Opportunistic Search with Semantic Fisheye Views

Paul Janecek and Pearl Pu

Swiss Federal Institute of Technology - Lausanne (EPFL), 1015 Lausanne, Switzerland,
{paul.janecek,pearl.pu}@epfl.ch,
WWW home page: http://hci.epfl.ch/

Abstract. Search goals are often too complex or poorly defined to be solved with a single query in a web-based information environment. In this paper we describe how we use *semantic fisheye views* (SFEVs) to effectively support opportunistic search and browsing strategies in an annotated image collection. We have developed a SFEV prototype enabling rapid, interactive exploration using both keyword similarity and semantic relationships derived using WordNet. 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 visualizations and information needs.

1 Introduction

Information seeking is often iterative, interactive, and opportunistic, especially when search goals are complex or uncertain. The results of a query may lead to the discovery of unfamiliar vocabulary and relationships that guide the future direction of search. This type of opportunistic search, which blends search and browsing behavior, is still a time-consuming and poorly-supported process in online search interfaces.

Visual Information Retrieval Interfaces (VIRI) allow users to rapidly shift between search and browsing tasks [8]. The tight coupling between visual representations and interaction make VIRIs powerful tools for discovering relationships between documents. However, as the amount of information displayed in a VIRI grows, it is increasingly difficult to represent and navigate over all of it within the constraints of screen size and access latency. As a result, a visualization is often a compromise between showing a small amount of information in detail, or a large amount of information abstractly. Furthermore, visual representations are often time-consuming to generate, but optimally supports only a small set of tasks [4, 23].

Semantic fisheye views (SFEVs) are a type of interactive *focus + context* visualization technique that attempts to manage the complexity of a visual interface by fluidly adapting the representation based on the user's current focus. SFEVs can use simple visual techniques to emphasize or increase the detail of the most important information, and de-emphasize or filter less important information, where *importance* is calculated relative to some measure of the user's current focus or task [7, 10]. By selectively reducing the complexity of a visualization, users are able to quickly understand and access local and global structure. For example, Fig. 1 uses these techniques to reveal a frequent theme in a collection.

Fig. 1. Reducing visual complexity with SFEVs. This simple example shows two images and the keywords used to annotate them positioned using a spring layout algorithm. On the left, the keywords are shown without any emphasis. The figures in the middle and on the right show the keywords emphasized by their relative importance in this small collection.

In this paper[1], we describe a prototype that uses SFEV techniques to support diverse search and browsing strategies within a large collection of professionally annotated images. Researchers have identified a wide range of browsing behaviors with different goals, strategies, and at different levels of granularity [14]. For example, Bates describes information seeking as an iterative, interactive process that evolves in response to the information found, and that encompasses both directed search and browsing [2]. Furthermore, the results of information seeking are not limited to documents, but also include the *knowledge* accumulated during the search process. This knowledge is essential for understanding and using the information discovered during search [21]. This model of search differs from the more classic query/document similarity in both the diversity and granularity of the information collected, as well as its evolving nature.

The prototype directly supports two alternative search strategies. The first technique emphasizes images and keywords that are similar in content to the focus. In contrast, the second technique emphasizes information that is *conceptually* related to the focus based on WordNet, a general lexical ontology of the English language. These strategies correspond to classic search strategies used for opportunistic search over heterogeneous collections of information [1, 2].

A significant contribution of this research is that it demonstrates a general method for integrating semantics directly into the interaction and visualization of information to support opportunistic search. This approach can easily be extended to support new search strategies that take advantage of metadata and ontologies.

In the following section we describe our general framework for designing semantic fisheye views, followed by an overview of the prototype architecture. We then explain how the prototype supports similarity- and semantic-guided browsing in greater detail, and discuss how we integrated WordNet into our interactive browser. We then briefly discuss some lessons learned and findings from user studies. Finally, we discuss related research and our conclusions.

[1] Color version available at http://hci.epfl.ch/website/publications/2004/janecek-wise04.pdf

2 Semantic Fisheye Views

We have designed and developed an interactive environment for exploring a large collection of professionally annotated images using *semantic fisheye view* techniques. The prototype directly supports different search strategies with separate types of fisheye views, one based on similarity metrics and another based on semantic relationships. The prototype was designed based on our SFEV framework [10], which we will describe generally in this section. In a previous workshop paper we described how this framework could be used to support query-based opportunistic search [11]; in this paper we focus primarily on browsing.

The semantic fisheye view framework describes two components that influence the *degree of interest* (DOI) assigned to each object in the collection. The first is *a priori* interest (API), which models "landmarks" in the collection, such as frequently used keywords or representative images. Conceptually, these landmarks provide a global context for the user to browse within. The second component, the set of interest metrics, models the conceptual distance of objects from the current focus. The combination of these two components results in a DOI value for each object in the collection.

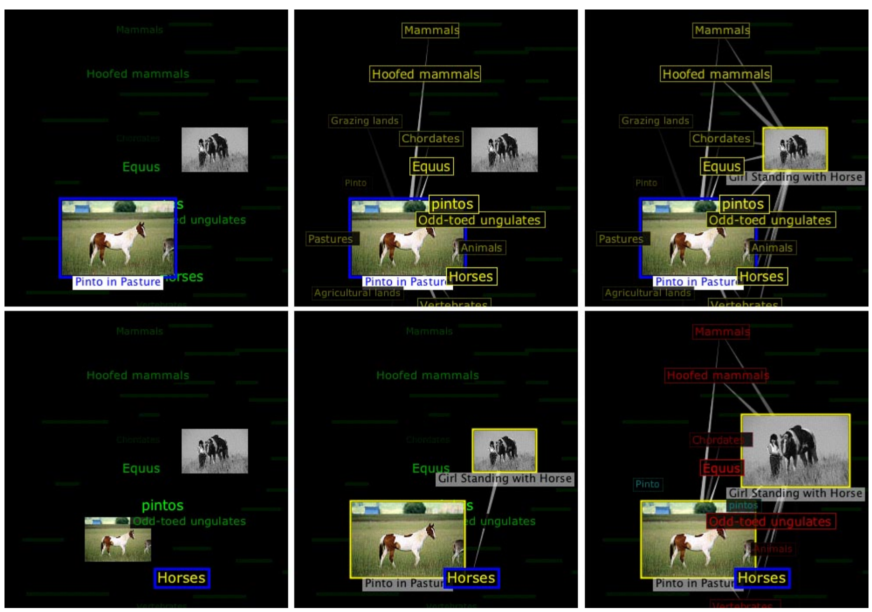

Fig. 2. Interacting with similarity (top row) and semantic (bottom row) SFEVs. Moving the mouse over an object causes related objects to smoothly increase in size and brightness based on the SFEVs interest metrics.

A wide range of semantic fisheye views can result from simply changing the balance and composition of these components. For example, Fig. 2 shows browsing over a collection using two different types of interest metrics. In the top row, the user focuses on an image and a similarity-based SFEV guides him to another image that is similar by content. In the bottom row, the user focuses on the keyword “Horses” and a semantic-based SFEV reveals other keywords that are semantically related, such as the “pinto” type of horse. In both cases, when the user brushes the mouse over an object it immediately begins to grow in size and detail and related information is progressively emphasized after a short delay. When the mouse moves away from an object, the related objects will slowly fade to their original state.

This example also shows how 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 DOI: the most interesting objects are immediately apparent, and less interesting objects fade into the background. Emphasis techniques impose a visual order over the objects in the collection and control the global contrast and complexity of the display. These techniques become progressively more important for exploration as the size of the collection increases. Another important aspect of SFEVs is the smooth animation that makes transitions between states understandable. In the following sections we provide greater detail about the general architecture of the prototype.

3 Architecture

The prototype is a complex Java application developed in our lab over several years to allow us to investigate the effectiveness of SFEV techniques for solving complex opportunistic search tasks in a semantically rich environment. Although the prototype does support keyword- and concept-based queries, it is primarily designed for rapid interactive exploration over the set of results. System responsiveness is critical for this type of task.

The architecture of the prototype can be divided at a high level into two parts: a client for launching queries and browsing over results, and a server that contains the image and WordNet databases. Only the relevant subset of these two large databases is loaded into the client browser to maintain system responsiveness and minimize traffic between the client and server.

Fig. 3 shows a general overview of how the prototype handles a simple query for “pintos.” The bold arrows trace the optimized path from a keyword-based query to the database (steps 1 and 2), followed by the construction of a graph from the resulting images and their keyword annotations (3), which is displayed in the view (4). The graph models images and keywords as nodes, and the links between them as edges. The importance of a keyword in an image is assigned to the weight of the edge connecting them. Small result sets will display in a few seconds and then animate as the layout algorithm places the nodes. At this point, the user can immediately begin interacting with the objects shown in the display.

The information required for semantic browsing is loaded into the browser in a second, background process shown with the smaller gray arrows (steps 5-7). For each keyword, the prototype loads the different word senses and all of their superconcepts (5) and constructs the WordNet subgraph (6). This is a query-intensive process that is done

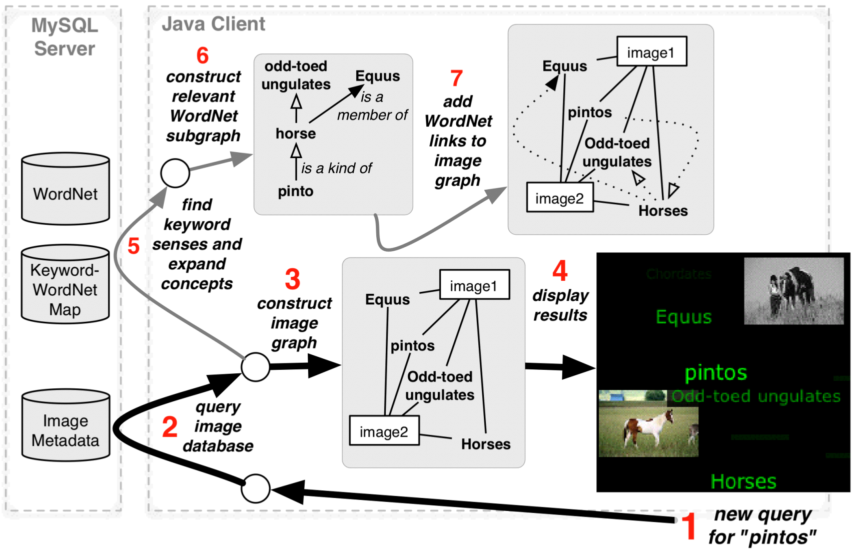

Fig. 3. The asynchronous architecture adds images to the workspace in the foreground and prepares for semantic interaction in the background. The images that match a query are quickly added to the display (1-4), while the subgraph of relevant concepts from WordNet is constructed in the background (5-7). These concepts are used during interactive exploration of the images.

in batches rather than one keyword at a time. As concepts are loaded into the browser, they are linked to the keywords (7). The WordNet subgraph is not shown directly in the display, but is used for rapid, interactive semantic browsing.

3.1 The Annotated Image Collection

The image database contains the metadata for a diverse collection of over 56,500 images provided to us by Corbis Corp for this research. 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.

Corbis uses a proprietary thesaurus to annotate their images, which is not distributed with the image collection. The thesaurus aids the annotators in attributing the most specific concepts relevant to the image, maintain a controlled vocabulary, and allows them to expand each concept with a limited number of common synonyms and the flattened hierarchy of keywords upward to progressively more general concepts. However, there is no explicit indication within the annotations of how the keywords are related. The annotations are unstructured lists of words, but are generally ordered from most to least important. Therefore, a user looking at the list would have to know and recognize how the keywords are related. Each image has, on average, 23 keywords, which is sufficient to use traditional information retrieval techniques with few modifications.

3.2 Visual Layout

An important property of SFEVs is that they are independent of a particular visual representation. We have implemented both tabular and map-based layouts of image collections that use the same interest metrics and emphasis techniques [19, 11]. The prototype described in this paper uses a spring layout to position the image graph.

Our layout algorithm is adapted from the implementation of the spring layout in the graphical model of the KAON library [16], however we do not use any of its functionality for managing ontologies. 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.

4 Similarity- and Semantic-guided Browsing

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 exploration. The prototype directly supports two different strategies for a user to browse over the collection (shown earlier in Fig. 2). The first strategy is based on content *similarity* and the direct relationships between objects in the collection. This corresponds to common keyword-based search and browsing behavior. The second strategy uses WordNet to find concepts related to the focus, which corresponds to *semantic* search and browsing behavior. In the following sections we describe how the prototype supports these strategies with different combinations of interest metrics.

4.1 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, and is derived from the direct links between images and keywords. The metrics calculate interest by searching outward from the focus to gather a set of related images and keywords. The interest of each object in the set is based on the weight of the links that lead to it. We derive the weight of these links from the order the keywords are used to annotate an image. This heuristic works well in this collection because the keywords are generally ordered from more specific to more general concepts as a result of the thesaurus-based term expansion.

These metrics reveal similarities between objects, 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, as well as the vocabulary of the collection. In the following section, we describe metrics for finding information that is conceptually related, but not similarly annotated.

4.2 Semantic-guided Browsing

The second type of SFEV implemented in the prototype allows a user to search and browse over the image collection using semantic relationships that are modeled in an external general ontology, WordNet. WordNet was developed at Princeton in 1985 to test psycholinguistic theories of the English language [15], and has continued to evolve since then. The version this research is based on (1.7.1, August 2002) contains approximately 140,000 unique word forms with 111,000 different senses.

While searching for information, users will often apply a wide range of strategies to find conceptually related information. 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. These strategies were described by Bates as generalization and specialization tactics [1]. A problem with these strategies is that they require domain knowledge: 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.

In order to use WordNet interactively, we only load a relevant subset of the entire graph into the browser (shown in steps 5-7 of Fig. 3). There were several significant steps required to prepare WordNet and the Corbis database to support rapid interactive browsing. The first step was to create a mapping between the vocabularies of WordNet and Corbis. We discovered that there is a significant overlap between the vocabularies and relationships in WordNet and the thesaurus used to annotate the images in the Corbis collection. This allowed us to derive links between the two vocabularies for over 90% of the most frequently occurring keywords in the Corbis collection [11].

The second step was to extract the different hierarchical relationships in WordNet (e.g., kind-of, part-of, member-of) and store them in separate tables. This allows us to rapidly access the entire path from a concept to its root in a single query over a single table, which is a critical step in constructing the subgraph related to a set of keywords. The final step was to precalculate values for images and concepts that are used in lexical and semantic queries, such as keyword frequency and inter-concept distance.

We use a single composite metric to support three of the basic search strategies described by Bates [1]: SUPER, SUB, and SIBLING. Each of these strategies is modeled as a directed search in WordNet along a particular type of hierarchical relationship. The SUPER and SUB strategies trace upwards or downwards along the kind-of, part-of and member-of hierarchies, while the SIBLING strategy finds concepts that share the same superconcept (e.g., if "pinto" was the focus, it would find other types of horse).

Fig. 4 shows a data state model tracing the flow of information in the composite metric that supports these search strategies. From left to right, this model is divided into four vertical regions: WordNet, the Image Collection, the Graphical Model, and the View. The leftmost vertical region shows operations over the subgraph of WordNet loaded into the workspace. The Image Collection region shows operations over the image graph loaded into the workspace as a result of one or more queries. The Graphical Model and View regions show operations that transform the visual representation.

When a user focuses on a keyword, the composite metric maps the keyword to different possible senses in WordNet (steps 1,2), and then expands these senses out-

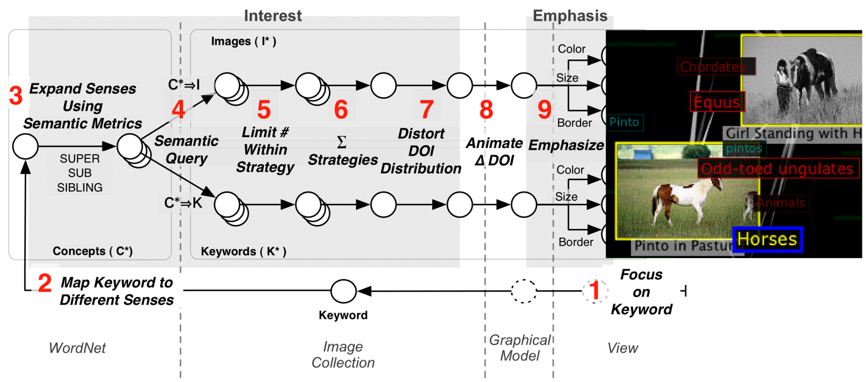

Fig. 4. Calculating interest with the SUPER, SUB and SIBLING strategies using a composite interest metric.

ward to find a limited set of the most closely related images and keywords using the SUPER, SUB and SIBLING strategies (steps 3-6). The composite metric coordinates information common to these three strategies, such as the different senses of a keyword and aggregates their results (step 6), but each strategy operates in a separate thread to optimize performance. The results of each strategy are limited so that the final view has a representative sample from each strategy while maintaining a relatively constant amount of visual complexity in the interface.

The keywords resulting from each type of strategy are displayed in different colors: more general concepts are shown in red, more specific concepts are shown in cyan, and siblings are shown in magenta. Fig. 5 shows a part of the results from focusing on the keyword "horses" in a collection of images. The colors allow the user to find related concepts quickly without adding lines to the complex display. Showing the conceptual neighborhood helps users opportunistically discover related concepts without having to explicitly search for them or navigate over a complex semantic model.

4.3 User Evaluation

Interaction is a critical component of fisheye views that was not originally addressed in our framework. During the development of the prototype, feedback from early users highlighted a significant problem of interaction that had to be overcome before we could conduct a formal user evaluation. Users were confused if too much information changed at the same time, but frustrated if information changed too slowly. Smoothly animating transitions and improving responsiveness with multiple threads helped, but the real solution was more subtle. We eventually found that we could support both rapid scanning detailed analysis by immediately showing the focus in greater detail and then progressively adding information after a short delay.

Once the prototype was usable, we conducted a formal evaluation comparing the effectiveness of the two types of browsing behavior. The results of this experiment suggested that the WordNet-based fisheye is more effective for complex exploration

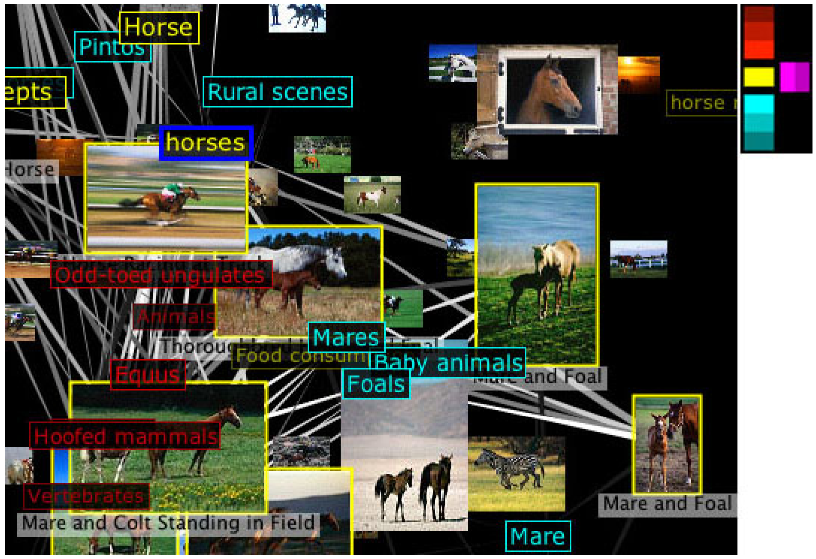

Fig. 5. Representing a sample of concepts most closely related to "horses" calculated with multiple strategies. The legend on the top right shows how colors are used to encode SUPER, SUB, and SIBLING relationships between keywords.

tasks [12]. The vast majority of users in the experiment strongly preferred the semantic interface over the similarity interface for this type of task because of the structure it provided for navigating over the collection and discovering new concepts. However, several people commented that it was difficult to estimate relative distances with the current emphasis techniques. Although the similarity-based SFEV was easy to use and understand, many people complained that it was more difficult to discover new information. In general, the comments were highly supportive and suggest that the SFEV approach is very promising for opportunistic search.

5 Related Work

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 and relevance to the user's current task [7]. Furnas suggested that this general technique could be used to create compact views in a variety of different domains by redefining the function that calculates *degree of interest (DOI)*.

Researchers have since developed a wide range of fisheye or *focus + context* techniques. *Distortion techniques* [13] use geometric transforms to magnify the area near the focus in the view. *Graphical* fisheye views [22] increase the size or detail of information related to the focus within the structure and constraints of a graphical model. The effectiveness of both distortion techniques and graphical fisheye views for specific search tasks depends largely on whether distance within the view or graphical model corresponds to the information needs of the user.

Semantic fisheye views, on the other hand, calculate interest based on *conceptual* distance from the focus to each object within one or more related data models, and then correspondingly modify the detail or emphasis [7, 10]. In this case, the usefulness of the emphasized information depends on the data model rather than any particular visual representation. Our approach to SFEVs extends Furnas' original proposal in two important ways: the direct support for multiple contexts and emphasis techniques.

There are a number of examples of these techniques applied to hypertext. The ScentTrails [17] prototype calculates the *DOI* of links in a Web page based on the user's current search goals, the hypertext path distance to relevant information, and a spreading activation model of Information Scent [18, 6]. The visual weight of the links is then adjusted to indicate relative interest by modifying the underlying HTML. The ZoomIllustrator [20] calculates *DOI* based on user interaction within a hypertext of human anatomy, and then adjusts the transparency and detail of relevant objects in a related illustrative graphical model.

The prototype we describe in this paper differs from these examples in several important ways. First, we use both similarity and semantic metrics to calculate *DOI*. The semantic metrics, in particular, are designed to provide a diverse sample for browsing rather than a narrow match. Second, the collection is dynamically constructed from the results of a query rather than a static hypertext. Third, the *focus* of the fisheye dynamically follows the mouse to support rapid brushing over the result collection rather than deliberate hypertext navigation, similar to the behavior of glosses in Fluid Documents [5].

Hollink, et al. recently developed a prototype that integrates four ontologies for searching within and annotating an image collection [9]. The link between ontologies is generated by hand, which increases precision over our technique but also limits its scalability. Our approach, in contrast, compensates for incomplete or ambiguous annotations using visualization techniques that place similarly annotated images near each other and interaction to support rapid exploration in both the collection and related ontology.

6 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.

7 Acknowledgements

This research was funded by grant 2000-066816 from the Swiss National Science Foundation. Corbis Corporation generously supported this research by giving us direct electronic access to their archive of annotated images. We greatly appreciate the helpful comments of the anonymous reviewers.

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. Stephen M. Casner. A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, 10(2):111–151, 1991.
5. Bay-Wei Chang, Jock D. Mackinlay, Polle T. Zellweger, and Takeo Igarashi. A negotiation architecture for fluid documents. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Enabling Architectures, pages 123–132. ACM Press, 1998.
6. 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.
7. G. W. Furnas. Generalized fisheye views. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 16–23. ACM Press, 1986.
8. 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.
9. Laura Hollink, A. Th. (Guus) Schreiber, Jan Wielemaker, and Bob Wielinga. Semantic annotation of image collections. In *Proceedings of the KCAP'03 Workshop on Knowledge Capture and Semantic Annotation*, Sanibel, Florida, 2003. ACM Press.
10. Paul Janecek 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.

11. Paul Janecek 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.
12. Paul Janecek and Pearl Pu. An evaluation of semantic fisheye views for opportunistic search in an annotated image collection. *Journal on Digital Libraries*, 4(4), October 2004. Special Issue on Information Visualization Interfaces for Retrieval and Analysis.
13. 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.
14. G. Marchionini. *Information Seeking in Electronic Environments*. Cambridge University Press, Cambridge, MA, 1995.
15. George A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
16. Boris Motik, Alexander Maedche, and Raphael Volz. A conceptual modeling approach for semantics-driven enterprise. In *Proceedings of the First International Conference on Ontologies, Databases and Application of Semantics (ODBASE-2002)*, LNAI. Springer, 2002.
17. 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.
18. 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.
19. Pearl Pu and Paul Janecek. Visual interfaces for opportunistic information seeking. In Constantine Stephanidis and Julie Jacko, editors, *10th International Conference on Human Computer Interaction (HCII '03)*, volume 2, pages 1131–1135, Crete, Greece, 2003. Lawrence Erlbaum Associates.
20. 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.
21. 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.
22. Manojit Sarkar and Marc H. Brown. Graphical fisheye views. *Communications of the ACM*, 37(12):73–84, 1994.
23. Jiajie Zhang and Donald A. Norman. Representations in distributed cognitive tasks. *Cognitive Science*, 18:87–122, 1994.