentially disorienting and difficult to understand if too much information changes at the same time. The prototype uses animated transitions and multiple threads to immediately begin or immediately interrupt metrics when the focus changes. A second critical element of SFEV is controlling the visual complexity of transitions and final views. We used hierarchical goal decomposition to analyze the sequence of actions and information required at different steps in the opportunistic task [5]. We noticed that in the process of analyzing the images in the view, users looked at information in a predictable sequence. Users would first rapidly scan over the images and captions in the collection. When they found an image that was interesting, they would pause to look at the keywords and compare them to the visual content of the image. The hierarchical goal decomposition allowed us to identify the knowledge required at each step of this process, and limit the complexity of the view to the minimum information that is needed. Although subtle, the pause effectively separates the rapid course analysis of images and their captions from the more detailed analysis of correlating keywords and images.

6 EXPERIMENTAL COMPARISON OF SIMILARITY-GUIDED AND SEMANTIC-GUIDED SEARCH

We conducted a formal user evaluation comparing the effectiveness of the two SFEV interfaces described in this paper for a complex opportunistic search task. The experiment was a within-subject design with two independent variables: type of SFEV and image collection. The task was to complete a partially filled concept hierarchy of either types of hoofed mammals or types of sea mammals, depending on the image collection. Both image collections were the same size, and we attempted to have similar distributions of images representing the concept hierarchies. We had sixteen participants in the experiment, four in each experimental condition.

The initial results of this experiment found that the participants using the semantic-guided interface found significantly more concepts ($p < .01$), made significantly fewer errors ($p < .016$), and were significantly more confident in their results ($p < .01$). The participants also overwhelmingly preferred the semantic-guided approach over the similarity-guided approach, even though it is visually more complex. A number of participants commented that similarity-guided search was easy to use, but they often felt they were “turning in circles” because it was difficult to break out of a cluster of similar images.

7 CONCLUSION

Opportunistic search and sensemaking in large information collections are highly interactive tasks that are poorly supported in current interfaces. These types of search activities require rapidly discovering, analyzing, and navigating between the relationships within information collections. However, a significant obstacle for users to effectively search over unstructured collections is their lack of domain knowledge, such as the vocabulary and semantic structure of the collection. To overcome this obstacle, researchers have proposed that interfaces should support search strategies to guide users over information in a collection.

The relationships within information collections are often too complex to be displayed in a single representation. We propose semantic fisheye views as an interactive visualization technique to support effective guided exploration over unstructured collections of information. Fisheye views reduce the visual complexity of displays by selectively emphasizing the most important information in a representation and deemphasizing or filtering less important information. The measure of importance is based on the user’s current focus and activity. An advantage of fisheye view techniques is that the metrics to determine importance are flexible, and can therefore

interactively support a wide range of search strategies over the same visual representation.

The main contribution of this research is the extension of focus + context techniques to effectively support multiple search strategies within a visualization. Initial experimental results suggest that semantic fisheye views are promising techniques for opportunistic search, and that semantic-guided search may be more effective than similarity-guided search for complex sensemaking tasks.

8 ACKNOWLEDGEMENTS

This paper is based in large part on unpublished research in [20]. This research was funded by grant 2000-066816 from the Swiss National Science Foundation, and was supported by the generous donation of annotated images by Corbis.

REFERENCES

- [1] M. J. Bates. Information search tactics. *Journal of the American Society for Information Science*, 30:205–214, 1979.
- [2] M. J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5):407–424, 1989.
- [3] M. J. Bates. Where should the person stop and the information search interface start? *Information Processing and Management*, 26(5):575–591, 1990.
- [4] Benjamin B. Bederson and James D. Hollan. Pad++: A zooming graphical interface for exploring alternate interface physics. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Visualization I, pages 17–26, 1994.
- [5] Suresh K. Bhavnani and M. J. Bates. Separating the knowledge layers: Cognitive analysis of search knowledge through hierarchical goal decomposition. In *Proceedings of the American Society for Information Science and Technology Annual Meeting*, volume 39, pages 204–213, 2002.
- [6] Eric A. Bier, Maureen C. Stone, Ken Pier, Ken Fishkin, Thomas Baudel, Matthew J. Conway, William A. S. Buxton, and Tony DeRose. Toolglass and magic lenses: The see-through interface. In *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, volume 2 of *VIDEOS: Part II – Techniques for Improved Human-Computer Interaction*, pages 445–446, 1994.
- [7] Dominique Brodbeck and Luc Girardin. Design study: Using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In *Proceedings of International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pages 104–111, London, UK, 2003. IEEE Computer Society.
- [8] Stuart K. Card and David Nation. Degree-of-interest trees: A component of an attention-reactive user interface. In *International Conference on Advanced Visual Interfaces (AVI02)*, Trento, Italy, 2002. ACM Press.
- [9] Stephen M. Casner. A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, 10(2):111–151, 1991.
- [10] Ed Huai-hsin Chi, Peter Pirolli, K. Chen, and James Pitkow. Using information scent to model user information needs and actions on the web. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 490–497, Seattle, WA, 2001. ACM.
- [11] Ed Huai-hsin Chi and John T. Riedl. An operator interaction framework for visualization systems. In *Proceedings of IEEE Symposium on Information Visualization (InfoVis98)*. IEEE Computer Society, 1998.
- [12] William S. Cleveland. *The elements of graphing data*. Wadsworth Advanced Books and Software, Monterey, Calif., 1985.
- [13] William S. Cleveland and Robert McGill. An experiment in graphical perception. *Int. J. Man-Machine Studies*, 25:491–500, 1986.
- [14] G. W. Furnas. Generalized fisheye views. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 16–23. ACM Press, 1986.
- [15] George W. Furnas and Samuel J. Rauch. Considerations for information environments and the navique workspace. In *Proceedings of the third ACM conference on Digital libraries*, pages 79–88. ACM Press, 1998.
- [16] Mountaz Hascoet-Zizi and Nikos Pediotakis. Visual relevance analysis. In *Proceedings of the first ACM International Conference on Digital Libraries (DL96)*, Human-Computer Interaction: Browsing and Visualization, pages 54–62, Bethesda, Maryland, 1996.
- [17] Beth Hetzler, Paul Whitney, Lou Martucci, and Jim Thomas. Multi-faceted insight through interoperable visual information analysis paradigms. In *Proceedings of IEEE Symposium on Information Visualization (InfoVis98)*, pages 137–144, Research Triangle Park, North Carolina, 1998. IEEE Computer Society.
- [18] Peter Ingwersen. *Information Retrieval Interaction*. Taylor Graham, London, 1992.
- [19] Peter Ingwersen. Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 101–110. Springer-Verlag New York, Inc., 1994.
- [20] Paul Janeczek. *Interactive Semantic Fisheye Views for Information Workspaces*. PhD thesis, Swiss Federal Institute of Technology - Lausanne (EPFL), in review.
- [21] Paul Janeczek and Pearl Pu. A framework for designing fisheye views to support multiple semantic contexts. In *International Conference on Advanced Visual Interfaces (AVI02)*, pages 51–58, Trento, Italy, 2002. ACM Press.
- [22] Paul Janeczek and Pearl Pu. Searching with semantics: An interactive visualization technique for exploring an annotated image collection. In Robert Meersman and Zahir Tari, editors, *OTM Workshops*, volume 2889 of *Lecture Notes in Computer Science*, pages 185–196, Catania, Italy, 2003. Springer.
- [23] Y. K. Leung and M. D. Apperley. A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction*, 1(2):126–160, 1994.
- [24] George A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
- [25] Boris Motik. The karlsruhe ontology and semantic web framework (kaon), 2003.
- [26] Gary Kwok-Chu Ng. *Interactive Visualisation Techniques for Ontology Development*. PhD thesis, University of Manchester, 2000.
- [27] Christopher Olston and Ed Huai-hsin Chi. Scenttrails: Integrating browsing and searching on the web. *ACM Transactions on Computer-Human Interaction*, 10(3):177–197, 2003.
- [28] Peter Pirolli. Computational models of information scent-following in a very large browsable text collection. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, volume 1 of *PAPERS: Finding What You Want I*, pages 3–10. ACM, 1997.
- [29] Stuart Pook, Eric Lecolinet, Guy Vaysseix, and Emmanuel Barillot. Context and interaction in zoomable user interfaces. In *Proceedings of AVI 2000*, pages 227–231. ACM, 2000.
- [30] J. J. Rocchio. The smart retrieval system – experiments in automatic document processing. In *Relevance Feedback in Information Retrieval*, pages 313–323. Prentice Hall, 1971.
- [31] Michael Ruger, Bernhard Preim, and Alf Ritter. Zoom navigation: Exploring large information and application spaces. In *Workshop on Advanced Visual Interfaces (AVI96)*, pages 40–48, New York, 1996. ACM Press.
- [32] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. The cost structure of sensemaking. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 269 – 276, Amsterdam, 1993. ACM.
- [33] Manojit Sarkar and Marc H. Brown. Graphical fisheye views. *Communications of the ACM*, 37(12):73–84, 1994.
- [34] Jiajie Zhang and Donald A. Norman. Representations in distributed cognitive tasks. *Cognitive Science*, 18:87–122, 1994.