

Concept Expansion Using Semantic Fisheye Views

Paul Janecek¹, Vincent Schickel², Pearl Pu²

¹ Computer Science and Information Management, School of Advanced Technologies,
Asian Institute of Technology (AIT),
P.O. Box 4, Klong Luang
Pathumthani 12120, Thailand
`paul@cs.ait.ac.th`

² Human Computer Interaction Group, Faculty of Computer and Communication Sciences,
Swiss Federal Institute of Technology in Lausanne (EPFL),
CH-1015 Lausanne, Switzerland
`{vincent.schickel-zuber, pearl.pu}@epfl.ch`

Abstract. Exploratory search over a collection often requires users to iteratively apply a variety of strategies, such as searching for more general or more specific concepts in reaction to the information they encounter. Rich semantic models, such as WordNet, are potentially valuable aids for users to make sense of this information. However, these large complex models often contain specialized vocabularies and a detailed level of granularity that makes them difficult to use as an aid for opportunistic search. In this paper, we describe how Semantic Fisheye Views (SFEV) can be designed to transparently integrate rich semantic models into the search process, allowing users to effectively explore a diverse range of related concepts without explicitly navigating over the underlying model. The SFEV combines semantic guided search with interactive visualization techniques, creating a search tool that we have found to be significantly more effective than keyword similarity alone for exploratory tasks.

1 Introduction

Similarity-based search models (such as the vector space model and relevance-feedback algorithms) are often very effective for precise queries, but less effective when search goals are not easily defined, such as a search to learn about an unfamiliar domain of knowledge or to discover the diversity of “interesting” information in a collection. This type of search is not simply a series of independent iterative queries, each of which is progressively refined towards more relevant information. On the contrary, it is an interactive, opportunistic process that evolves in response to the information found, the users’ knowledge, and their search strategies [1]. An important component of this process is “sensemaking,” where users construct and refine their mental schemas of the concepts and relationships they encounter in the documents of a collection [2]. In this paper, we describe the implementation details of Semantic Fisheye View (SFEV) [3], a *focus + context* technique that interactively guides a user’s attention over a potentially dense visualization of information to the objects that are the most *semantically* related to their current focus.

Bates described the following characteristics that distinguish opportunistic (or “berry-picking”) search from a series of separate queries [1]:

1. *Nature of the query.* Queries are not static, but rather evolve as the user analyzes the result sets. It is important to note that this evolution is not an increase in precision, but rather a change of focus (e.g., exploratory, explanatory, exhaustive).
2. *Nature of the overall search process.* Information is gathered in bits and pieces instead of in a single set of results (e.g., lists of keywords, an author, documents).
3. *Range of search techniques used.* Users employ a wide variety of search strategies during the search process (e.g., keyword matching, concept expansion) [4].
4. *Information domain where the search is conducted.* More than one type of information is consulted during the search process (e.g., text, figures, cross-references).

Bates’ model inspired us to develop a prototype, VisAmp, which implements several information sensemaking strategies as *interest metrics* in the SFEV framework [5]. We described the architecture of SFEV in [5], focusing on the strategies used to support real time display of semantic fisheye views and how interaction can be rapidly and smoothly handled. In a related paper [3], we described an experimental evaluation where users were significantly more effective at sensemaking tasks with an interface that revealed semantically related information rather than one that revealed keyword co-occurrence. The results of the experiment do not discount the usefulness of keyword-similarity or imply that semantic models are more effective in all cases. However, the results do suggest that users greatly appreciate the opportunity to access information using semantic models. Our research is exploring how to integrate the increasingly available semantic models (alongside other similarity models) into highly interactive visual interfaces for information retrieval.

The contribution of this research to the domain of information retrieval is in the integration of semantic models into a highly interactive visual tool for strategically exploring, accessing, and understanding collections of information. The contribution of this paper over our previous work is a detailed description of how *degree of interest* functions combine multiple concept similarity metrics at the keyword and document levels to guide exploration over a collection.

In the following sections, we first examine existing work in several domains relating to our work. We describe the general framework of SFEV, including the interest metrics, the concept expansion and goal refinement mechanisms, and the visualization component of SFEV. We then illustrate our framework with a user scenario of information sensemaking within a large professionally annotated image collection. Finally, we conclude this paper with a short summary of the work achieved.

2 Related Work

This work can be compared with others on three themes: alternative search paradigms, visual information retrieval interfaces (VIRI), and fisheye view visualization techniques.

Researchers have sought ways to incorporate additional information, such as metadata and ontologies, into search tools. Some researchers have focused on tools for semantically organizing gathered information, for example with annotations [6] or

concept maps [7]. Our research, on the other hand, focuses primarily on the difficulty of encountering relevant information. This is especially true with image collections, which typically have few keyword annotations. Researchers have utilized semantic models to expand both the metadata related to an image as well as the query. A number of researchers have had some success using WordNet [8] to improve the effectiveness of keyword-based queries [9][6] and interactive browsing [10] over image collections. Researchers have also used similar techniques with other large ontologies [11] and combinations of ontologies [12].

However, there are several unresolved problems with these approaches. First, general solutions have had limited success in large, complex image collections and large ontologies. Large collections and ontologies are often inconsistent in their level of detail and incomplete in their coverage, which amplifies the difficult problem of matching the annotations in the image collection to the specific concepts in an ontology (i.e., lexical/semantic disambiguation). Interfaces that do not allow users to interactively adjust the matching algorithms between the collection, the ontology, and the query limit the users' ability to adapt in response to the results they find [1].

A third problem is that they are not visual. Furnas identified a number of advantages of Visual Information Retrieval Interfaces (VIRI) over more traditional query/result-list interfaces [13]. One of the most significant was the synergy between search and browsing. Displaying results in a persistent and meaningful location allows users to accumulate knowledge through navigation. One significant obstacle for effective visualizations is how to handle visual complexity as the amount of information in a representation increases.

Furnas [14] first described fisheye view as a technique for selectively reducing the information in a display to show only 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. Furnas suggested that this general technique could be used to create compact views in a variety of domains by redefining the function that calculates the degree of interest. Researchers have developed a wide range of fisheye view or focus + context techniques. Many of these use geometric distortions to magnify objects near the focus [15]. SFEVs, on the other hand [3], calculate conceptual distance from the focus within one or more data models, and are therefore independent of a particular visual representation [5].

Our research combines the main strengths of semantic-guided search, VIRIs, and *focus + context* visualization techniques in one framework. This combination allows users to visually explore the semantic relationships between documents as they refine their search goals.

3 The Semantic Fisheye View Framework

Fisheye views are based on the observation that, from the user's perspective, the importance or utility of information at any given moment is a function of two general components: *a priori* interest (*API*), and interest metrics. *API* reflects the importance of an object within a particular structure, task, or domain, and is independent of the user's current focus. Interest metrics determine the relative interest of every object in

the collection with respect to the user’s current focus and task. In a semantic fisheye view, both *API* and interest metrics can be generally described and combined [5]. When using SFEVs to implement search strategies, we model a user’s current search goal as a focus, and the system’s reaction in terms of *degree of interest (DOI)* and emphasis. The *DOI* is the relative importance of every object in the information space, and emphasis is a mapping between the *DOI* of an object to a visual property used to display that object, such as size.

We use the following general function to calculate the *DOI* in a particular context of an object x , given the focus f_i :

$$DOI_{context}(x | f_i = y) = f \left(API_i(x), \sum_{j=1}^n w_j dist_j(x, y) \right). \quad (1)$$

This general function highlights an important characteristic of this framework: both *API* and the conceptual distance between objects may be derived from one or more distance metrics (*dist_j*). The weight, w_i , is the weight associated with the distance metric being used and n is the number of metrics being used. We define a focus f_i , as a tuple of one or more weighted objects:

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

The focus may include objects from different domains: the history of queries (Q); the keywords (K) and images (I) extracted from the Corbis annotated image collection as a result of these queries; the Lemmas (L) and Concepts (C) extracted from WordNet that correspond to these keywords; and the history of previous foci (F).

The way in which these distance metrics are combined depends on the context being modeled. By orienting the framework to support multiple metrics, it is able to support richer models of user interest that may span multiple domains, such as related concepts and history of interaction. The prototype, which we have developed for exploring image collections, uses multiple *API* and distance metrics to model relative interest. The focus transitions between objects in the domains of queries, keywords, images and concepts.

3.1 A Priori Interest

Conceptually, the *API* establishes the global context in which the user searches. In our framework, the *API* is used to model the information that should remain stable as the focus changes. For example, when the user moves their 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* of each object will always return to its *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* interactively in two different ways. First, the *API* may be defined by the result of a query (3.a). In this case, we model the user’s focus as a lexical or semantic query (Q), and the *DOI* of the objects in the workspace reflect their relevance to the query. The prototype calculates the relevance of the keywords as a function of their frequency in the results of the query. By default

we use relative frequency, which emphasizes common themes in the collection. Alternatively, using the inverse document frequency (*idf*) emphasizes infrequent keywords, which is effective for highlighting unique words such as names.

The user may also define the *API* from the *DOI* calculated using a previous foci (3.b). In this case, the *API* is utilized to accumulate important objects in the workspace (like Bate's berry-picking strategy, or relevance-feedback algorithms).

$$API_i(x) = \begin{cases} DOI_{context}(x | f_j) & 0 \leq j < i \quad (a) \\ DOI_{context}(x | f_{i-1}) & (b) \end{cases}. \quad (3)$$

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.

3.2 Distance metrics

The SFEV models semantic queries for complex combinations of concepts as a distance metric between concepts within one or more related semantic model. The approach we used for semantic queries calculates the minimum distance between collections of concepts, based on research by [16]:

$$sim(c_1, c_2) = \frac{depth(c_a)}{depth_{max}}. \quad (4)$$

In this equation, c_1 , c_2 , and c_a are concepts in WordNet, *depth* is measured in one of the WordNet hierarchies, and c_a is the lowest common ancestor between c_1 and c_2 . This metric calculates the distance between concepts based on the generalization structure of WordNet. We precalculated the similarity between all concepts in the subset of WordNet that is related to the keywords in the image collection. The result of this time-consuming calculation is stored in a concept similarity table. To build this table, we first extracted the subset of concepts that have an exact or inexact match to keywords in the image collection. We then iteratively calculated the similarity of all combinations of concepts in the *kind-of*, *part-of*, and *member-of* hierarchies in WordNet using equation 4. Finally, we normalized the similarity values to the range [0,1].

We calculate the similarity between a query, Q , and the collection of keywords annotating an image, A , using equation, (5), proposed by [17] and also used by [16]:

$$sim(Q, A) = \frac{\sum_{i=1}^n \max_{j=1, \dots, m_i} \{sim(q_i, a_j) \cdot w_{q_i} \cdot w_{a_j}\}}{n}. \quad (5)$$

In this equation, $sim(q_i, a_j)$ is the precalculated similarity described in (4), w is the weight of the concept in $Q = \{ \langle q_1, w_{q_1} \rangle, \dots, \langle q_n, w_{q_n} \rangle \}$ or $A = \{ \langle a_1, w_{a_1} \rangle, \dots, \langle a_m, w_{a_m} \rangle \}$, and n is used to normalize the similarity measure. Essentially, this equation sums the similarities of the closest matching concepts between a query and an image. The images and keywords that are found as the result of lexical and semantic queries are loaded into the local workspace of the prototype.

3.3 Concept expansion

In addition to searching for images by similarity over its concepts, we have also developed a method for concept expansion that interactively reveals the semantic neighborhood of a concept using several of the search tactics identified by Bates [4]: SUPER (finding more general concepts), SUB (finding more specific concepts) and SIBLING (finding concepts from the same parent), which Bates referred to as RELATE). Each concept expansion command encapsulates an algorithm for iterating over the WordNet structure and generating queries for related concepts. The information seeker is only presented with related concepts that have instances in the collection (i.e., images and keywords), and therefore is not required to navigate over the complex structure and vocabulary of WordNet. This approach allows the user to direct the search process at the tactical and strategic level, as suggested by Bates [18].

3.4 Goal Refinement

As the user's focus changes, the interface calculates the DOI of all objects in the workspace and smoothly animates changes in their representation. The DOI is computed based on equation 1, and the interest metric is based on conceptual similarity. Thus, the new DOI will be computed as follows:

$$DOI_{context}(x | f_i = y) = API_i(x) + \sum_{j=1}^n w_j sim_j(x, y) \cdot \quad (6)$$

In this equation, the results of a number of different similarity metrics, n , are combined to determine DOI within a particular context. By exploring the information revealed by the SFEV, users learn the vocabulary and conceptual relationships within the collection and are able to interactively refine their search goals.

3.5 SFEV supporting Information Sensemaking

We now describe the interest metrics we have developed to support the sensemaking process. Figure 1 traces the flow of information from the user's interaction on the bottom right through the composite metric that calculates semantic similarity and back to the updated view on the top right. The model is divided into three vertical regions, from left to right: WordNet, the Image Collection, and the View. The WordNet model is the subgraph of WordNet that is related to the keywords in the image collection. The image collection is modeled as a graph, where images and keywords are nodes and the links between them are edges. The output of an interest metric is a table that assigns a new DOI to each object in the related collection.

When a user brushes over a keyword in the graphical model, the associated keyword object is passed as a focus to an interest metric that calculates the DOI for related images and keywords as follows:

1. *Map keyword to different senses ($k \Rightarrow C^*$):* A semantic disambiguation metric maps the keyword to one or more related concepts from WordNet and assigns a

weight, w , to each based on a precalculated keyword/concept mapping. The weight indicates the confidence that the mapping is correct.

2. *Concept expansion*: These concepts are passed to metrics implementing the SUPER, SUB, and SIBLING strategies. Each of these metrics traverse the WordNet graph in parallel, gathering a weighted collection of relevant concepts. The weight of each concept is inversely related to the distance.
3. *Semantic query*: The concepts found from each strategy are then passed to metrics that find the most relevant images and keywords:
 - $C^* \Rightarrow K^*$: Calculates the similarity between a concept and a keyword based on lexical/semantic disambiguation.
 - $C^* \Rightarrow I^*$: Calculates the similarity between a concept and an image by finding the related keywords, and then by calculating the sum of the weights between keywords and the images they are related to.
4. *Limit # within strategy*: These limits maintain a relatively constant amount of visual complexity, and avoid having the results of one strategy dominate the other.
 - K^* -Limit Number: the number of relevant keywords is limited for each strategy. By default, the SUB strategy is allowed more keywords than the others because it will tend to branch more quickly.
 - K^* -Limit Range: The distributions of interest returned by each strategy are scaled to a similar range (e.g., 0.3 – 1.0). Giving a preferred strategy a higher upper limit will make it more visible than the others.
 - I^* -Limit Number: number of relevant images can be limited for each strategy.
5. *Combine results of strategies*: The results of the different strategies are aggregated.
6. *Distort DOI distribution*: The distribution of interest for the sets of images and keywords are scaled to the range of 0.1 and 1.0. The distribution may also be distorted to increase the contrast between min and max values.

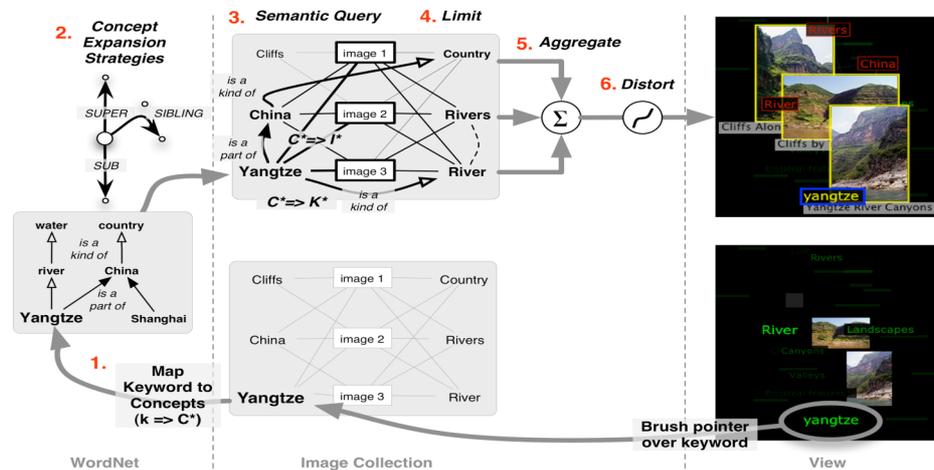


Fig. 1. SUPER, SUB and SIBLING strategies implemented using a composite semantic metric

The interface smoothly animates changes in the degree of interest value for each object, which cascades through different emphasis techniques to affect the representation of objects in the view.

3.6 Emphasis Techniques

Visualizations often attempt to show as much information as possible within the constraints of the available space. However, Pirolli et al. [19] point out that “squeezing” more information into the display does not necessarily “squeeze” more information into the mind, but that strong information scent cues and *focus + context* techniques can enable users to navigate and search through information at more than twice the rate of the user of a normal browser.

VisAmp uses several emphasis techniques to align the visual weight of objects with their semantic 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 (e.g., through dimming, shrinking in size and detail, or filtering) [3]. The relative contrast creates a visual ordering that allows a user to rapidly and opportunistically access the most important contextual information, i.e., visual emphasis corresponds to information “scent”.

4 An Image Retrieval Example

We demonstrate the prototype with a scenario where a student uses the prototype to learn about China before he attends a conference there. The user starts with an initial query for “China” to see what kinds of images are available. The system populates the workspace with several hundred pictures matching the query, positions them using a spring layout, and resizes them to reflect their relevance as shown in Fig. 2a. The layout organizes images so that similarly annotated images are near each other, such as the images of flags clustered at the bottom.

As he brushes the cursor over different images, they smoothly grow in size so he can read the captions and then fade slowly back to their original size as he moves to another. Pausing over an image reveals related keywords, and moving the cursor over a keyword reveals related images and concepts. For example, in Fig. 2b he pauses the cursor over *China*, which reveals subconcepts (in cyan) such as the *Great Wall* and the *Yangtze*, superconcepts (in red) such as *Asia*, and siblings (in magenta) such as *Nepal* and *New Zealand*.

Looking closer at the images and keywords related to the *Great Wall* leads him to the keyword *Building* (Fig. 2c), which expands to reveal a diverse range of subconcepts, such as *cafes*, *courtyards*, *skyscrapers*, *temples* and *ruins*. He decides to return to an unfamiliar word he saw earlier, *Yangtze*. Brushing over the keyword (Fig. 2d) reveals the superconcepts *River* and *China*, and siblings such as the *Great Wall*, as well as a cluster of images of water. This leads him to believe that the *Yangtze* is a river in China, and zooming in to look closer at the images confirms this. The responsiveness of the prototype allows him to rapidly discover and explore the concepts and images that capture his attention, and the underlying semantic interest metrics allow him to access the information in an opportunistic but well structured manner.

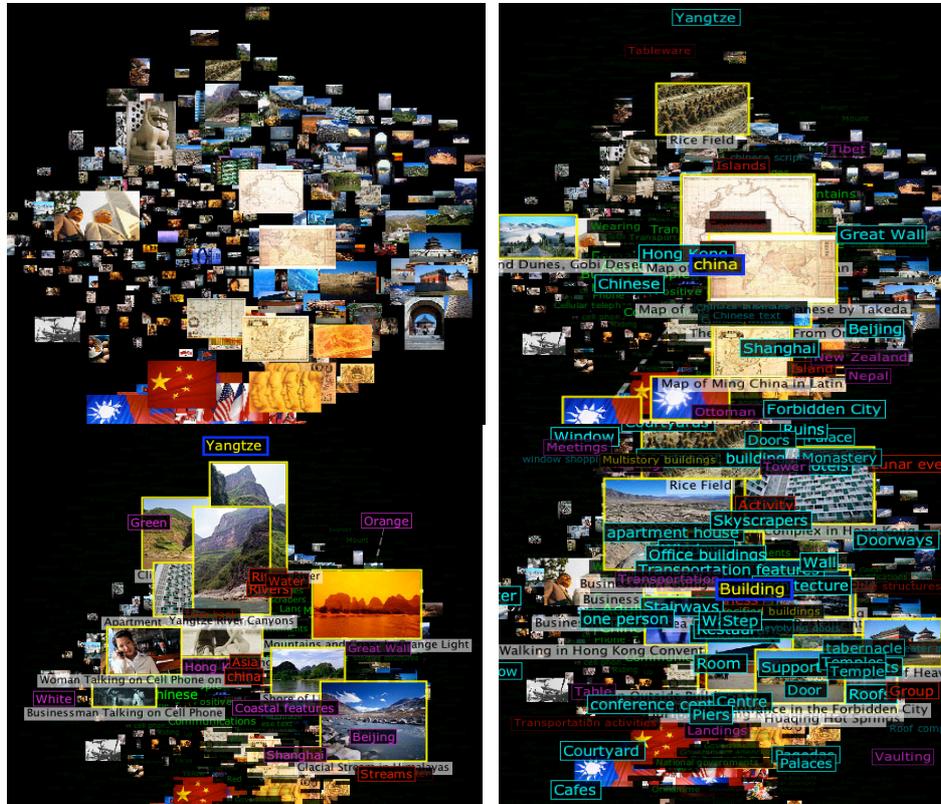


Fig. 2. Exploring images of China using VisAmp (clockwise from top left): (a) overview of search results; (b) brushing over the keyword *China* reveals subconcepts such as *Yangtze*, superconcepts such as *Asia*, and siblings such as *Nepal*; (c) *Building*; (d) *Yangtze*

5 Conclusions

In this paper, we have described the underlying interest metrics that allow SFEVs to support concept expansion and goal refinement using one or more semantic models into the interface. Users are able to rapidly and interactively discover new concepts and refine their current search goals by simply brushing over objects in the interface, without having to create queries or navigate through an explicit representation of the related semantic model. In many cases, large semantic models such as WordNet are too detailed, inconsistent, or confusing for users to navigate over explicitly.

Preliminary results extending VisAmp to scientific literature are encouraging. Using the ACM classification model as a domain ontology, users were able to explore concept hierarchies for learning about the domain. Future work includes supporting multiple concepts and refining the interest metrics to take into account multiple connections between concepts, as discussed by Andreasen in [20].

References

1. Bates, M. J., The Design of Browsing and Berrypicking techniques for the online search interface, *Online Review*, 13(5):407-424, 1989.
2. Russell, D., et al. The cost structure of sensemaking, in *Proc. of the ACM Conf. on Human Factors in Computing Systems (INTERCHI '93)*, 269-276, 1993.
3. Janecek, P., and Pu, P., An Evaluation of Semantic Fisheye Views for Opportunistic Search in an Annotated Image Collection. *Int. Journal of Digital Libraries*, 4(4). Special Issue on "Information Visualization Interfaces for Retrieval and Analysis." 2004.
4. Bates, M., J., Information Search Tactics, *Journal of the American Society for Information Science*, 30:205-214, 1979.
5. Janecek, P., and Pu, P., A Framework for Designing Fisheye Views to Support Multiple Semantic Contexts, in *Proc. of Int. Conf. on Advanced Visual Interface (AVI02)*, Trento, Italy, ACM Press, 51-58, 2002.
6. Sereno, B., Buckingham Shum, S., Motta, Enrico, ClaimSpotter: An environment to support sensemaking with knowledge triples. In *Proc. of the ACM Conf. on Intelligent User Interfaces (IUI'05)*, San Diego, CA, USA, January 2005.
7. Leake, D. B., et al., Aiding Knowledge Capture by Searching for Extensions of Knowledge Models. In *Proc. of the ACM Int. Conf. On Knowledge Capture (KCAP)*, 2003.
8. Miller, G. A., WordNet: A Lexical Database for English. *Com. of the ACM*, 38(11), 1995.
9. Yee, K.-P., et al., Faceted Metadata for Image Search and Browsing, in *Proc. of the ACM Conf. on Human factors in computing systems (CHI '03)*, 401-408, 2003.
10. Aslandogan, Y.A., et al., Using Semantic Contents and WordNet in Image Retrieval, in *Proc. of the 20th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*. p. 286-295, 1997
11. Liu, H., and Lieberman, H., Robust Photo Retrieval Using World Semantics, in *Proc. of the LREC 2002 Workshop on Creating and Using Semantics for Information Retrieval and Filtering: State-of-the-art and Future Research*, 15-20, 2002.
12. Hollink, L., et al., Semantic Annotation of Image Collections, in *Proc. of the ACM Workshop on Knowledge Capture and Semantic Annotation (KCAP'03)*. 2003.
13. Furnas, G., W., and Rauch, S. J., Consideration for Information Environments and the NaviQue Workspace, in *DL'98: Proc. of ACM Int. Conf. on Digital Libraries*, 79-88, 1998.
14. Furnas, G., W., Generalized Fisheye Views, in *Proc. of the SIGCHI Conf. on Human factors in computing systems*, ACM Press, 16-23, 1986.
15. Leung, Y., and Apperley, M., A Review and Taxonomy of Distortion-Oriented Presentation Techniques, *ACM Transactions on Computer-Human Interaction*, 1(2): 126-160, 1994.
16. Yang, J., et al., Thesaurus-Aided Approach for Image Browsing and Retrieval, in *Proc. of 2nd IEEE Int. Conf. on Multimedia and Expo (ICME 2001)*, 313-316, 2001.
17. Smeaton, A. F., & Quigley, I., Experiments on Using Semantic Distances between Words in Image Caption Retrieval, in *Proc. of the 19th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, pp. 174-180, 1996.
18. Bates, M., J., Where Should the Person Stop and the Information Search Interface Start?, *Information Processing and Management*, 26(5), 575-591, 1990.
19. Pirolli, P., Card, S., and Van Der Wege, M., The Effects of Information Scent on Visual Search in the Hyperbolic Tree Browser, *ACM Transactions on Computer-Human Interaction*, 10(1), 20-53, 2003.
20. Andreasen, T., Bulskov, H., and Knappe, R., From Ontology over Similarity to Query Evaluation, in *Proc. of 2nd Int. Conf. on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems (ODBASE)* (Catania, Sicily, Italy, 2003).