Research on Library Information Visualization Retrieval Technology Based on Readers' Needs

Jianfei Hu¹, Di Wang¹, Bin Zhang¹

¹Library, University of Jinan, Jinan, Shandong 250022, China

Abstract: With the advancement of social informatization and the widespread application of network, the resources accumulated by libraries are becoming larger and larger. In addition to the need to store, transmit, retrieve and classify massive data, there is an urgent need to understand the relationship between data and its development trend. With the continuous development of modern information technology, digital libraries have also made great progress. And to introduce visual information retrieval technology into digital libraries is an inevitable trend in the future development. This paper first analyses the characteristics of readers' information retrieval needs in the new period, and then introduce library information visual presentation technologies to design the library information resource visualization retrieval system. The system can visually display the collection structure, popular books, readers' interests and so on, to truly reflect the structure, coverage, utilization and other indicators of books resource, and provide convenience for the visual sharing of library information resources.

1. Introduction

Information visualization technology is to display, interpret or transmit the original abstract data information in a more vivid way, such as graphics or images, while summarizing the existing laws of data. To display the structured and unstructured environment more intuitively with the help of computer technology, multimedia technology and digital technology is the focus of information visualization.

Digital libraries are developing with the progress of modern information technology. It can not only save massive multimedia information in electronic format, but also provide corresponding retrieval, addition, modification, deletion and preservation services. Introducing visualization technology into library information retrieval will inevitably arouse users' interest, simplify retrieval operation and improve retrieval efficiency. On the one hand, visual retrieval technology can display library information resources more intuitively and vividly in graphical way, which is convenient for users to analyze and understand book information; on the other hand, it can also provide more intuitive guidance for readers in information retrieval, to simplify retrieval operation steps and improve retrieval efficiency.

This paper first analyses the characteristics of readers’ information retrieval needs in the new period, and then introduce some library information visual presentation technologies to design the library information resource visualization retrieval system, and interpret the dimensional navigation and visualization design of each module, including library information resource structure visualization module, coverage visualization module, reader knowledge discovery module and literature knowledge discovery module.
2. Reader information retrieval requirements in the new period

In the past, traditional libraries have been limited in their service ability and means because of the imperfect infrastructure and technical level. At present, due to the emergence of new technologies, many libraries have gradually adopted information technology and network technology to improve their services and facilities, so that readers' information needs are no longer limited by single source, limited content and single access, but turn to the diversification of information, the enrichment of information content, and the acquisition of information is not limited by space and time. With the continuous popularization and application of new technologies, the carriers of knowledge and information units are gradually diversified, and the corresponding readers' information needs are gradually changing. Their demands put forward higher requirements for the forms and carriers of knowledge and information. Compared with printed documents, electronic documents contain a large amount of information resources, and readers have more extensive access to information resources and more diverse types of information resources. Therefore, the visualization of library information retrieval is important.

Visual library information retrieval system can greatly broaden the scope of influence of the original library. It can promote the growth of traditional libraries and thoroughly change the service form. For example, the visual library system makes the service of the library have no time and geographical limitations, but the traditional sense of the library can only serve in a fixed location in a fixed time. Compared with traditional libraries, the visual library information retrieval system greatly increases the utilization efficiency of books in libraries, and brings great experience to readers. Visual library information system can save a lot of books because of its infinite and expandable electronic storage space, which greatly increases its collection capacity. Moreover, it can also facilitate the sharing of resources between libraries.

3. Library information visual presentation technology

Combining the advantages and disadvantages of various visualization technologies, choose the most appropriate visualization technology for each library resource model, or use a variety of techniques to achieve the desired results of the visualization of library resources in various theme models, so as to build a visualization service based on library resources.

3.1 Information dimension and attribute presentation technology

Parallel coordinate method can be adopt for information dimension and attribute presentation, which can present multi-dimensional and large amount of book information. Moreover, the visualization structure of the technology can also reflect the relationship among the dimensions and clearly show the trend of data among the dimensions. Information dimension and attribute presentation by parallel coordinate method is shown in Table 1.

| Information dimension | Information attribute                                      | Information attribute                                      |
|-----------------------|-----------------------------------------------------------|-----------------------------------------------------------|
|                       | Category and quantity                                     | Proportion of category and quantity                        |
|                       | Per capita category and quantity                          |                                                           |
| Time                  | Category and quantity vary with time                      | Proportion of category and quantity vary with time         |
|                       | Per capita category and quantity vary with time           |                                                           |
| Press                 | Category and quantity vary with press                     | Proportion of category and quantity vary with press        |
|                       | Per capita category and quantity vary with press          |                                                           |
| Subject               | Category and quantity vary with subject                   | Proportion of category and quantity vary with subject      |
|                       | Per capita category and quantity vary with subject        |                                                           |
| Bookseller            | Category and quantity vary with bookseller                | Proportion of category and quantity vary with bookseller   |
|                       | Per capita category and quantity vary with bookseller     |                                                           |
| Time-press            | Category and quantity vary with time of different press   | Proportion of category and quantity vary with time of      |
|                       | Per capita category and quantity vary with time           |                                                           |

Table 1. Information dimension and attribute presentation by parallel coordinate method
In order to realize the visualization of utilization rate of book resources, dimensions and attributes of statistics need to be clarified, such as time, subject classification, press, bookseller, etc. Similarly, to realize the visualization of book resource coverage, corresponding attribute information should also be included, such as current collections, target collections, etc. Obviously, these data have the characteristics of more dimensions and larger data volume. In addition, based on the analysis of readers' needs, the final visualization must also reflect the trend of the collection structure, coverage and utilization over time, subject classification and other dimensions. Therefore, we propose to utilize parallel coordinate visualization method, taking information dimension as horizontal axis, taking the attribute value that needs to be counted as the vertical axis, so as to clearly show the multi-dimensional and multi-attribute resources of the collection.

3.2 Icon visualization technology

Icon-based visualization technology is also a common visualization technology, which presents the data needed to be visualized on a meaningful icon. As far as possible, icons similar to the subject of the target visual data should be selected. For some data which can be easily deployed on two-dimensional plane, this technology has obvious advantages over other visualization technologies.

In the knowledge discovery model of library collection resources, the visualization of reader's interest obviously has the characteristics of this visualization technology requirement. Reader's interest is directed at book subject words, which are displayed in a better form through visualization. Icon visualization technology could help to show the primary and secondary, heat and time changes of reader's interest. In icon-based visualization technology, choosing an appropriate and meaningful icon is the most important part in the visualization process. E.g. when visualizing the classification of books on a "CSS technology" topic, the visualization effect can be to combine some popular book covers to form a "CSS" shape. There are many successful visualization cases in the field of book information visualization. Tag cloud is probably one of the most frequently used visualization services, as shown in Figure 1.

The visualization method in the form of "tag cloud" can facilitate the visualization of data with a small number of text attributes. And the visualization method can provide a variety of fonts, colours and layout, and generate beautiful visual icons. In the visualization of reader's interest in library collection knowledge discovery, the subject text of reader's interest is excavated. These limited
number of interest text labels are the basis of icons. The model of knowledge discovery in this paper uses these interest tags to control the position, color and size of the tags according to the reader’s interest, and dynamically visualize the form of reader’s interest as “interest cloud”.

3.3 Clustering visualization technology
For data sets with hierarchical structure, their specific information can be presented hierarchically and classified. Through hierarchical presentation, data sets with hierarchical relationships can reveal the overall layout, the potential association and interaction between data and data.

The relationship of library information resources obviously has the above characteristics. We define the association rules of books and documents as follows: if some books have the same subject terms, then they have an association; if Book A is associated with Book B, Book B is associated with Book C, then Book A is associated with Book C. Based on the above facts, if a book has more books associated with it, it will be more conspicuous when he is presented. When using the hierarchy-based visualization technology to express the book association relationship, firstly cluster analysis of the book literature, select some subject books that can reflect the characteristics of the library, and then expand the associated books of the selected books, thereby finally, a mesh relationship diagram is presented. This kind of book relevance degree visual analysis method not only shows the related structure of a certain kind of subject literature, but also can find some interdisciplinary subjects to help readers expand their professional knowledge. The effect of by clustering visualization technology is shown in Figure 2.

![Figure 2. Clustering visualization technology visualization effect](image)

4. Design of library information resource visualization retrieval module

4.1 Overall system framework
Based on the needs of readers, this paper designs a library information visualization retrieval service system, which includes a front-end browser section for displaying various effects of visualization. The overall framework of the system is shown in Figure 3.

![Figure 3. Overall framework of library information visualization retrieval service system](image)
According to the above framework, the specific module design of the system includes browser page visualization various presentation forms design, data model and data management design and data communication interface module design.

4.2 Library information resource structure visualization module
We utilize parallel coordinate method to realize library information resource structure visualization, in the form of report. According to the requirements of the parallel coordinate method, multiple dimensions correspond to multiple coordinate axes, in the design of this module, a coordinate axis is used, and the corresponding dimension information is displayed in a page in the form of browser navigation. The navigation bar of the dimension in the module and the visual effect corresponding to each dimension are designed as Table 2.

Table 2. Design of library information resource structure visualization module

| Information retrieval dimension | Information visualization design | Data analysis design |
|---------------------------------|---------------------------------|---------------------|
| Time                            | Time as abscissa, legend attribute value as ordinate; histogram and corresponding scale graph | Analyze the visualization result data and give the overall trend of the collection resources based on the time dimension |
| Subject classification          | Subject classification as abscissa, legend attribute value as ordinate; histogram and corresponding scale graph | Analysis the book discipline classification characteristics and trends |
| Press                           | Press information as abscissa, legend attribute value as ordinate; histogram and corresponding scale graph | Analysis of the characteristics of books belonging to press |
| Bookseller                      | Bookseller information as abscissa, legend attribute value as ordinate; histogram and corresponding scale graph | Analysis of the characteristics of books belonging to booksellers |
| Time-subject                    | Time as abscissa, legend attribute value as ordinate, subject classification pie chart option; display the book resources of the time and subject | Analyze the trends of books in each subject over time |
| Time-press                      | Time as abscissa, legend attribute value as ordinate, press pie chart option; display the book resources of the time and press | Analyze the trends of books in each press over time |
| Time-bookseller                 | Time as abscissa, legend attribute value as ordinate, bookseller pie chart option; display the book resources of the time and bookseller | Analyze the trends of books in each bookseller over time |

Legend information of visual report includes species number, number of books, proportion of species number, proportion of books number, proportion of per capita species number and proportion of books number per capita. The column chart shows the number of species and volumes, and the curve chart shows the proportion information.

4.3 Library information resource coverage visualization module
Visualization of library coverage can be in the form of dynamic reports, which mainly visualizes the proportion of book resources to the coverage of similar institutions and publishers. The dimensional navigation and visualization design is shown in Table 3.
Table 3. Design of library information resource coverage visualization module

| Information retrieval dimension | Information visualization design | Data analysis design |
|---------------------------------|---------------------------------|---------------------|
| Coverage of similar institutions |                                 |                     |
| Bibliographic coverage          | Specified subject as abscissa,  | Analyze bibliographic |
|                                 | attribute value of the          | coverage of specified |
|                                 | corresponding bibliography     | subjects in similar   |
|                                 | as ordinate                    | institutions         |
| Author coverage                 | Specified subject as abscissa,  | Analyze bibliographic |
|                                 | attribute value of the          | coverage of specified |
|                                 | corresponding author as         | author in similar     |
|                                 | ordinate                       | institutions         |
| Coverage of publisher           |                                 |                     |
| Bibliographic coverage          | Specified subject as abscissa,  | Analyze bibliographic |
|                                 | attribute value of the          | coverage of specified |
|                                 | corresponding bibliography     | subjects in the       |
|                                 | as ordinate                    | publisher            |
| Author coverage                 | Specified subject as abscissa,  | Analyze bibliographic |
|                                 | attribute value of the          | coverage of specified |
|                                 | corresponding author as         | author in the         |
|                                 | ordinate                       | publisher            |

Legend information includes the number of bibliographies in the library, the number of bibliographies in similar institutions, and the total number of cross-bibliographies between our library and similar institutions, which are displayed by histogram

4.4 Reader knowledge discovery module
Visualization of library reader knowledge discovery can be done in the form of report forms and visual programming. The dimensional navigation and visualization design is shown in Table 4.

Table 4. Design of reader knowledge discovery visualization module

| Information retrieval dimension | Information visualization design | Data analysis design |
|---------------------------------|---------------------------------|---------------------|
| Cluster analysis of readers     | Clustering display of pie chart with reader's active degree; click on each cluster block to display the specific readers of the cluster | Analyze the reader's activity |
| Analysis of reader interest     | Based on the reader's interest tag, show interest "cloud tags"; recommend books of interest to specific readers | Analyze the reader's interest characteristics |

4.5 Literature knowledge discovery module
Visualization of literature knowledge discovery can be done in the form of report forms and visual programming. The dimensional navigation and visualization design is shown in Table 5.

Table 5. Design of literature knowledge discovery visualization module

| Information retrieval dimension | Information visualization design | Data analysis design |
|---------------------------------|---------------------------------|---------------------|
| Literature clustering           | Analysis of the popularity of the literature through the borrowed information, display, clustering display in the form of pie graph; click on each cluster block to display the specific literature of the cluster | Analyze the popularity of literature |
| Literature association analysis | Based on the relevance of literature, displayed by reticular graphs | Analyze the relevance and intensity of the literature |

5. Conclusion
In this paper, we utilize the visual presentation technologies, including information dimension and attribute presentation technology, icon presentation technology, and clustering presentation technology, to construct the library information resource visualization retrieval system. Through the dimensional...
analysis and cluster statistics of book information, the system can visually display the collection structure, popular books, readers' interests and so on, to truly reflect the structure, coverage, utilization and other indicators of books resource, and provide convenience for the visual sharing of library information resources.

References

[1] Feng S. Application of Information Visualization Technology in Public Library Retrieval System[J]. Electronic Technology, 2012.
[2] Li Q, Zhang, Huang. Application of Information Visualization Retrieval in Digital Library[J]. New Technology of Library & Information Service, 2012.
[3] Koshman S. Visualization-based information retrieval on the Web[J]. Library & Information Science Research, 2007, 28(2):192-207.
[4] Zhou S, Qiang B, Gao J, et al. A Method of Information Retrieval Results Visualization Based on Social Network Analysis[J]. New Technology of Library & Information Service, 2013.
[5] Brusilovsky P, Ahn J W, Rasmussen E. Teaching Information Retrieval With Web-based Interactive Visualization[J]. Journal of Education for Library & Information Science, 2010, 51(3):187-200.
[6] Hoeber O, Yang X D. Supporting Web Search with Visualization[M]// Web-based Support Systems. 2010.
[7] Sukhadan K, Vijayarajan V, Krishnamoorthi A, et al. Image based book cover recognition and retrieval[J]. IOP Conference Series: Materials Science and Engineering, 2017, 263:042088.
[8] Benkoussas C, Bellot P. Information Retrieval and Graph Analysis Approaches for Book Recommendation[J]. The Scientific World Journal,2015,(2015-9-30), 2015, 2015(4):1-8.
[9] Julien C A, Guastavino C, Bouthillier F. Capitalizing on Information Organization and Information Visualization for a New-Generation Catalogue[J]. Library Trends, 2012, 61(1):149-161.
[10] Jia Z. A Textbook Embedded with Knowledge of Scholarly Publishing: Review on the Book Literature Information Retrieval & Paper Writing (5th edition)[J]. Library Journal, 2016, 35(10).
[11] Liu L, Zhao Y, Chen J. Book page retrieval method based on convolutional neural network[J]. Journal of Huazhong University of Science & Technology, 2017, 45(11):22-28 and 67.
[12] Pesch O. Library Standards and E-Resource Management: A Survey of Current Initiatives and Standards Efforts[J]. Serials Librarian, 2008, 55(3):481-486.
[13] Novytska T L, Levchenko Y S. Information Resource Management Model Of Scientific Institution Digital Library[J]. Information Technologies & Learning Tools, 2014.
[14] Wilson K. Beyond Library Software: New Tools for Electronic Resources Management[J]. Serials Review, 2011, 37(4):294-304.