3D Visualization of Document Retrieval Spaces using VRML

Michael J. Nelson
mnelson@julian.uwo.ca
Faculty of Information and Media Studies
Middlesex College
University of Western Ontario
London, Ontario N6A 5B7
Canada

Abstract

With the advent of more powerful workstations on the Internet capable of fast graphical displays, the possibility of using a three dimensional visualization of a database of documents is greater than ever before. Much previous work on the visualization of documents has concentrated on the two dimensional representation of document spaces. The advent of the Internet and VRML (Virtual Reality Modeling Language) has given us an easier method for three dimensional visualization by using the client/server model in a distributed way to display models of document space. Since VRML was originally designed to build models for worlds with everyday 3D solid objects such as trees, chairs, cars, buildings and people or even cartoons, this paper investigates the capabilities of VRML such as zooming, objects shapes, colors and text labels for building an interactive model of a document space. An evaluation of the some possible 3D models and their metrics, such as Salton's vector space model, and their appropriateness for 3D world implementations is also included. A front end for a retrieval system is being built to represent the retrieved set in a 3D space where the user can manipulate this space using a standard VRML plugin to a standard World Wide Web browser.

Introduction

Most searching and retrieval of documents is achieved in text mode by typing in several keywords for a query the receiving a list of documents back (sometimes in ranked order). This is a one-dimensional view of a retrieved set of documents which actually form a complex web of documents with many relationships between them. With all the graphical user interfaces on the Internet, it seems natural to seek a more graphical method of displaying collections of documents or retrieved sets of documents. A screen is essentially two dimensional, so the next logical step is to create maps of documents in 2D and several projects have taken this approach (see below for descriptions). Another possibility is to model the natural three dimensional (3D) world on the screen and simulate 3D by using the usual tricks of perspective and changing viewpoints. The advantages of these visualization approaches are summarized in Table 1. Two of the categories that especially relate to traditional information retrieval are "reduced search" and "enhanced recognition of patterns" which relates directly to the browsing model of finding documents. Thus a 3D world allows for much more complex representations and more interactive browsing of a collection or answer set. For more discussion of the cognitive and perceptual aspects of visualization see the short summary in chapter 1 of Card, McKinlay and Schneiderman (1999).

Table 1. How information visualization amplifies Cognition
(from Card, McKinlay and Schneiderman, 1999, p16)

Increased Resources
- High-bandwidth Hierarchical Interaction
- Parallel perceptual processing
- Offload work from cognitive to perceptual system
- Expanded working memory
- Expanded storage of information

Reduced Search
- Locality of processing
- High data density
- Spatially indexed addressing

Enhanced Recognition of Patterns
- Recognition instead of recall
- Abstraction and aggregation
- Visual schemata for organization
- Value, relationship, trend

Perceptual Inference
- Visual representations make some problems obvious
- Graphical computations

Perceptual Monitoring

Manipulable Medium

Why VRML?

Most of the systems reviewed below have used customized interfaces to implement the visualization of text and documents spaces. In order to make visualization systems more standard and to make it easier to send simulated worlds over the WWW, the VRML (Virtual Reality Modeling Language) was developed for rendering 3D worlds of any kind. VRML works by sending a text based description of a 3D world over the Internet to a browser addin which then displays a 3D view of the world and allows the user to freely navigate through the simulation using the mouse. Most the initial development of the language was at SGI Corporation. Unfortunately, SGI has removed its direct support but VRML 2.0 still lives on and much of their work has been taken over by Cosmos Corporation. There is also an ISO standard which is often referred to as VRML-97. The development of standards is currently managed by the Web 3D Consortium (http//:www.vrml.org), a group which includes the Visual Information Technology Group at NRC. The Web 3D Consortium is currently developing the next standard called VRML-NG (Next Generation).

Why use the VRML language instead of some other way of viewing 3D representations of document spaces? First of all it is a standard, secondly there are millions of viewers already installed on workstations around the world and thirdly, no graphics development needs to done to implement a 3D document space viewer. VRML files sent to the viewer describe the 3D world in a text file which contains an object-oriented descriptions of objects and their properties. Thus the retrieval system can automatically create the VRML from of the documents in the database. To

implement such a system we need to evaluate the various techniques used in VRML descriptions to have some idea of the possibilities for representing document spaces in 3D. Before looking at the specifics of VRML the more general questions of document space representation and visualization are reviewed.

Models of document spaces.

One dimensional models of documents are very common. The ranked list of documents produced by a retrieval system, a bibliography and a library classification are all designed to produce lists of documents in a single dimension.

At the other extreme the most common model used in information retrieval is the vector space model (Salton and M. McGill, 1983). In this model each index term or word used to index documents in the database is a dimension and a document is represented by a point in this space. This gives a vector space of thousands of dimensions. Representing this visually in two or three dimensions in a meaningful way is very difficult and the most common solution resorts to some form of document clustering.

One way to reduce the number of dimensions is to use one of the methods use in statistical analysis related to principal components and factor analysis. In fact Deerswater et. al. (1990) tried this and called their method *latent semantic indexing*. They reduced the space to approximately one hundred orthogonal vectors which were used to represent documents for retrieval purposes. This work was aimed at better retrieval using the traditional text query and list of documents returned and did not try to use a visual representation. This still leaves about one hundred dimensions but they did show a 2D example of nine documents and 12 index terms. Another problem with this method is that it is computationally intensive and does not scale well to large databases.

A recent improvement on this method comes from the WEBSOM project (Honkela et.al., 1996; Kohonen, 1998; http://websom.hut.fi/websom/). SOM stands for self organizing maps which are clickable maps of various document spaces in 2D which form a hierarchy of maps, that is clicking on the highest level map displays a new map on a larger scale with more detail until you get a list of documents at the end. Each map is colored to show the density of documents near each keyword.

A different type of 2D system is the Envision system (Nowell et.al., 1996). This system allows the user to display documents with two different axes of their choice. Some of the possibilities are: author, relevance, year of publication, etc. Documents are represented by colored icons which can encode factors not included in the original two axes. A user questionnaire rated the system highly.

Another 2D approach is the VIBE system (Korfage , 1991; Olsen et.al. 1993) which displays the output from a text query where the terms in the query become vertices on the screen (three terms gives a triangle) and documents as points located by their similarity to each term. For example, if a document only relates to two terms on the screen with equal weight the document point will be half way between the two terms. There are several different variations of this system developed at the University of Pittsburg.

Hemmje, Kunkel and Willett (1994) generalized this idea to a 3D system they call LyberWorld. The points from 2D became points on a sphere which they called relevance spheres. The documents are then distributed throughout the sphere according to similarities between documents and similarity of documents to terms on the sphere. This work is especially interesting because it relates the capablities of the system back to the basic functions of retrieval as outlined by Agosti, Gradenigo and Marchetti (1991). They claim their 3D approach supports the *semantic association, associative reading, sequential reading, navigation,backtracking* and *search strategy development* functions.

Another approach to 3D representation is to model the document space as a relief map with mountains and valleys often called a *landscape.* This is the method chosen by Chalmers (1994) who characterizes his approach as "2.1"D and called his system *Bead.* Documents are placed on the map according to similarity measures based on word occurrences but no details were given. Users were allowed to type in keywords and the matching documents represented by small triangles were highlighted by color. More information such as keywords and titles are displayed when the user clicks on a document.

A very interesting paper is one by Wise et.al. (1995) who use both a starfield (which they term a galaxy) and a landscape approach which they call a *themescape.* Their system is call SPIRE, the Spatial Paradigm for Information Retrieval and Exploration. They claim that each approach has unique advantages for seeing patterns in document collections. One test database was one weeks worth of CNN stories. One of the techniques they use is a *temporal slicer* which shows documents for different dates in an animation.

One of the few papers found to discuss the use of VRML in the context of document searching is by Rohrer and Swing (1997). They discuss several database and searching applications including document searching. They describe a prototype system for WWW searching which organizes the ranked results from a search engine into a 3D VRML world. The individual items are organized into clusters and the distance from the origin is related to the rank of the item in the list. Initially each cluster was represented by a sphere and as the user approaches individual documents in the cluster become visible. Details of how this was accomplished are not given.

Representing Document Spaces in VRML

To demonstrate the capabilities VRML for representing document spaces a simulation of a small set of documents was created. There are many choices to be made when choosing a 3D representation. One of the most important is the metaphor for representation. Some possibilities described in the articles reviewed include the geographic terrain (or landscape), the starfield of bright objects in a dark space, a 3D room and points on a sphere. VRML is very flexible and could represent any of these metaphors very easily. My system will start with the starfield metaphor.

Another choice is how to represent the documents, what shapes and colors, individual documents or clusters, with how much detail. The choice made for this system was to represent clusters by spheres with the size of the sphere related to the number of documents. Colors where chosen to be bright and different to aid in orientation and navigation. To make the 3D world more efficient, individual documents are not represented directly in the world. The user can access a list of the documents by clicking on a cluster. This is implemented by using a VRML feature called a *touchsensor* and HTML frames (See figure 1).

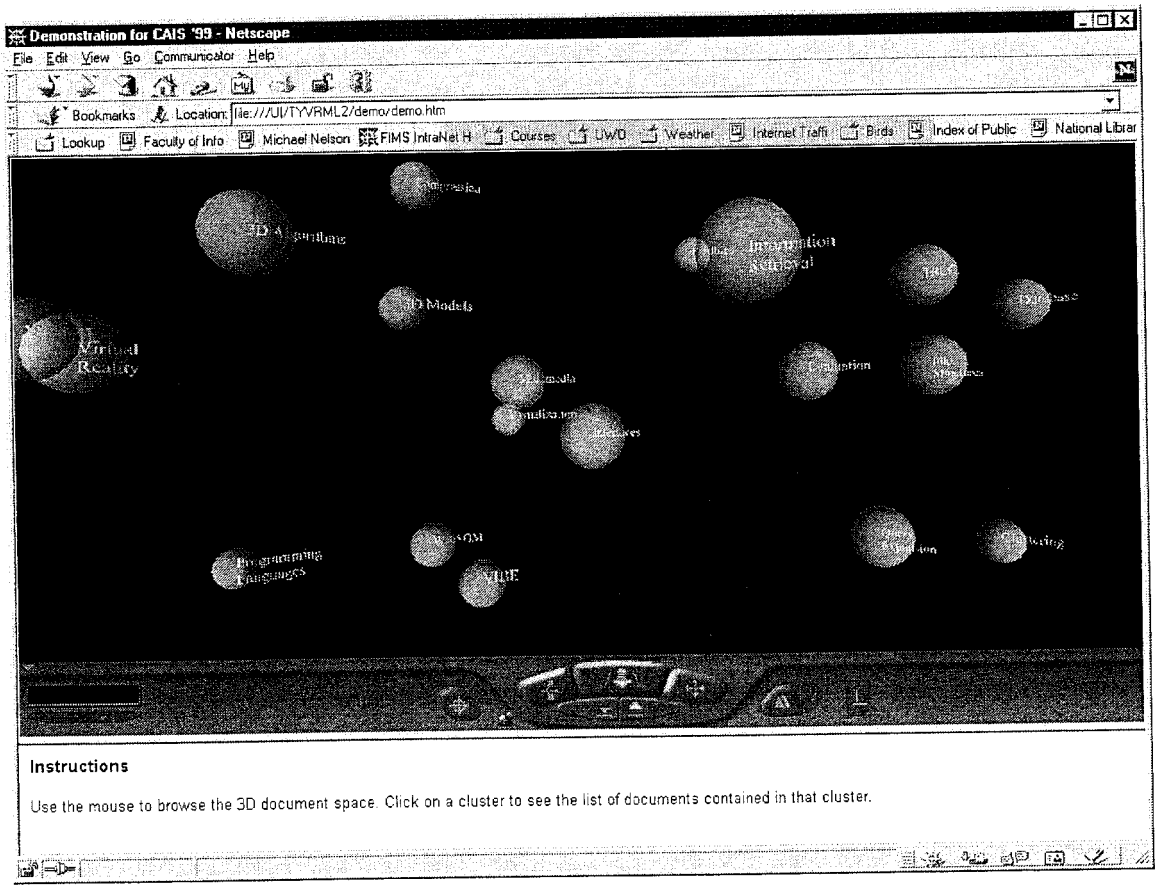
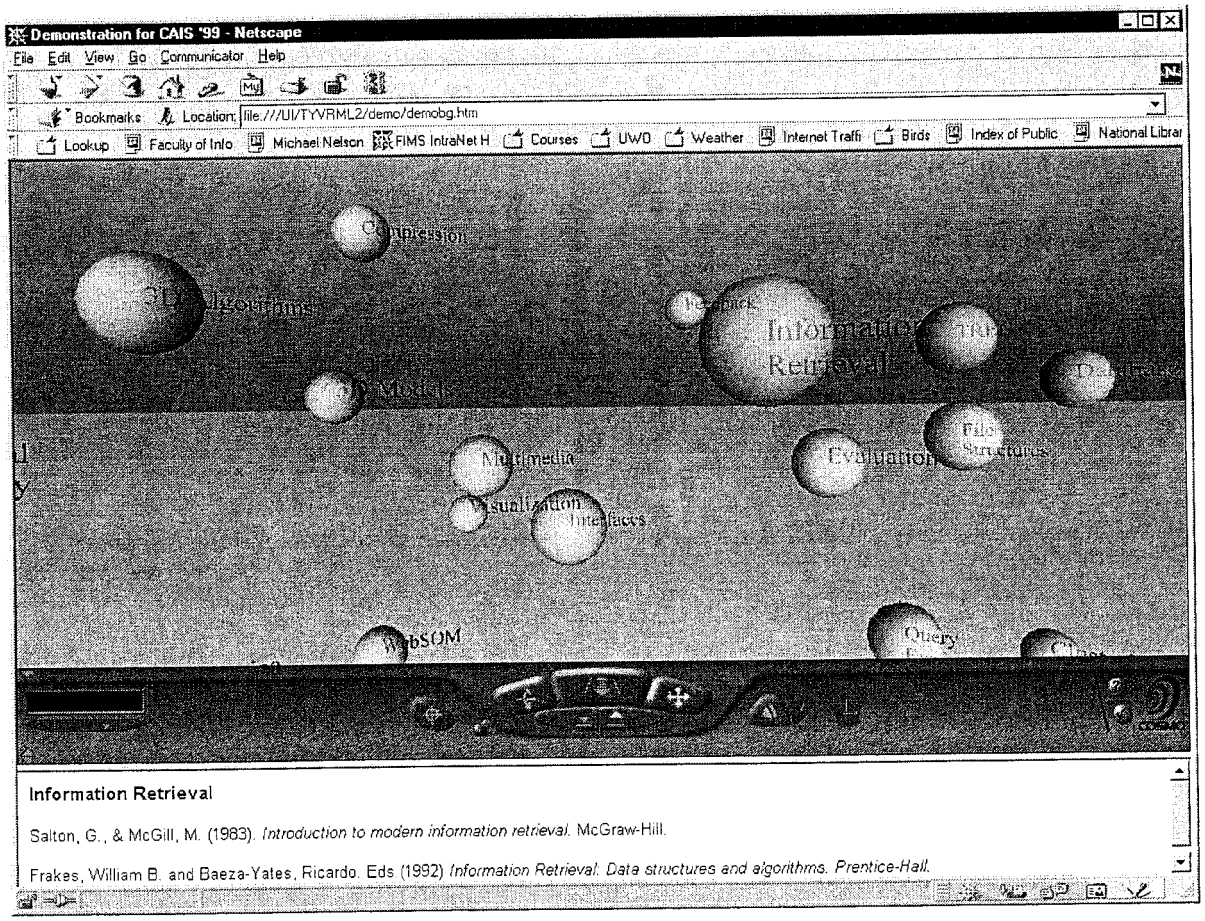


Figure 1

One of the more difficult things to accomplish was to put text labels on each cluster. First, in a 3D world only the closer clusters are readable. VRML only has 2D flat planes of text. The solution was to make each sphere partially transparent and to put the text inside the sphere where possible. To make the text visible from any angle a special feature of VRML called a *billboard* which turns to face the user as they move through the 3D world. The more basic problem is exactly what label to place on the cluster as this is the main clue the user has to decide on which cluster to investigate further. This is an automatic indexing problem which needs much more research.

One of the other calculations which must be done by the retrieval engine is to position each cluster or document in space. Previous solutions have been briefly discussed above. For the current project work is being done on taking a set of inter-cluster similarities and calculating 3D coordinates for each cluster.

In a starfield, navigation is a problem because there are no signposts for users to tell which way is up, down or how far one has travelled. One partial solution is to use a VRML *background* which gives a sky, a sun, and a ground which is always far away from any objects in the space. This is halfway towards a terrain or map, but gives the user a better orientation within space (See figure 2) at the expense of more processing for the viewer.



Evaluation

No evaluation of this system has been attempted, but others have started to develop methods for visualization systems. For example, Morse et.al. (1998) compared several different interfaces for information retrieval. They found that about 60% of the users (students) preferred the visual displays but that the text interface performed better for the retrieval questions used. This suggests that much more work needs to be done to carefully design the interface to see if the performance can be improved. Preference and performance testing should be done throughout and this is one of the upcoming tasks for this project.

Summary

Currently the VRML generator is separate from the retrieval engine but work is progressing to integrate all the components into a system which can start with an initial text query, then display the retrieved documents in a 3D space for browsing. From the existing implementations, it seems that VRML is very easy to use and can represent any of the current ideas for representing document collections in 3D. It has the advantage of being a standard which can be used in any browser with a VRML addin. The performance is quite adequate for smaller spaces (twenty to thirty clusters) and

collections in 3D. It has the advantage of being a standard which can be used in any browser with a VRML addin. The performance is quite adequate for smaller spaces (twenty to thirty clusters) and more investigation needs to be done into larger spaces. There are several techniques that can be used to make the VRML files smaller. One used in the examples is to use something called a prototype for the spheres, then each sphere is built with one function call to the prototype. In addition one could use very short variable names and do without any fancy formatting to save space. The files can also be sent in gzip compressed format if they are being sent over the internet. Another factor which has not been mentioned yet is the user interface in the browser which is currently not very intuitive. In the figures the viewer from CosmosWorld is depicted but the system might benefit from a custom viewer with an interface built expressly for this task. One can also imagine a full virtual reality interface with goggles and gloves, which is entirely possible using VRML.

Bibliography

Agosti, M., Colotti, R., & Gradenigo, G. (1991). A two level hypertext retrieval model for legal data. 14th Annual Conference on Research and Devlopement in Information Retrieval, Chicago 1991,. Washington, D.C.: ACM.

Card, S.K., Mackinlay, J. D. & Shneiderman, B. Eds. (1999). *Readings in information visualization: Using vision to think.* San Francisco, CA: Morgan-Kaufmann.

Chalmers, M. (1994). Ongoing work on Bead, an information visualiser. In L. Ruben (Ed.), *Information retrieval: New systems and current research.: 15th Research Colloquium of the British Computer Society Information Retrieval Specialist Group 1993,*. London: Taylor Graham.

Deerswater, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391-407.

Hemmje, M., Kunkel, C., & Willett, A. (n.d.). LyberWorld - A visualization user interface supporting fulltext retrieval. In W. B. Croft, & C. J. van Rijsbergen, (Eds). In *Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval 1994*, (pp. 249-259). London: Springer-Verlag.

Honkela, T., Kaski, S., Lagus, K., & Kohonen, T. (1996). Exploration of full- text databases with self-organizing maps. *International Conference on Neural Networks, Washington 1996, June 2-6,*.

Kinnucan, M. T. (1992). Fisheye views as an aid to subject acces in online catalogues. *Canadian Journal of Information Science, 17*(2), 25-40.

Kohonen, T. (1998). Self-organization of very large document collections: State of the art. In L. Niklasson, M. Boden, & T. Ziemke, (Eds). In *ICANN98, the 8th International Conference on Artificial Neural Networks 1998, Vol. 1*, (pp. 65-74).

Korfhage, R. R. (1991). To see or not to see? - is that the query? *Fourteenth Annual Conference on Research and Development in Information Retrieval, Chicago 1991*, (p. 134). Washington, D.C.: ACM.

Lin, X. (1992). Visualization for the document space. *IEEE Visualization '92 1992*, (pp. 274-281).

Marrin, C., & Campbell, B. (1997). *Teach yourself VRML 2 in 21 days*. Indianapolis: Sams.

Morse, E., Lewis, M., Korfhage, R. R., & Olsen, K. (1998). Evaluation of Text, Numeric and Graphical Presentations for Information Retrieval Interfaces: User Preference and Task Performance Measures. *IEEE International Conference on Systems, Man, and Cybernetics 1998*, (pp. 1026-1031).

Morse, E. L., & Lewis, M. (1997). Why information retrieval systems sometimes fail. *IEEE International Conference of Systems, Man, and Cybernetics, Orlando, Fl. 1997, October 12-15*, (pp. 1680-1685).

Mukherjea, S., Hirata, K., & Hara, Y. (n.d.). Towards a multimedia World Wide Web information retrieval engine. *Sixth International WWW Conference, Santa Clara, CA 1997*,. http://www.scope.gmd.de/info/www6/technical/ paper003/paper3.html: .

Nowell, L. T., France, R. K., Hix, D., Heath, L. S., & Fox, E. A. (1996). Visualizing search results: Some alternatives to query-document similarity. *19th Annual International Conference on Research and Development in Information Retrieval, Zurich 1996*, (pp. 67-75). Washington, D.C.: ACM.

Olsen, K. A., Korfhage, R., Sochats, K., Spring, M. B., & Williams, J. (1993). Visualization of a Document Collection: The VIBE System. *Information Processing and Management, 29*(1), 69-81.

Robertson, G. G., Card, S. K., & Mackinlay, J. D. (1991). Information visualization using 3D interactive animation. *Communications of the ACM, 36*(4), 57-71.

Rohrer, R. M., & Swing, E. (1997). Web-Based information visualization. *IEEE Computer Graphics and Applications, 17*, 52-59.

Salton, G., & McGill, M. (1983). *Introduction to modern information retrieval.* McGraw-Hill.

Spoerri, A. (1993). InfoCrystal: A visual tool for information retrieval and management. In B. K. Bhargava, T. W. Finin, & Y. Yesha, (Eds). In *Conference on Information and Knowledge Management, Washington, D.C. 1993*, (pp. 11-20).

Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., & Crow, V. (1995). Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *IEEE Information Visualization '95 1995*, (pp. 51-58).