

# Application Study of Information Visualization in Digital Library

Zhijiang Liao  
Library  
Hebei United University  
Tangshan, China

Min Gao  
Library  
Hebei United University  
Tangshan, China

Feng Yan  
Department of Scientific Research  
Management  
XingTai University  
Xing Tai, China,

**Abstract**—Information visualization in digital library can solve some problems faced during its development. In formation visualization consists of four parts: determine the object, visualization mapping, space structure and graphics rendering. This article introduces application of information visualization in information resources, information retrieval process, information retrieval results and knowledge domain in digital library.

**Keywords**- Information visualization; Information retrieval; Digital library; knowledge domain

## I. INTRODUCTION

The advent of the information age, the digital information showing a trend of exponential growth, excessive information has become barriers to access to effective information. Easy to find the information from these massive amounts of digital information resources, and discover inside information, the relationship between the information and the development trend of information, in order to make better use of these data, information visualization is applied in the digital Library. The information visualization is a process of transform data information and knowledge into a visual form to make full advantage of people's natural capacity for Visual mode<sup>[1]</sup>. Information visualization as an service mode of information effective organization, analysis, reveals, can fundamentally change information acquisition, information organization, information storage, information present and information access methods, information image visualizations, scientifically reveals poi between the direct and indirect relations, the user can give full play to their creativity and imagination of the flexibility to search for information, explore the great value of the information resources.

Digital Library is the use of advanced information technology digitized literature resources, the orderly organization of massive information intact, and with the knowledge of the efficient use of network technology, information center. With digital information resources, information access network, the proliferation of information services and other features. Information visualization applications in the digital library of abstract information and its intrinsic link can be presented to the user graphical information into meaningful and understandable visual representations, thereby promoting user information Search and cognitive processes, the vast amounts of data to be used effectively in the digital library.

## II. INFORMATION VISUALIZATION BASIC PROCESS

Information visualization can be thought of as information from data to visual form to people's perception system of adjustable mapping, Card<sup>[2]</sup> proposed simple reference model as shown in figure 1. The key problem to be solved of Information visualization is the mapping, transformation and interactive control of reference model.

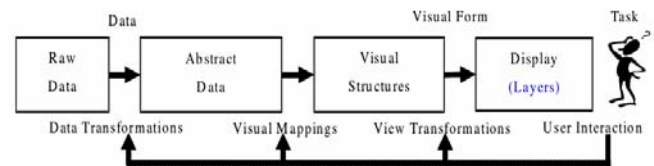


Figure 1 information visualization reference model

Information visualization model has a lot of kinds, The main have the Multiple Reference Point Based Models(MRPBM), the Euclidean Spatial Characteristics Based Models(ESCBM) and the Self-organizing Maps Models(SOM), the Path-finder Associative Network Models(PFNET), the Multi-dimensional Scaling Models(MDS).

In this model, information is through a series of data transformation from the original data to a visual display. In the diagram from left to right, each arrow represents is likely to be a series of transformations. From the user to each of the transform (from right to left) arrow indicates that the user operation control adjustments to these transformations. Data transform the original data mapping for structured data; visual mapping converts data table into visual structures; view transformations by defining the position, scale, proportion, reduction of graphics parameters such as create a Visual structure view. The user's interaction is used to control the parameters of these transforms. Visualization and their interactive control service to a work task at last.

### A. Determine the Visual Object

The structure of visual object extraction and virtual structure is the foundation of information visualization, to realize the information must demonstrate the effective extraction right visual object. Determine the visual object is refers to the analysis data set of internal characteristics (semantic structure, link relations, the reference relationship,

etc.), extract structured information (such as link, citing relationship), context information, metadata, use information (such as browsing sequence, operation) and semantic information (field, keywords, abstract, title, etc.) and so on, provide the basis for the selection of suitable visual form.

After determine the Visual object, can build a virtual structure based on the visual objects and relationship between visual objects. So-called virtual structure refers itself does not exist in the original data, require a certain amount of extract, analytical methods are abstract. Virtual structure often can help users to discover hidden patterns in the data and associated information<sup>[3]</sup>. To in II dimension space in the displayed out document between of similar degrees, needs on high dimension document space drop dimension, currently by used of drop dimension method more is reference has science Visual of in the of mathematics method, as used Factor Analysis method analysis different author of terms habits<sup>[4]</sup>, and uses Multi-dimension scaling constructed concept map<sup>[5]</sup>, and uses Latent Semantic Analysis to document visualization dimensionality<sup>[6]</sup>.

### B. Information Visualization Mapping

In information visualization, dealing with abstract information, so first of all visual objects and virtual structures need to be converted to visual attributes, such a transformation process is called information visualization maps. Due to the scientific visualization of visual objects themselves have some space or physical attributes, and information visualization to achieve in the visual object itself does not have an abstract information in space properties, therefore, need to convert the abstract information properties into visual properties. Such a process is the difficult point of the abstract information visualization compared with scientific visualization.

Visual structure has three basic parts: space base, mark, mark graphic attributes. In the visualization process, data structure is mapping for visual structure, visual structure in a space base with mark and graphic attributes of information coding. In order to get a good visual structure, the mapping must be able to maintain the original data information, and only data the data in the table can be expressed in visual structure. To get a good mapping and is not an easy thing, because in the visual structure is easy to appear the unnecessary data. Therefore, how to seek a good visual structure, information visualization is one of the key problems.

### C. Visualization Space Structure

After the visual object mapping into visualization of visual properties, need to form a visualization space according to the virtual structure and their discrete visual property organization together at first step, the key is to determine an effective visual metaphor form, this metaphor forms should be able to real effectively contain and express the information and the relationship between information, which allows users to use its inherent spatial relationships to understand the ability to understand abstract information.

The structural design of the visualization space means to determine the number of dimensions of a visual space, and define the axis of the coordinate system. The number of the dimension of the visual space can be one-dimensional, two-

dimensional or three-dimensional. In order to utilize the advantages of the spatial structure, most of the information retrieval visualization models are based on two-dimensional or three-dimensional. The coordinate system can be rectangular coordinates, polar coordinates or parallel coordinates. Among them, the use of the Cartesian coordinate system is most widely used. The parallel coordinate multiple relationships can be converted into a model. Based on the attribute of the extracted data, the coordinate axis of the coordinate system can be qualitative, sequential or quantitative. It must be noted that the axes form, in some cases are not necessarily identical.

### D. Visual Graphics Rendering

Drawn visual graphics is the last step, the final completion of the abstract information presented to the user with intuitive graphical way after graphing information visualization. Visual rendering how to determine how much the distance between the position of the node in the display screen, the node with the node, using the side connected between which nodes, etc. need to be determined in advance before the draw visual graphics.

The difference between the visualization and the traditional graphical display is the visualization need to visualization of large amounts of information. so the traditional drawing method can't be revealed complex information space. How to ordinary computing power of computer realize graphics rendering and realize real-time interactive graphics is the key, in the graphic drawing can use existing graphics software generate graphics, also can use certain graphics layout display algorithm is proposed to realize. The current graphics software mainly has AT&T Graphviz, 3D graphics software OpenGL, etc. And at present relatively successful graphic display algorithm have Spring Embedder to realize the visualization of information space display<sup>[7]</sup>, based on sample structure analysis database visualization algorithm, based on hierarchical structure visualization of Mesh algorithm and its optimization algorithm AMR (Adaptive Mesh Refinement).

## III. INFORMATION VISUALIZATION APPLICATIONS IN THE DIGITAL LIBRARY

### A. Information Resources Visualization

Digital Library is a digital information resource center, the digital library's information resources are mainly three categories of text information, audio and video information

1) *Text information visualization*: To the contents of the text information in graphical form to show icons mapped out the most simple method can be used. Depending on the text according to their classification criteria are assigned with a different icon, Users simply browse icon to generally understand the content of the text. This method is very simple to operate, only need to establish a icon sets, and the fixed mapping function can be established between the relevant standards and icon sets. But the icon can reflect the meaning is limited and rough. In addition, you can also use keywords sorted word frequency tree method will retrieve the resulting from a large number of users at a time out of all the keywords in the literature collected for statistical analysis, and these documents can be classified according to the key words.

According to the number of occurrences of the search keywords in the text for each page, using different colors to indicate the number of occurrences of the word in the page, so that the user can see a glimpse of keywords in this paper the distribution, and then directly select a page to read and check.

2 ) *Non-text information visualization*: For non-text information visualization, the main purpose is to provide users with multiple viewing angle to prompt potential is difficult to convey the meaning of the document information surfaces better, some music files, its main theme is the soul of can be in the form of the picture presented to the user, the user will be able to listen to the song before it there is a general awareness.

Pictures itself has to people intuitive table meaning of role, and uses Visual of technology will its placed a virtual space, by user to select observation of distance and angle can deepened user of impression and understands as LVis used 3D technology stereo to show various painting works, let user has immersive of feeling<sup>[8]</sup>. Video information of Visual of is can through Browse key frame of method to achieved.

### B. *Information Retrieval Process Visualization*

Information retrieval visualization refers to the information resources, the user questions, information retrieval model, the retrieval process and the retrieval results of invisible interior semantic relationship convert into graphics or images, and displayed in a two-dimensional, three-dimensional or multidimensional visual space, these intuitive image of spatial information to help users understand the search results, grasp the retrieval direction to guide the retrieval process, in order to improve the efficiency and performance of information retrieval.

Retrieval process visualization can provide information feedback, allow retrieval process reversible, support retrieval strategy control, and user's interaction etc. Information retrieval process visualization mainly adopts the Shneiderman information visualization classification method. By increasing the search path to information space, and other user retrieval other path to provide valuable information. In the process of information retrieval visualization in the visualization technology are mainly Pathfinder Network Scaling and Multidimensional Scaling technology<sup>[9]</sup>. Retrieval process visualization of the most representative of research results is the development of Advanced Query User Interface Architecture (AQUA) system supported European Community. AQUA is based on achieving retrieved chain mode, mainly used in digital library Dienst for literature inquires to provide a powerful, easy-to-use visual interface. The system of visualization window from left to right display several panel, the left panel is used to input term and the corresponding retrieval results, on the right panel used to further limited search range. All panel will form the whole retrieval process retrieval chain, and this chain is a continuous process, in this process the user can need according to oneself arbitrarily amend, add, delete retrieval chain of face plate, in order to optimize retrieval strategy

### C. *Information Retrieval Results Visualization*

The traditional information retrieval results are linear array, each time to give the user the information provided is very limited, and the retrieval results cannot whole browsing. In addition, keywords and the relevance of the document, the document retrieval the relationship between the can't clear expression comes out. The retrieval results of visualization is mainly devoted to design an interface show all the retrieval results, will the retrieval results with graphics, images, animation and so on visual way display, expressing the relationship between retrieval content, retrieval content reveals the profound meaning, so that users to search results to browse and operation. Use of statistical, clustering, correlation analysis, and other means to hit the collection of information analysis processing, reveals the retrieval results rules hidden in and contact, and prescribed view forms to show users.

For information retrieval results visualization research has made many achievements, especially in the digital library research Xerox PARC in a typical research achievements, such as Tilebars and Scatter/Gather system<sup>[10]</sup>. TileBars system is developed by the Berkoley university digital library researchers. Tilebars system allows users to use complete information, based on the document retrieval word distributed behavior, decide which documents and document what part for browsing, fast and simple to display: the document correlation length; In the document retrieval words the frequency of the occurrence, search terms in the distribution of documents and the relative distribution between the search terms and search term.

### D. *Domain of knowledge visualization*

Along with the digital library theory and practice continues to mature, to obtain knowledge is in constant into geometric exponential growth, the study knowledge and scientific structure, scientific development power (such as development speed, development mode, etc.), looking for the forefront of science, more and more get people's attention. Information visualization technology for the development of research in the field of knowledge related problem to provide the new train of thought, especially in recent years, the information visualization technology and scientific metrology method, combining generation has all kinds of attributes of scientific map, expression disciplines and fields, professional, literature, author, the relationship between the interpretation in the field of knowledge structure, mapping the development trend of the field of knowledge, and to promote the access to information, so that the knowledge structure more apparent, helping the user to achieve their purpose.

Knowledge Visualization includes intelligent user interfaces, such as semantic knowledge of browser or browser, it allows people to share knowledge through Visual. Semantic link network and cognitive maps can reduce the divide between knowledge representation and knowledge visualization. Digital library information visualization system enables people to browse a lot of complex data, quickly find the information you want, easier to navigate and interact with the data, identifying patterns and trends, as well as a better understanding of information. The contents and forms of knowledge visualization and Visual descriptions are different. On the content, it gets more than just facts and figures, there are ideas,

principles and relationships. Formally speaking, knowledge visualization rely on direct communication, not only raise active Visual sense and motivates people to complete your own picture.

#### E. *The man-machine interface visualization*

Interactive interface (user interface) is an important part of the digital library. It is communication window between user with system when users find, search and retrieving the information in digital library. The visualization of the user interface helps users to enhance the affinity of the retrieval system. Retrieval process visualization helps users to further understand, recognize and correct their own information needs, express their own needs semantic layer close to the actual needs of intent and context level. Interface windows design need to pay attention to the following aspects: a friendly, intuitive, convenient, and has the effect of humane and intelligent use of graphics and voice, so that the user handy; able to attract the attention of the user, to provide users with the most convenient mode of operation.

#### IV. CONCLUSION

Information visualization technology in many fields have begun to play an important role, how to play its role in the digital library, will existing huge amounts of information maximum limit development and utilization, solve the information demand and the service of individuation, information to provide personalized, it is still a worthy of exploration and research. This needs to strengthen the construction of information visualization and information theory, method, technology and means of research, at the same time to research how to effectively use the visualization

technology achievements to realize the information architecture theory, thus the maximum degree realizes digital library the value of information visualization.

#### REFERENCES

- [1] N. D. Gershon, S.G. Eick. "Information Visualization". IEEE Computer Graphics and Applications. Los Alamitos, Vol.17,pp.29-31, July 1997.
- [2] S. K. Card, J.D. Mackinlay, and B. Shneiderman, Readings in Information Visualization: Using Vision to Think. San Francisco: Morgan Kaufmann, 1999.
- [3] M. A. Hearst, "TileBars: visualization of term distribution information in full text information access," In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing System, Denver, pp. 59-66, May 1995.
- [4] T. Snelgrove, "A Method for the Analysis of the Structure of narrative Texts," Literary and Linguistic Computing. Oxford , Vol.5,pp. 221-225,May 1990.
- [5] E. Brunet, What Do Statistics Tell Us ? Research in Humanities Computing, Oxford: Clarendon Press, 1991, pp. 70-92.
- [6] P. Mutton, P. Rodgers, Spring embedder preprocessing for WWW visualization. Sixth International Conference on Information Visualization, IEEE, pp.744-749, July 2002.
- [7] T. K. Landauer et al., "From paragraph to graph: Latent semantic analysis for information visualization," Proceedings of the National Academy of Sciences of the United States of America, Vol.101(Suppl), pp. 5214-5219, April 2004.
- [8] E.J. Isaacson, "Content visualization in a digital music library," Third International Workshop on Information Visualization Interfaces for Retrieval and Analysis (IVIRA), Houston, May 2003.
- [9] J.Y. zhou, T. Sun, "Application of Information Visualization in Digital Library," New Technology of Library and Information Service. Beijing, pp.5-8, January 2005.
- [10] M. Hearst, G. Kopec and D. Brotsky, "Research in Support of Digital Libraries at Xerox PARC," D-Lib Magazine, June 1996.