Information visualization and retrieval using stereoscopic display of document and term relations

Richard H. Fowler and Wendy A. L. Fowler

Department of Computer Science University of Texas - Pan American Edinburg, TX, USA 78539

1. ABSTRACT

Scientific visualization brings human perceptual processes to bear in organizing and understanding data about physical phenomena. Information visualization has a similar goal for elements in often semantic domains. Document Explorer is an information visualization and retrieval system that displays three dimensional associative network representations of document and term relations. The system's networks of documents can be constructed from existing measures of association, e.g., link structure in hypertexts, or derived by the system using content similarity among documents. The system also maintains a network of terms which is available to the user for query formulation.

Recently, we have added stereoscopic display of the system's several networks to enhance users' perception of structure. Additionally, users' head movements are tracked and used to change point and angle of view to further enhance structure perception and allow additional user interface mechanisms for navigation in three dimensions. By viewing and interacting with these networks using head-tracked stereoscopic display the user is better able to perceive relationships in the networks and, thus, better able to distinguish clusters of documents and categories of terms during the information retrieval process.

Keywords: information visualization, interaction, stereoscopic display, information retrieval, document retrieval, document visualization, human computer interaction

2. INTRODUCTION

Scientific visualization brings human perceptual processes to bear in organizing and understanding data about physical phenomena. Information visualization has a similar goal for elements in often semantic domains. Additionally, information visualization systems have typically focused on tasks in which the user is engaged in information seeking behavior. This has led to closely related goals for providing users interactive environments tailored to the tasks which they are performing. The visual representation of structure to facilitate understanding of document collection organization, e.g., clustering, and assist in finding information has a relatively long history in information retrieval^{1, 2}. More recently, the development of new visualization and interaction techniques coupled with a deepening understanding of the paradigm of information visualization has led to a significant increase in the number of systems for visually representing the large text-based information spaces common in information retrieval. Many of the general challenges of information visualization are evident in visually based information retrieval systems, e.g., perception of structure in a document collection and orientation to global structure while examining local detail. Other issues are more focused on the information retrieval process, e.g., formulating and modifying queries and the role of the system's keyword vocabulary. A number of systems have demonstrated the utility of visual adjuncts to conventional search³, as well as new ways of conceptualizing the document retrieval process which are uniquely suited to visual representation⁴.

Information visualization and scientific visualization both seek to provide understanding and insight using visual representations. However, information visualization is often faced with challenges not encountered in scientific visualization. For scientific visualization the visual representation is derived from a physical phenomenon and there typically exists a natural visual representation based on temporal, spatial, or other physical properties of the phenomenon. For information visualization elements to be visualized typically have no physical component which might naturally supply a basis for visualization⁵. Instead, elements to be visualized often only have semantic properties having no spatial analog from which to create a visual representation. Any spatial ordering for semantic relations must be created as part of the information visualization process. From this perspective, visualization of semantic information spaces must provide: 1) spatialization of the abstract data, which may entail both a) data organization and b) derivation of a visual spatial representation of the data, 2) presentation of the spatial representation in a display space for user interaction and viewing, such as Euclidean two- or three-dimensional space or a distorted space, and 3) techniques and tools for user interaction with the visual representation.

One aspect of spatialization is the process of organizing the information, or data, which is to be visualized, e.g., extracting keywords to provide a multidimensional vector space. In some applications the data organization is available in the data to be displayed, e.g., a file system hierarchy. Spatialization also creates a viewable representation, a spatial-visual representation. It typically is at this step that the final spatial nature of a data organization is derived by a transformation to a representation useful for information visualization. The term spatial-visual emphasizes that it is a form which is both spatial and also visual, i.e., viewable in three dimensions. This distinguishes it from representations which may be spatial, e.g., vector space, which is spatial in n dimensions, but which are not viewable without a reduction of dimensionality through a mapping into a visualizable, typically three dimension, space. The derivation of a spatial-visual representation is at the core of information visualization.

Display space in this context is the visual space used for presentation to the user. The separation of display space and a spatial-visual representation reflects the practice that a system or technique is really the combination of spatialization methods, together with a particular display space and set of interaction methods which create the unique technique. Most frequently, display spaces are simply three dimension perspective projection, e.g., the Xerox PARC cone tree⁶ is radial tree layout in three dimensions presented in Euclidean three-dimensional space.

The third set of components is user interaction mechanisms. These are the ways of interacting with the spatial display and sometimes interactively changing the spatial-visual representation, e.g., changing the fisheye view filtering⁷ or expanding and collapsing nodes in a network. Most systems include such user-manipulable filtering mechanisms which affect the spatial-visual representation to varying degrees, e.g., recalculating degree of interest or simply hiding or showing elements.

3. DOCUMENT EXPLORER: AN INFORMATION RETRIEVAL AND VISUALIZATION SYSTEM

3.1 System overview

Document Explorer is an information visualization and retrieval system using associative network representations of document and term relations for user display and interaction. Recently, we have added stereoscopic displays of the system's several network displays to enhance users' perception of structure. Additionally, users' head movements are tracked and used to change point and angle of view to further enhance structure perception and allow additional user interface mechanisms for navigation in three dimensions.

The system is typically used to display a network of documents which is derived either from measures of association, or content similarity, among documents or from existing relations, e.g., hypertext links. Figure 1 below shows some of the system components. Most recently, the system has been used together with emerging Internet tools to display networks of World Wide Web (WWW) documents based on both semantic and link properties ^{8,9}. The system also maintains a network of terms which is available to the user for query formulation. For display, network nodes are positioned in three dimensions using a graph layout algorithm based on a spring metaphor in which semantic strength of association between elements is used to determine spring strength. Other parameters, such as spring length and elastic properties, are also varied, either automatically or under user control, in arriving at the final node positions to visually reveal clustering and connectivity among documents.

Display and interaction mechanisms focus on facilities for orientation to global structure together with navigation and retrieval tools for exploring local detail. Overview diagrams track the view volume of the detailed view, and cluster centroid nodes are distinguished for display and navigation. Nodes in view at a particular time can be determined both by the standard 3D viewing projection and user controlled tools to vary display density based on network structure. For example, the user can use fisheye viewing techniques to control density by specifying threshold values for the degree of interest or path distance from the focus node. Tools are provided for fluid navigation in the space, such as zooming to documents or document clusters selected manually, or through a natural language questions matched to a document cluster in which node coloring are used to indicate parts of the network which have been viewed or are of high similarity to the user's queries.



Figure 1: Document Explorer screen. The content based document network is shown in the main window with overview diagrams to the left. Windows below the document network show visual bookmarks the user has set. At the upper right is the user's natural language query and its visual representation.

3.2 Organizing text-based information for visualization using associative structures

For the system we have implemented there are three reasons for using statistically-based associative structures. One reason follows from the view that information retrieval systems should supply the user with a variety of tools and retrieval techniques. Statistically-based associative information structures provide one class of retrieval tools that can complement other retrieval aids. For example, an associative thesaurus based on term co-occurrence in documents presents a structure of term relationships quite different than presented in the thesaurus showing term hierarchies. The associative thesaurus can encourage browsing and exploration, as well as bring the user's own associations into play. For information needs in which the user is not familiar with the domain, and indeed may not even know what his or her information needs are, the associative structures provide one means to explore and gain information to better define the information need.

A second reason for using statistically-based associative structures is the desire to have a representation that can be derived automatically in an interactive system, rather than through knowledge-engineering efforts. Associative structures can also serve as one component of a hybrid system incorporating both deep and shallow representations¹⁰. Networks are naturally represented visually and can provide a common representation for several information retrieval system components.

3.2.1 Extracting content by deriving document and term associations

Several networks are available to the user, as shown in Figure 2. The network of documents is typically used to show relations among documents by deriving and displaying an associative network based on similarity among documents, as detailed below. The current system implementation operates with document networks up to a few thousand nodes. The system can also be used to display existing network structures, e.g., existing hypertext or WWW links. The network of terms, or associative the-saurus, is based on similarity among terms in a particular document set and is typically less than a thousand terms. Single document networks supply visual representations of relations among terms within a document. These relatively small networks supply an adjunct to the text of a document and can be used to rapidly scan a list of documents. Finally, the user's query can be represented as a network for modification by direct manipulation using terms shown on the display.

For each database the system uses a separate set of terms that includes the most frequently occurring word stems, excluding function words. For some forms of retrieval this simple procedure suffers from the limitation that frequently occurring terms have relatively little value for discriminating among documents¹¹. However, one function of the associative thesaurus is to give a picture of the all concepts in a document set. The most frequently occurring terms tend to be general terms that provide useful information about the domain of the document collection.

To derive distances among documents used to construct the network of documents content similarity among documents is used, as measured by the number and frequency of terms common to documents. To derive the distances between terms used to construct the other networks, text is analyzed by first finding the sequence of term stems present in the text. This sequence is used to assign distances between terms based on lexical distance and co-occurrence in syntactic units with a metric similar to that used by Belkin and Kwasnik¹². Term pair similarity is calculated as the sum of values added when terms are adjacent, or occur in the same sentence, paragraph or document. These similarities provide the associations used in deriving the networks displayed by the system.

3.2.1 Forming associative networks

Pathfinder networks¹³ (PfNets) have a number of properties useful in creating associative structures to be visually represented¹⁴. PfNets are a class of minimum cost networks which add a path length constraint to more well known minimum cost networks. For example, MCSTs are constructed with only the constraint that the sum of weights is a minimum, and, of course, there is no cycle (so the number of edges is a minimum). The Pathfinder algorithm adds a constraint that the longest path considered in deriving a connected network from a set of distance data be less than or equal to some maximum, q. So, with small values of maximum path length the networks can be quite dense. Additionally, PfNets incorporate a data transformation when specifying the network. The Minkowski *r*-metric, *r*, gives the power to which a date element is raised, and its value may vary from to 1 to ∞ . The typical notation for a Pathfinder network is PfNet (r, q), where r is the value of the Minkowski r metric and q is the longest length of paths considered in constructing the network. PfNet (∞ , n-1), where n is the number of nodes in the network, is the union of all MCSTs, i.e., the MCST with all edges with tie distances retained. As such, it is typically quite sparse and supplies a representation closely related to single link clustering. This sparse network is used in the Document Explorer for global navigation and overview displays. With the data sets and layout algorithm we have explored it has produced visual representations in which "clusters" of elements are quite evident.

Another useful property of PfNets is that as either r or q is manipulated, networks produced using smaller values are subnetworks of any network produced with a larger value. This inclusion property provides the basis for systematically varying the density of networks derived from a common data set. In the Document Explorer networks with differing values of qare stored, and the user can switch between them in order to vary density, i.e., to add links which show other "weaker" relations than are evident in the most sparse network. Varying density in this way allows more detail about the data set to be revealed, while maintaining the links that were shown in the sparser PfNet.



Figure 2. Document Explorer networks. Text is the basis for each of the four types of networks displayed by the system. Depending on the network, distances are derived among documents or among terms. Minimum cost networks are derived from the distances. Nodes of the network are positioned in two or three dimensions and the network displayed to the user.

3.3 Creating visual representations of the networks

Creating displays which simultaneously 1) reveal structure, 2) allow rapid, effective navigation, 3) provide orientation of the part being viewed in detail to the overall structure, and 4) support techniques for direct manipulation of objects used in query formulation are among the challenges for the system's visual displays. To automatically create displays which reveal structure, nodes of the associative networks are positioned using a graph layout algorithm based on a spring metaphor which extends the 2D algorithms of Eades¹⁵ and Kamada and Kawaii¹⁶ to 3D. Nodes are considered as connectors and the length and strength of springs among connectors is derived from path distances. Nodes are positioned at the points which minimize the total energy in the system of springs as the nodes are allowed to vary in three dimensions¹⁷.

3.4 Navigation and interaction tools and techniques

Several viewing tools are available to the user. Standard interaction techniques using a six degree of freedom input device (Spaceball) are used to move through the networks' 3D space. Nodes in view at a particular time are determined both by the standard viewing projection of the three-dimensional space and user-controlled tools that vary display density based on network structure. Most simply, the user can select a node and expand or collapse connected nodes. The user can enter the asso-

ciative thesaurus from any node displayed, e.g., from a query or document network. Similarly, any document, e.g., from the results of a search, can serve as an entry point into the network of documents.

Display mechanisms for changing network views facilitate orientation to the overall structure while examining local detail. The availability of two networks which differ in density, as described above, supplies a type of network abstraction enhancing users' perceptual extraction of network structure, and also provides a mechanism for navigation. The sparse network is essentially a tree, and so has the fewest links necessary to connect all nodes. This sparse network is used to position nodes in 3D space. The layout of nodes on this basis supplies a characteristic clustering of nodes in which structure is relatively easy to perceive, such as shown in the leftmost bookmark of Figure 1. However, browsing in the derived semantic space of documents can be enhanced by exploring relations which are relatively weak. Such relations are shown in the denser networks. The user usually finds a single document or document region by browsing or traversing links from a known document in the sparse network, and then makes weaker links visible. These links are of a different color and the point at which the selection is made is identified by a marker in the network which is relatively large and can be easily zoomed back to.

3.4.1 Overview Diagrams

Overview diagrams¹⁸ are a common means of supplying a user with 1) knowledge about the organization of the complete network, 2) a means for navigating the network, and 3) orientation within the complete network. In overview diagrams a small number of nodes, selected to provide information about the organization of the complete network, are displayed to the user. Additionally, the nodes typically provide entry points for traversing the network. These nodes provide orientation by serving as landmarks to assist the user in knowing what part of the network is currently being viewed.

In the document collections we have used, PfNets derived for associative thesauri and networks of documents have a characteristic structure. There tend to be a small number of nodes that have many nodes directly connected and there are relatively short paths between these highly connected nodes. There are relatively few nodes of high degree and the diameter of the network is small. This form of network suggests criteria for selecting nodes to include in overview diagrams. Overview diagrams displayed by the system include those nodes of highest degree in the complete network. The terms selected for overview diagrams of associative thesauri tend to be general terms that provide a guide to the content of the database; for a network of documents overview nodes tend to be broad, general articles about a subdomain. They are landmarks in that they supply information about both the content and structure of the database. High degree reflects frequent co-occurrence with many other terms or documents.

The overview diagram can also be used for navigation. By selecting one of the overview nodes the main window is zoomed to a display centered on the selected node. The location of the subgraph view volume in relation to the overview diagram is indicated by shading any node in the overview diagram that is displayed in the detailed view. If no term from the overview diagram is present in the detailed view, the overview term closest to the center of the main window display is shaded.

3.4.2 Other navigation and orientation tools

A number of navigational tools allow users to return to previous viewing points or provide information about the context of the current viewing position. Visual bookmarks, shown in the bottom of Figure 1, can be created by the user at any point in his or her exploration of the network. The bookmark is shown as a snapshot of the user's view. Selecting a bookmark returns the user to the viewpoint when the bookmark was set. Visual anchors, also shown in Figure 1, are displayed as colored arrows and can be created at any time. Selecting an anchor also returns the user to the viewpoint at the point the anchor was set.

Document Explorer signposts supply context information about a network region when it is being viewed in detail. Selecting nodes in the overview is often used as the initial method of exploration, and in practice users quickly become familiar with the relatively small set of overview document labels and terms. Signposts exploit users' familiarity with overview nodes as part of a mechanism which provides global context while viewing a detailed local region. A signpost is a 3D arrow, similar in visual form of an anchor, but with each arm of the arrow labeled with the nearest overview node. The labeled arms of the signpost show the direction of the nearest overview node. In principle the same orientation information about the location of the detailed view in relation to the complete network is available through the relation of the overview window to the subnetwork being viewed in detail. Yet, the signpost provides a complementary form of this information, information about the global context (the location of the overview nodes) presented within the detailed view.

Further display and orientation facilities are available. Link and node colorings indicate parts of the network of high similarity to the user's query. A backtracking facility allows users to retrace the path taken to arrive at particular viewpoint. Users can select a node representing a single document or term and expand or collapse connected nodes. Users can enter the associative thesaurus from any node displayed, e.g., from an individual document network or the query. Similarly, any document, e.g., from search results, can serve as an entry point into the network of documents. All navigation maintains fluid movement in the space, always zooming to new viewpoints.

3.5 Query formulation and modification

Typically, the user enters a natural language statement of information need to begin the browsing and retrieval process. The system converts the natural language to a weighted vector representation, and performs conventional weighted vector matching to rank documents in the document set. The user can then select a document from the ranked list as the entry point in the network of documents. By doing this the viewpoint is positioned near the document selected and the document is distinguished by color.

The association map of keyword relations supplies a visual representation of term space, and can be used to assist users in formulating queries. The relations among terms as shown in an association map are typically quite different than relations found in a conventional thesaurus and serve as an alternative ordering of term relations. The system provides direct manipulation facilities for constructing queries, as shown in Figure 1. A visual, manipulable representation of the user's query is displayed together with its original natural language form. This visual representation can be changed by moving terms from the term network or a single document network into the query window and connecting them to the query graph. The vector space retrieval is based on the query graph and is performed by converting it to a weighted vector representation.

4. HEAD-TRACKED STEREOSCOPIC DISPLAY IN DOCUMENT EXPLORER

Recent attention to immersive interfaces has focused on head mounted displays (HMDs). Such displays typically provide the user a wide field of stereoscopic view and use head position tracking to determine view. Though HMDs can be quite effective in evoking the feeling of presence necessary for total immersion, the apparatus and restricted view preclude their use in most environments. Further, the relatively low resolution of current HMDs makes text display difficult.

Head-coupled stereoscopic (HCS) viewing offers an alternative to HMDs. It supplies a relatively immersive interface, yet allows the user to work comfortably at a desk and switch among tasks. HCS viewing can also provide the resolution needed to display the large quantities of text typical of document retrieval tasks. In HCS 3D objects appear to be positioned in a volume bounded by the screen perimeter and extending from somewhat in front of the screen to somewhat behind the screen. Much of the utility of 3D display is obtained from the user's active interaction which changes the viewpoint and allows the user to confirm or disconfirm perceptual hypotheses. In HCS simple head movements easily and naturally change viewpoint to "look around" the objects. This viewing paradigm has been called "fish-tank virtual reality" and human factors experiments indicate enhanced perception and interaction in 3D tasks for HCS viewing compared to flat screen presentations¹⁹. More recently, Ware and Franck²⁰ performed a series of studies in which enhanced perception and interaction in networks similar to those used by Document Explorer was demonstrated.

The HCS system we have used employs LCD shutter glasses for stereoscopic viewing and untethered head position tracking. The enhanced 3D perception and interaction afforded by this display addresses the needs in our system to aid the user in perceiving structure and interacting with the system's large networks. This is most evident in viewing large networks of documents in which clusters of documents are more readily perceived and user interaction benefits from enhanced depth perception. The system also provides sufficient display resolution for text-based tasks typical of document retrieval. In using the various system displays the user moves between stereoscopic 3D displays for some tasks, e.g., viewing the document collection and term networks, and 2D viewing, such as reading a document abstract, without removing any apparatus.

5. CONCLUSIONS

Though many of the display and interaction techniques employed in information visualization and visually based document retrieval have been known for some time, their implementations in systems which support the full range of facilities needed for information retrieval is relatively new. Document Explorer is one recent system which explores the design space of visually based information retrieval systems. Recent systematic study of the ways in which stereoscopic display can facilitate perception of large network structure has assisted in design of the system. Together, the use of stereoscopic display of large information structures coupled with appropriate interaction techniques can facilitate users' utilization of large information spaces in information retrieval.

6. ACKNOWLEDGEMENTS

Work on this project has been supported by NASA grants NAG9-551 and NAG9-842.

7. REFERENCES

- 1. L. Doyle, "Semantic roadmaps for literature searchers", J. A. C. M., 8(4), pp. 367-391, 1961.
- 2. D. Crouch, "The visual display of information in an information retrieval environment", *Proc. ACM SIGIR*, pp. 58-67, 1986.
- 3. A. Spoerri, "InfoCrystal: A visual tool for information retrieval", Proc. IEEE Visualization '93, pp. 150-157, 1993.
- 4. K. A. Olsen, R. R. Korfhage, K. M. Sochats, M. B. Spring, and J. G. Williams, "Visualization of a document collection: The VIBE system", *Info. Processing and Management*, **29**(1), pp. 69-81, 1993.
- 5. N. D. Gershon (Panel chair.), "Information visualization: The next frontier", *Proc. ACM SIGGRAPH* '94, pp. 485-486, 1994.
- 6. G. G. Robertson, S. K. Card, and J. D. Mackinlay, "Information visualization using 3D interactive animation." *Comm. ACM*, **36(4)**, pp. 57-72, 1993.
- 7. G. W. Furnas, "Generalized fisheye views", Proc. Human Factors in Comp. Sys., pp. 6-23, 1986.
- 8. R. H. Fowler, A. Kumar, and J. L. Williams, "Visualizing and browsing WWW semantic content", Proc. IEEE/ACM Conf. on Emerging Tech. and Apps. in Comm., pp. 110-113, 1996.
- 9. R. H. Fowler, W. A. Fowler, and J. L. Williams. "3D visualization of WWW semantic content for browsing and query formulation", J. Network and Comp. Apps., 20(2), 1997
- 10. W. B. Croft and R. H. Thompson. "I3R: A new approach to the design of document retrieval systems", J. Am. Soc. Inf. Sci., 38, pp. 236-261, 1987.
- 11. K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval", *J. Doc.*, **28**(1), pp. 11-20, 1972.
- 12. N. J. Belkin and B. H. Kwasnik. "Using structural representations of anomalous states of knowledge for choosing document retrieval strategies", *Proc. ACM SIGIR*, pp. 11-22, 1986.
- 13. D. W. Dearholt and R. W. Schvaneveldt, "Properties of pathfinder networks". In R. W. Schvaneveldt (Ed.), *Pathfinder* associative networks: Studies in knowledge organization, pp. 1-30, Norwood, NJ: Ablex., 1990.
- 14. C. Chen, "Structuring and visualising the WWW by generalized similarity analysis", Proc. Hypertext '97, 1997.
- 15. P. Eades, "A heuristic for graph drawing", Congressus Numerantium, 42, pp. 149-160, 1984.
- 16. T. Kamada and S. Kawai, "An algorithm for drawing general undirected graphs", Info. Proc. Letters, 31, pp. 7-15, 1989.
- 17. A. Kumar and R. H. Fowler, *A spring modeling algorithm to position nodes of an undirected graph in three dimensions*, Dept. of Comp. Sci. Tech. Rep., Univ. of Texas Pan American, Edinburg, 1994. http://www.cs.panam.edu/info_vis.
- 18. M. D. Apperly, I. Tzavaras, and R. Spence, "A bifocal display technique for data presentation", *Proc. Eurographics*, pp. 27-43, 1982
- 19. K. W. Arthur, K. S. Booth, and C. Ware "Evaluating 3D Task Performance for Fish Tank Virtual Worlds". ACM Trans. on Info. Sys., **11(3)**, pp. 239-265, 1993.
- 20. C. Ware and G. Franck, "Evaluating stereo and motion cues for visualizing information nets in three dimensions", *ACM Trans. on Graphics*, **15(2)**, pp. 121-140, 1996.