NEAR: VISUALIZING INFORMATION RELATIONS IN THE MULTIMEDIA REPOSITORY A•VI•RE

Victor Y. Chen[1], Cheryl Z. Qian[2], and Robert F. Woodbury[3]

ABSTRACT

This paper describes the NEAR (Navigating Exhibitions, Annotations and Resources) panel, a novel interactive visualization technique designed to reveal implicit relations such as sharing, reference and similarity. NEAR is implemented on A•VI•RE, an online information repository. A•VI•RE supports semi-structured collections (exhibitions) containing various resources and annotations. Users are encouraged to contribute, share, annotate and interpret resources in the system by building their own exhibitions and annotations. We present a visual panel that implements new navigation and communication approaches that support discovery of relations. By quickly scanning and interacting with NEAR, users can see not only implied relations but also potential connections among different data elements. NEAR was tested by several users in the A•VI•RE system and shown to be a supportive navigation tool. In the paper, we further analyze the design, report the evaluation and consider its usage in other applications.

KEY WORDS

measure similarity, trace reference, information sharing, inherent relation, information visualization, online multimedia repository

INTRODUCTION

The architecture of collection/resources appear in a variety of application domains such as: shopping carts and shopping items, papers and their citations, people in groups and their activities and schedules, classes and methods in Java programming. Being an extended online information repository system, A•VI•RE contains great amounts of semi-structured data such as resources, annotations and exhibitions in this collection/resources structure. By using graphic nodes, icons and links, we developed a visualization panel "NEAR" (Navigating Exhibitions, Annotations and Resources) as a small but supportive tool for users to navigate through A•VI•RE smoothly, revealing information and relations among different data elements and interpreting data elements from different perspectives.

This paper explores the visualization of unstructured relations such as information sharing, similarity and reference graphically among data elements in A•VI•RE. Principles of the

[1] Graduate Student & Research Assistant, Interactive Arts & Technology, Central City Tower, Simon Fraser University, Surrey, BC, Canada V3T 5X3, Phone +1 604/805-7112, FAX 604/437-7112, yvchen@sfu.ca

[2] Graduate Student & Research Assistant, Interactive Arts & Technology, Central City Tower, Simon Fraser University, Surrey, BC, Canada V3T 5X3, Phone +1 778/389-9008, FAX 604/437-7112, cherylq@sfu.ca

[3] Professor, Interactive Arts & Technology, School, Central City Tower, Simon Fraser University, Surrey, BC, Canada V3T 5X3, Phone +1 604/268-7501, FAX 604/268-7478, rw@sfu.ca

design could be adopted to visualize other information domains such as similarity ranking, co-citation in scholar literatures, online shopping systems, "find-similar-document" on WWW and groupware of awareness and coordination.

INFORMATION STRUCTURE IN A•VI•RE.

A•VI•RE (a Visual Rete, URL: http://www.avire.ca), a generic repository for visual material related to cultural disciplines, is designed as an interactive online space where users in different roles (such as curators, exhibitors, critics and viewers) work together to create a larger social entity. Users can upload *resources*, organize *exhibitions* and *annotate* resources. In A•VI•RE, each exhibition and resource has metadata as its attributes.

The information structure in the system has three primary object types:

- A resource can be an image, an video clip, a file of any type or even an annotation or an exhibition.
- An exhibition is a collection of exhibitions, annotations and resources.
- An annotation is a mixture of text and reference to exhibitions and resources.
- One resource can be used in multiple exhibitions. Every exhibition or resource may have one or more annotations. Annotations may also be annotated.
- Metadata exists on all of these object types.

In the system, one exhibitor may set up an exhibition, add resources and write annotations. Another exhibitor might use the same resources but interpret them from a different perspective. A•VI•RE thus provides a potentially complex information structure among exhibitions, resources and annotations. Users can potentially gain from multiple views and interpretations of A•VI•RE's objects. The problem is providing access to the web users have made. Our premise is that local links, that is, link paths no more than a few links in length, can provide meaningful access to a community of interpretation.

PROBLEMS OF VISUALIZING INFORMATION AND REVEALING RELATIONS IN A•VI•RE

In A•VI•RE, it is hard to provide multiple viewpoints of a resource. A resource can be collected in multiple exhibitions and quoted in multiple annotations, so it naturally bridge different interpretations and connect interesting ideas. The system should provide a panorama view of a resource from different perspectives, but its design hasn't met the goal.

It is also hard to see the relations among exhibitions. An exhibition might share resources with other exhibitions or belong to another exhibition as a resource. Although references information among exhibitions is valuable, it could not be explored in the original interface.

It is easy to find one resource or exhibition through key word searching, but hard to access specific relevant data elements. To search for a specific image among over 3000 images, the user has to bear heavy cognitive load and remember too much related information.

WHY TO APPLY GRAPH VISUALIZATION TO A•VI•RE

Metadata aside, a principal source of information about A•VI•RE's objects is what users have done with them. Amazon's model of recommendation based on the collective actions of buyers is strong evidence that using such information may be effective. NEAR visualizes local co-reference and inclusion relations in the A•VI•RE structure. Coreference is the phenomenon where two objects (exhibitions or annotations) both refer to or include the same resource. Inclusion refers to the relation between two collections (exhibitions or annotations) that exist when all members (or related objects) of the first are also members (or related objects) of the second. Using these two relations we can construct higher order relations such as similarity and mutual interests in a group. In A•VI•RE, interpretation connections between resources and exhibitions linked by annotations are valuable information because they provide later users with access to the carefully constructed thoughts of others. Our design aims to use specific local interconnections as a basis for overview and discovery.

KEY QUALITIES

Outlined here are several qualities we consider to be essential in effectively visualizing resources, exhibition and annotations, along with a brief discussion of their value.

- **Relations** provide context.
- **Interpretation** gives nuance.
- **Reuse** hints at importance.
- **Compactness** yields immediacy.
- **Interaction** highlights structure.
- **History** structures narrative.
- **Orthogonal codings** distinguish queries.
- **Chronology** renders the origin of ideas.

EXISTING RELATED PROJECTS & RESEARCHES

We review literatures from four perfectives: measuring similarity, tracing reference, group information sharing and awareness, and visualizing database information.

MEASURING SIMILARITY

Many applications require a measure of “similarity” between objects. One obvious example is the “find-similar-document” in the World-wide Web (Baeza-Yates & Ribeiro-Neto 1999). More generally, a similarity measure can be used to cluster objects. “Similar” users and items can be grouped based on the users’ preferences for collaborative filtering in recommender systems (Konstan et al. 1997, Shardanand & Maes 1995. Bibliometrics studies the citation patterns of scientific papers (or other publications), and relationships between papers are interred from their cross-citations. Most noteworthy from this field are the methods of co-citation (Small 1973) and bibliographic coupling (Kessler 1963). These methods have been applied to cluster scientific papers according to topic (Popescul et al. 2000). More recently, the co-citation methods have been used to cluster web pages (Larson 1996). However, few of these projects intend to present their final outcomes recognizably or efficiently. For example, SimRank (Jeh & Widsom 2002) described an efficient way to measure the structural-context similarity. But, the similarity outcomes are demonstrated through many simple node-link graphs, which are hard to understand at a glance.

TRACING REFERENCES

A•VI•RE is an information and reference archiving system. It archives the information of how resources are contributed, shared, interpreted and annotated as exhibitions or annotations, and also archives this reference information as new resources. Techniques used in email archiving system such as node-link graphs can represent references and demonstrate overviews of links (Ware 2004). Email conversations can be visualized through a mixed model of sequences and trees (Venolia & Neustaedter 2003), a circular graphical model (Hershkop & Stolfo 2004) or Thread Arcs (Kerr 2003). Compared with other node-link forms, Thread Arcs are elegant because they demonstrate how a simple node-link diagram can represent complex email relations clearly and smoothly. The fundamental goal of visualizing emails is to keep track of simultaneous email conversations. Thread Arcs provide a useful way for users to see the whole structure and background information.

Information Sharing and Awareness

In A•VI•RE, resources can be shared by multiple exhibitions and annotated in multiple annotations. AwareCo (Elliot & Carpendale 2005) is a visualization application to provide a means for awareness and coordination between family members with diverse schedules. Multiple schedules for family members are shown simultaneously. Sharing, conflicts and overlapping appointments are easy to spot. The design of how schedules share and overlap in AwareCo inspires NEAR's design in terms of exhibition overlap and resource share.

VISUALIZING INFORMATION IN A IMAGE DATABASE

In most image database systems, image data are put into categories related to metadata and keywords. The result of a query in the image database systems is usually a set of images, displayed in an Image Browser or shown in a two-dimensional grid of thumbnails (Ogle & Stonebraker 1995). For document browsing, the Document Lens (Robertson & Mackinlay 1993)) uses a focus+context technique to display the document of interest in detail while compressing the rest of the document space. Image browsing systems also display the entire space of images either without hierarchies (the default for most file mangers) or with some structural information provided by annotations (Kang & Shneiderman 2000). Apart from giving an overview of the entire document/image space, recent projects such as Concentric Rings (Torres et al. 2003) and MoireGraphs (Jankun-Kelly & Ma 2003) could highlight the relations among documents/images. But in their detailed examples, images are overlapped and relation lines are entangled with each other. It seems that this design issue of visualizing large amounts of resources has not been satisfyingly solved.

VISUALIZATION

To design a panel that demonstrates the relations among data elements in A•VI•RE, we consider two essential parts: nodes (resources, exhibitions and annotations) and their links. The design is based on the key qualities we have determined and refer to related projects we have reviewed. In Figure 1, the NEAR panel shows the following: a list of all resources that belong to the current exhibition; all related annotations; and all exhibitions that share any resources in the list. It is important to make the abstract image meaningful. So, exhibitions, annotations and resources are represented by graphic icons or thumbnails. Due to the limited capacity of human's visual working memory (Ware 2004), only 3~4 variations for each attribute are provided in the design.

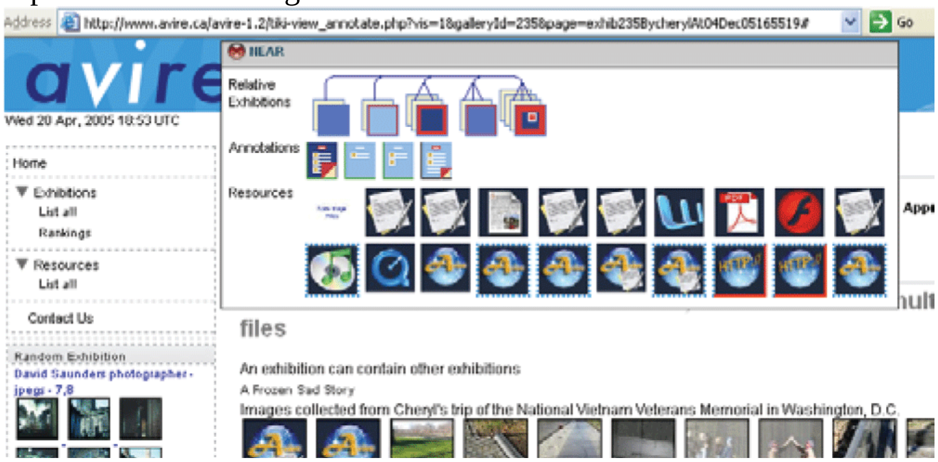

Fig. 1 Screenshot of the NEAR panel in A•VI•RE.

Variations multiply by the Cartesian product of the varying qualities. Orthogonality is important for user to distinguish different attributes for the large amounts of varieties (exhibition nodes and annotation nodes have 432 varieties, and resource nodes have 9 varieties to represent visit status and popularity). For example, different shapes and layers represent size and visit status. Blues of different saturations indicate popularity. Red color represents preattentive important elements such as current exhibition and implied relations.

NODES

Nodes of Resources

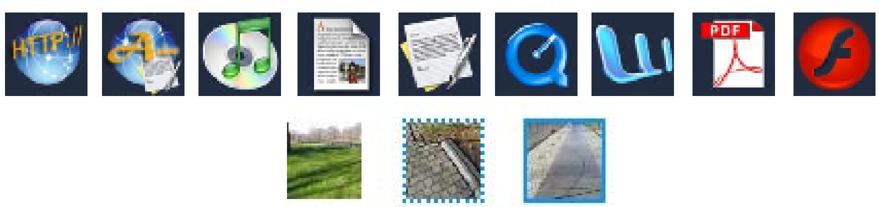

Fig. 2 Node of Multimedia resources and image resources
(Different border of the icon represents the frequency of the resource being shared)

Resources in A•VI•RE can be an image, a multimedia file, a website or any type of document. To meet the qualities of compactness and recognizability, the scale of thumbnail nodes should be small but easily recognizable. For files other than images, we adopt the design of the file icon from Windows and Mac OS X operation system. (Figure 2)

Nodes of Exhibitions

An exhibition is a collection of exhibitions and resources. Since exhibitions are actually document/image containers in the system, we designed the icon as shown in Figure 3.

Fig. 3 Examples of Exhibition Nodes.
Left Group: Nodes of different size (determined by number of resources).
Right Group: Nodes of different popularity (determined by the user visits)

From the exhibition nodes, following information can be represented: size of the exhibition (the number of pages in the node), popularity (different shades of cover color), chronology (from left to right in the order of time), reference (number of resources been shared, Figure 5), currency and visit status (Figure 6 & 8). To set a scale of measurement for a particular quality, we dynamically analyze all exhibitions and base rating on a normalized distribution.

Nodes of Annotations

Annotations combine text and resources into a linear narrative – in essence, they are papers and their thumbnail displays them as such. The qualities of content organization, chronology, popularity, currency are presented in Figure 5.

Fig. 4 Examples of Annotation Nodes
Left Group: Nodes of different organization (determined by number of resources & amount of annotation text). Middle Group: Nodes of different popularity (determined by user visits, color is darker means being more frequently visited). Right Group: Nodes of different importance (if the annotation is the default view)

LINKS AND RELATIONS

In NEAR, links are used to show the similarities, references and sharing among exhibitions, annotations and resources. Smooth continuous contours are used to connect nodes (Ware 2004). Any exhibition share the same resource/resources are linked. As in Thread Arcs (Kerr 2003), the amount of sharing is represented by the number of branches from one node (Figure 5), and icons are lined up in the order of creation time.

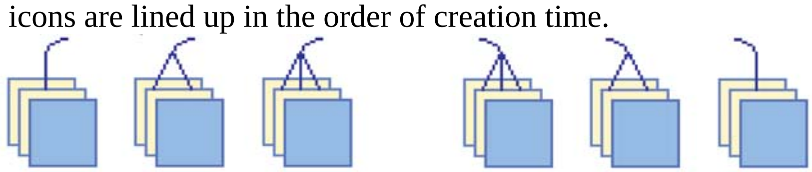

Fig. 5 Examples of Exhibitions Using Branched-Links to Share Different Amounts of Resources

Nodes can also be responsive when a related annotation or a shared resource is opened in the background browser window. Figure 6 shows the examples of an exhibition node and an annotation node which includes the currently opening resource.

Fig. 6 Examples of Exhibition Node and Annotations Node with Currency Relations

INTERACTION

Interaction is something that allows us to drill down and find more data about anything that seems important (Ware 2004). Ben Shneiderman also has called a "mantra" to guide visual information-seeking behavior and the interfaces that support it: "Overview first, zoom and filter, then details on demand (Shneiderman 1998)". In NEAR, information obtained through user's interaction is important. There are three kinds of interactions in NEAR: cursor over, click and double-click.

Since the relations among data elements are complex, "Brushing" (Becker & Cleaveland 1987) is adopted in NEAR interaction to enable visual linking of components of heterogeneous complex objects. When the mouse cursor brushes over a data element, the user can find more relations between data elements and read the description of the element. Sometime people like to move mouse when they read. To avoid flicking, a click on an element freezes the relation view. A click on empty space will re-active the mouse-over effect. Similar to any desktop applications, double clicking will forward the background page to display the full content of the element.

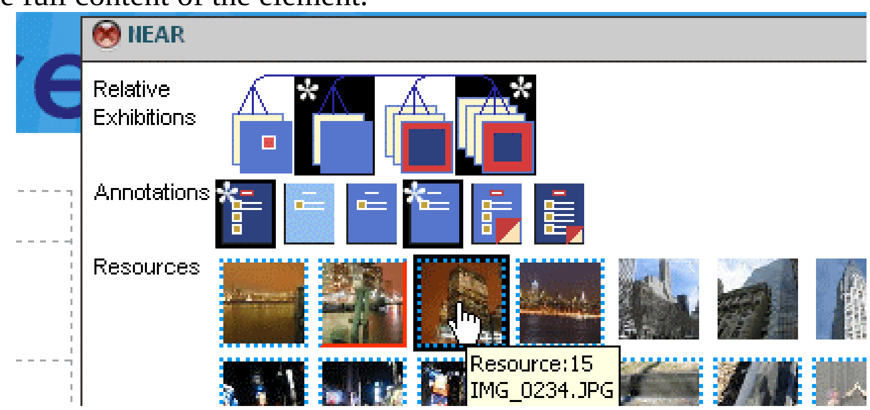

Fig. 7 Screenshot of NEAR panel (when the cursor is over a resource).
the * is added to this figure to indicate that so-marked items are colored red on screen

Figure 7 shows an instance of how NEAR is working in A•VI•RE: two exhibitions and two annotations are highlighted with red borders to indicate that they are sharing the highlighted resource. The screenshot shows the following: which exhibition(s) share this resource and which annotations quoted this resource. Cursor-over also shows some property information of this resource such as resource ID, name, contributor and comments. Similar things happen when the cursor is over an exhibition or an annotation. When the cursor is over an exhibition, all the shared resources and annotations would be highlighted. When the cursor is over an annotation, all quoted resources and related exhibitions will be highlighted.

Fig. 8 Examples of Nodes with Different Visit Statuses.
Left Group: Exhibition Nodes: unvisited, visited, just visited and current.
Middle Group: Annotation Nodes: unvisited, visited and current.
Right Group: Resource Nodes: unvisited, visited and current.

Furthermore, to make navigation easier, NEAR changes some quality of the node to display the visited status (Figure 8). Since the exhibition is at the top of the hierarchy structure, when an exhibition is changed, its related annotations and content resources will be changed correspondingly. The exhibition status of “just visited” makes it easier for the user to trace the relation between the current one and the previous one.

EVALUATION & DISCUSSION

We want to learn two things from the user evaluation: firstly, since every graphic node implies a lot of meaning and different relations, we want to optimize its design to maximally match user’s common senses. Secondly, we want to learn about the usefulness and effectiveness of NEAR on solving problems in a user’s navigation and interpretation.

Five people were invited to test the NEAR panel in different circumstances. They all used A•VI•RE before and found the original navigation paths not sufficient. We showed three participants the list of nodes examples and taught them the meaning of linkages before their navigation. The other two participants start to navigate without taking any tutorial. The result showed that the tutorial was essential because it took those two participants quite a while to understand the meanings of the graphic nodes.

All of the participants gave positive evaluations to the design. Three of them found it helps their understanding of resources and describes it as very useful in term of supporting navigation. However, they also provided critique on the panel design. Some important comments are as below:

- To the participants, the NEAR panel is more like a tool box instead of a visualization window. They want the panel can be movable and adjustable.
- One participant thinks that the high degree of detail in each icon precludes quick glances across the set to ascertain their meanings. Further, the high degree of detail also requires knowledge on the part of the viewer to ascertain the differences between the icons.
- Two participants (without tutorial) have different opinions on meaning of visit status of annotation nodes.

People from different culture backgrounds do interpret icons differently and use different graphic languages to describe the same object. We accept the diversity of interpreting nodes and made some changes on our original designs. There are still much to be improved such as the flexibility of the panel and the layout of node elements. The evaluation results show us clearly that the panel is useful and supportive but not self-expressive.

FUTURE APPLICATIONS

NEAR was developed to address the need for visualizing implied relations such as similarities, references and sharing among different data elements. We believe the techniques of NEAR can be applied to other applications with collection/resources structure. For example, Amazon.com has the function of suggesting new customers with previous customer’s shopping experience with lines of text and could be easily ignored by users. A visualization method of such of functions could be an interesting challenge.

NEAR principles can also be used to improve the applications of academic citation index. The ISI Web of Knowledge provides seamless access to current and retrospective multidisciplinary information from approximately 8,700 of the most prestigious, high impact research journals in the world (Thomson Scientific 2006). Its unique search method: *cited reference searching* helps users to navigate forward, backward, and through the literature, searching all disciplines. We believe our design of NEAR can help to represent those implied relations by graph visualization and enhance the navigation efficiency of cited reference searching and literature similarity comparing.

CONCLUSIONS

This paper describes the NEAR panel, a visualization technique that can display relationship and qualities simultaneously and reveal implicit relations such as similarity, inter-reference and sharing. NEAR explores the possibility of visualizing unstructured relations among structured data elements graphically in a visual material repository A•VI•RE. We present this visual panel to explore new navigation and communication approaches. By quickly scanning and interacting with NEAR, users can see not only implied relations but also potential connections among different data elements. Generally, our evaluation shows NEAR is a supportive navigation tool, but we also meet some unsolvable problems such as culture differences on icon interpretation. We plan to further develop this tool and implement the visualization design into other kinds of information repositories with collection/resources structure. We believe the idea of NEAR can bring users *near* to the implied relations of information regardless of whether the information has already been well-determined.

REFERENCES

Baeza-Yates, R. and B. Ribeiro-Neto. (1999). *Modern Information Retrieval*, Addison Wesley, Reading, Massachusetts.

Becker, R. A. and W. S. Cleaveland, (1987). “Brushing scatterplots”. In *Technometrics 29*, 2, 127-142.

Breese, J. S., Heckerman, D. and C. Kadie. (1998). “Empirical analysis of predictive algorithms for collaborative filtering”. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, Madison, Wisconsin, July 1998.

Elliot, K.and Carpendale, S. (2005), *Awereness and Coordination: A Calendar for Families*, research report, University of Calgary.

Herman, M. and M. Marshall, (2000). “Graph Visualization and Navigation in Information Visualization: a Survey,” in *IEEE Transactions on Visualization and Computer Graphics* Vol. 6, No. 1. 24-43.

Jankun-Kelly, T. J. and K-L. Ma, (2003). “MoireGraphs: Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes,” in *Proceedings of the 2003 IEEE Symposium on Information Visualization*, IEEE Computer Science Press. Seattle, Washington, 59-66.

Jeh, G. and J. Widom. (2002). “SimRank: A measure of structural-context similarity”, in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton*, Alberta, Canada, July 2002, 1-11.

Kang, H. and B. Shneiderman, (2000). "Visualization Methods for Personal Photo Collections: Browsing and Searching in the Photofinder," in *IEEE International Conference on Multimedia and Expo (III)*. 1539-1542.

Kerr, B. (2003). "Thread Arcs: An Email Thread Visualization," in *Proceedings of the 2003 IEEE Symposium on Information Visualization*, IEEE Computer Science Press. Seattle, Washington, 211-218.

Kessler, M. M. (1963). "Bibliographic Coupling between Scientific Papers", in *American Documentation*, Vol. 12, 10-25.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, L. L., Gordon, L. R. and J. Riedl. (1997). "GroupLens: Applying Collaborative Filtering to Usenet News," in *Communications of the ACM*, Vol. 40, No. 3, March 1997, 77-87.

Larson, R. R. (1996). "Bibliometrics of the World-Wide Web: An Exploratory Analysis of the Intellectual Structure of Cyberspace," in *Proceedings of the Annual Meeting of the American Society for Information Science*, Baltimore, Maryland, October 1996.

Li, W. J., Hershkop, S., and S. J. Stolfo, (2004). "Email Archive Analysis through Graphical Visualization," in *VizSEC/DMSEC'04*, Washington, DC, Oct. 29. 128-132.

Ogle, V. E. and M. Stonebraker, (1995). "Chabot: Retrieval from Relational Database of Images". In *IEEE Computer*, 28(9). 40-48.

Popescul, A., Flake, G., Lawrence, S., Ungar, L. H. and C. L. Giles. (2000). "Clustering and Identifying Temporal Trends in Document Databases", in *Proceedings of the IEEE Advances in Digital Libraries*, Washington, D.C., May 2000.

Robertson, G. G. and J. D. Mackinlay, (1993). "The Document Lens", in *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST93)*, Visualizing Information, 101-108.

Shardanand, U. and P. Maes, (1995). "Social Information Filtering: Algorithms for automating 'word of mouth'", in *Proceedings of the Conference on Human Factors in Computing System*, Denver, Colorado.

Shneiderman, B. (1998). *Designing the User Interface.* Addison-Wesley, Reading. MA.

Small, H. (1973). "Co-citation in the scientific Literature: A New Measure of the Relationship between Two Documents", in *Journal of the American Society of Information Science*, Vol. 24, 265-269.

Sweller, J. (1998). "Cognitive load during problem solving: Effects on learning," in *Cognitive Science*, Vol. 12. 257-285.

Thomson Scientific. (2006). *ISI Web of Knowledge Home*, [Online], Available: http://isiwebofknowledge.com/index.html (Accessed: 2006, Feb.20th).

Torres, R. S., Silva, C. G., Medeiros, C. B. and H. V. Rocha, (2003). "Visual Structures for Image Browsing," in *CIKM'03*, November 3-8, New Orleans, Louisiana, pp. 49-55.

Venolia, G. D. and C. Neustaedter, (2003). "Understanding Sequence and Replay Relationships with Email Conversations: A Mix-Model Visualization," in *CHI 2003*, ACM press. Vol 5. Issue 1. 361- 368.

Ware, C. (2004), *Information Visualization: Perception for Design* (2nd Ed.). Morgan Kaufman, San Francisco, CA.