情 報 学 基 礎 41－5
（1996. 4. 18）

The Interactive Multi-directional Information Displaying System

Haruo Kimoto

NTT Information and Communication Systems Labs.
309A, 1-2356, Take, Yokosuka-Shi, Kanagawa 238-03, JAPAN
E-mail: kimoto@syrinx.dq.isl.ntt.jp

Abstract

This paper describes a new method of multimedia information retrieval which shows multimedia information on a screen in a multi-directional way, using perspective, during the process of recursive interactive retrieval. This method supports a graphical, sensitive and individually tailored interaction between a user and multimedia data. The original work here in is the support of interaction between a user and multimedia data in a way that shows as many as multimedia data in a multi-directional manner on a screen. This work does not involve analytical techniques, although much current research into multimedia information retrieval focuses on processing and handling multimedia data. The prototype system INMUL(Interactive Multi-directional Information Displaying System) is now being implemented.

インタラクティブ多方向情報表示システム

木本 晴夫

ＮＴＴ情報通信研究所
〒２３８－０３ 横須賀市武１－２３５６
E-mail: kimoto@syrinx.dq.isl.ntt.jp

本稿では、繰り返し的なインタラクティブ検索の過程において、透視的な技法を用いて、画面上に多方向にマルチメディア情報を表示する新しい検索手法について述べる。本手法は、ユーザとマルチメディアデータの間でのグラフィカルで感性的な個人適合型インタラクションを支援するのが目的である。本手法における新提案は、可能な限り多くのマルチメディア情報を画面上に多方向に表示してユーザとマルチメディアデータのインタラクションを支援することである。従来のマルチメディア情報検索の多くの研究がマルチメディアデータそのものの認識処理や操作に焦点をあてているのに相反して、本研究ではそれらの解析的な手法は用いていない。現在はプロトタイプシステムであるＩＮＭＵＬの開発・実験中である。

1. Introduction

Much research and development has been done in the field of Information Retrieval, especially for text document IR. In recent years Natural Language Processing (NLP) techniques have been used for the development of the automatic indexing system, the natural language interface, and the message understanding system, and Neural Network techniques have been used for the document classification. In these systems the kinds of texts used were newspaper articles, patents, academic papers, manuals, books and others. The number of available texts is extremely large.

But these days multimedia data is increasing very rapidly, with the growth of the Internet, which can serve texts and multimedia data. The kinds of multimedia data are still images, videos, music, texts and combinations of these. There is now large need for multimedia Information Retrieval methods, but few methods are in use in the real world, and few techniques have been reported.

The problems of multimedia information retrieval are as follows:

(1)The handling and processing of multimedia data is difficult compared with that of text data.

(2)The amount of data is very large, and so takes a long time to process.

(3)As it is difficult to recognize multimedia data, it is also difficult to process its contents.

Multimedia content processing techniques include image recognition, image retrieval using images, speech recognition, music retrieval by melody and others. Only a few of these techniques are put into use, and most of them are at the research level.

For text retrieval characters, words, and other linguistic units can be used to analyze the content of text. Text analyzing techniques include morphological analysis, syntactic analysis, semantic analysis and discourse analysis. For multimedia retrieval, it is difficult to find any unit which can be used to analyze the content of the data.

Besides the above problems, the problem of handling sensation is important in multimedia information retrieval because multimedia information has sensory information in itself. And any individual, who is an information user, has his own sensations. This is very different from text retrieval. The only technique that deals with this problem is to use sensary words in retrieving. There is little research which deals with sensation in searching multimedia information.

This paper proposes a new method of multimedia information retrieval which shows multimedia information in a multidirectional way on a screen recursively during the process of interactive retrieval, which is the method that supports the interaction between a user and multimedia data in a sensitive and individual way. In this paper, the concept of multimedia information retrieval is the interaction between a user and multimedia data in an individual way.

The originality of this work is the supporting of the interaction between a user and multimedia data in a way which shows as many as multimedia data in a multidirectioal way on a screen. This work does not use analytical techniques, although many current researches of the multimedia information retrieval focuses on the analysis, the processing and the handling of the multimedia data. Current research includes object recognition in an image and melody analysis for retrieving similar melodies.

2.The problems of multimedia information retrieval

There are currently two models of multimedia information retrieval. One model uses word-indexes given to multimedia data. In this paper this is called the index model. The other model matches rough sketches to

multimedia data. In this paper this is called the data matching model.

Neither model is sufficient for multimedia information retrieval. The problems of both models are described below:

(1)Problems of the index model

The index model has the same problems as the model which uses indexes for retrieving documents, that is the traditional information retrieval model. These problems are the relevancy of indexing, the problems of using and maintaining a thesaurus, and other problems.

(2)Problems of the data matching model

(a)It is impossible to make a rough sketch of the warmness of spring.

(b)It is difficult to make a rough sketch of textile designs which have no object, nor shape. For example, if a design is a collection of dots, it is nearly impossible to make a sketch for matching with other images.

(c)As it is very difficult to make a good rough sketch using a mouse device, the system sometimes fails to retrieve the images that a user wants.

The features of database retrieval, information retrieval, multimedia information retrieval and text information retrieval are described below for comparison.

● The features of database retrieval:

(1)A user knows exactly what he wants and can describe it exactly.

(2)A user knows of the existence of the data in the database, for example, boys in a school, or wheels in a database of car parts.

● The features of information retrieval:

(1)A user does not know whether what he wants is in the database or not.

(2)A user can not describe what he wants exactly because he does not know what is in the database. For example, an economist may want to find a newspaper article about the president, but he can not know to exclude articles about the president's dog's illness when he does not know of the existence of such an article.

● The features of text information retrieval:

A user always describes what he wants to find using some linguistic unit, such as characters, words, phrases and so on. So text information retrieval might be refered to as descriptive retrieval, as opposed to non-descriptive retrieval.

● The features of multimedia information retrieval (MMIR):

(1)A user makes a selection that suits him.

(2)Usually a user does not describe what he wants. A user selects what he wants from among the things which he can see. So multimedia information retrieval can be called non-descriptive selective retrieval.

(3)It is useless to describe what he wants to get, because it is not certain that what he wants is in the database. The first thing the user must do is to view the database as effectively as he can.

Besides the aspects described above, several new techniques for MMIR are needed. Again, the fundamental requirement for MMIR is to see all of the data items effectively. For this the concept of browsing is essential. There are several techniques for browsing. They are interactive browsing, walk-through browsing, and directional browsing. For directional browsing, structuring the data is necessary.

For browsing all data, the techniques below are effective:

(1)Showing a perspective-projection image instead of the full image.

(2)Showing the images in a limited time span.

(3)Showing the images in a time flowing stream on a screen.

For effective viewing of all the items in a database, directional browsing is necessary. And showing the perspective-projection image in a limited time with the help of the time-flowing stream technique makes it

more effective, giving users the impression of wandering on the floor of a department store.

3.The design concept of the MMIR system

In this section we will describe the design concept of the multimedia information retrieval prototype system, which meets the needs mentioned in section 2.

A new paradigm of MMIR is to show all items effectively. To show the data in a directional way and to use perspective-projection are effective ways of showing data to users.

The concept of directional search is that a user has some directions in mind during the search process although he usually does not know the end point. So, the directional retrieval function provides users with a choise of directions which are displayed on screen by the system. By choosing the direction recursively, a user can proceed along those directions and will finally be able to find the items desired.

Since the human field of vision is limited, it is necessary to show on the display as much data as possible during the interaction between a human and a computer. Showing the perspective-projected image is effective for this purpose.

The idea of the directional search comes from the following user need: A user often thinks during the search process that there might be something which is near to the result which he has just got, but he does not know how to check for it and wants help from the system. This function meets this need.

This function suits the 3D technique because the 3D technique can show many directions clearly on a screen in a smart way. By using the directional search, the user can satisfy his mind that he should look for data in the neighbourhood of the data he has found, or that there would be something good in the neighbourhood, or that he should check the neighbourhood. Since the user checks by himself with the help of the system, he will be satisfied with the result. He might even be able to check all of the data if he wants. So, the 3D technique can provide a user with a superior browsing function.

Besides its appropriateness for directional search, the 3D technique can utilize the space on a screen, time control for showing the data and handling of the image data as an object on a screen.

From among the many functions that use the 3D technique, the directional search function is reported in detail in this paper. What are the directions in the directional search? The directions can be made in any way. The system can make directions by organizing data of a database in several ways. The manager of a database, for example a librarian, can organize the data in his own way and make corresponding directions. There are many way of making directions, such as classification of data by content, or attributes, by linking of bibliographical references, by chronology, region and so on.

Section 6 of this paper explains one way of making directions using a computer system. Since the directions are made by the system, they can be made and can be changed recursively and automatically through the search process. This automatic making of directions is one of the original techniques of this paper.

In the real world, directions have perspective. In this system, the perspective is made by differentiating the sizes of the boxes in which the data is shown. This means that, in a given direction, the larger box contains the more relevant data, and the smaller box contains the less relevant data. Of course, the boxes are aligned in a size order in a direction. The box size for each document is determined by its relevance point. The rel-

evance point is calculated by the degree of keyword matching with a query.

Following is the list of features included in the prototype system. The name of this system is INMUL (Interactive Multi-directional Information Displaying System).

(1)The ranking function.

(2)The varying size displaying function.

(3)The recursive display function.

(4)The multi-directional displaying function.

(5)The combination of the recursive display function and the multi-directional displaying function.

The author is thinking of including the following features in the next prototype system.

(1)The display-time control function.

(2)Handling of image data as objects in a display space.

(3)Perspective-projection of images.

(4)The combination of the above mentioned functions.

There are many systems which visualize data. Their functions are such as the categorization and the allocation of data on a screen[Lin, 1991].

The aim of making the INMUL prototype system is to evaluate the functions implemented there in with regard to their effectiveness and their ease of use.

4.Related Works

There are several related works. Hemmje displayed documents in a 3D space whose axes were keywords. The scale of these axes could be changed [Hemmje et al., 1994]. Rao showed several displaying methods such as the corn tree and the perspective wall using 3D techniques[Rao et al., 1995]. The hyperbolic tree browser, which is described in his paper, uses perspective to show the relationship between words. Shneiderman made a data visualization system with a dynamic query interface. The data in this system was housing data, such as distance from capital city, price of land and so on. So, this system is in the field of database search, not in the field of Information Retrieval, since the data is already well organized manually [Williamson and Shneiderman, 1992][Plaisant, 1994]. Lin made a semantic map for information retrieval using a neural network's unsupervised learning algorithm [Lin, 1991]. The semantic map, which categorizes and allocates data, is displayed on a 2D space and enables economical representation of the data.

In comparison, the prototype system INMUL is original in its use of a directional search and especially in its use of a multi-directional data displaying function. Also, these functions run recursively during the interaction between the user and the system.

5.The functions of the INMUL system

In this section, the functions of the INMUL prototype system are explained.

(1)The ranking function.

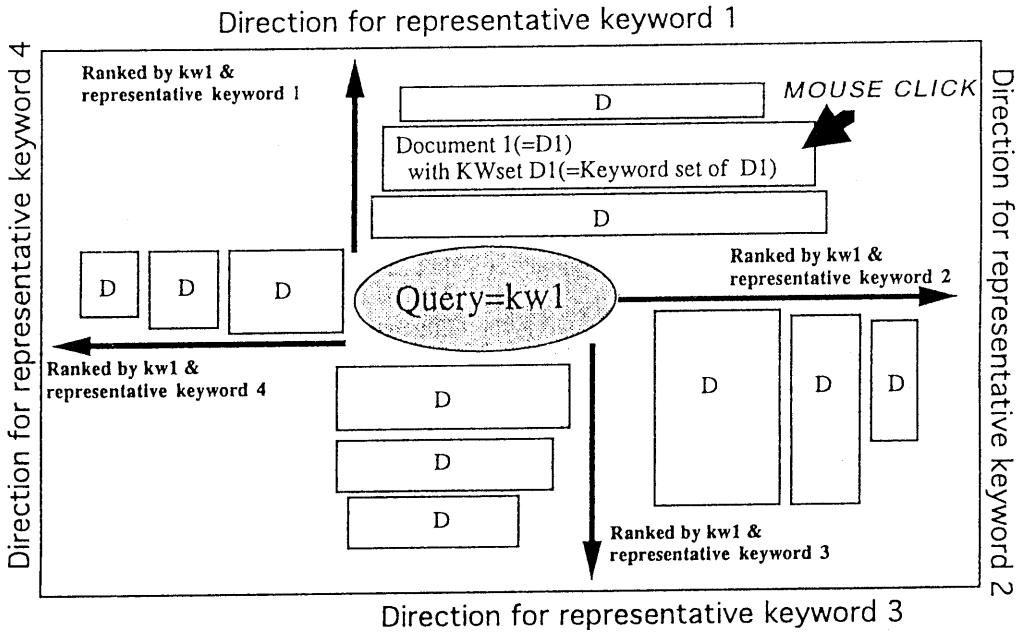
The data items, which are documents, are ranked according to the relevance points derived from queries issued by the user. The relevance points are calculated by the degree of matching between indexed keywords of the data and the keywords of the query.

(2)The multi-directional displaying function.

The data items are displayed in multiple directions and in order of ranking on a screen. Each direction is signified by one of the representative keywords from the database. The representative keywords are the ones appearing with high frequency in a database. The making of directions using representative keywords is described in Section 6. Examples of the multi-directional display are shown in Fig.1.

(3)The size-varying display function.

●The first layout on the display



●The next layout on the display

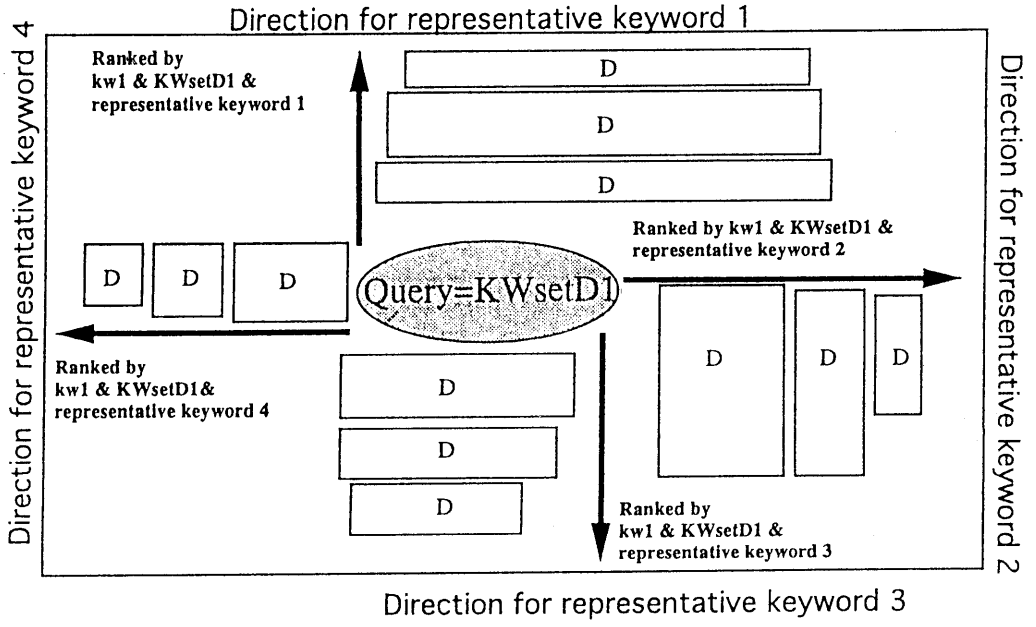


Fig.1 Multi-directional information display on a screen.

The size of each of the displayed data is derived from its relevance point. So, the display size of each datum is different, so that the more relevant the datum is, the larger it is displayed.

(4)The interactive search function.

The user can select the direction by clicking one of the data in that direction. The clicked datum will be located in the centre of the screen in the next screen layout. So, the user can change or select the direction by interacting with the system.

(5)The recursive display function.

The system displays the data recursively in a multi-directional layout during the interaction with the user. When the user selects one of the data on a screen using the interactive search function, the recursive display function displays that datum in the centre of the screen, and from this centre the new directions are displayed. This means that the focus of search can be selected by the user, who can change it with each interaction and follow his own directions of search, enabling a satisfactory search process.

The following are the functions that the author thinks, would be effective for an interactive MMIR with a GUI. In the next prototype they would be implemented.

(1)The display-time control function.

(2)The handling of image data as an object in a 3D space. This function includes texture mapping of data onto several streams and onto the surface of a rolling sphere.

(3)Perspective-projection of images.

(4)Interactive search between the user and the system in a 3D space

(5)The combination of the above mentioned functions.

6.Producing the multi-directional information layout on the screen

Each direction is derived from one of the representative keywords of a database. The representative keywords are the ones appearing with high frequency in the database. The schema for the extraction of the representative keywords is shown in Fig.2. The representative keywords are indexed to many data items in the database, so by using them for retrieval, many items are retrieved and displayed on a screen.

The following is the procedure for making the multi-directional information layout on the screen:

Step 1:

The representative keywords are extracted from the database.

Step 2:

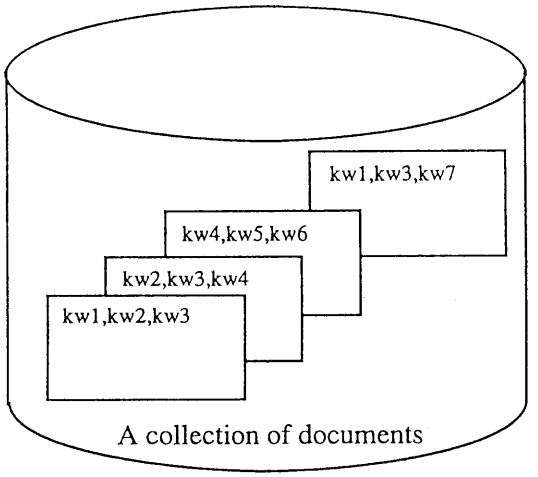
Keywords for search are input by the user. Here, these keywords are called KW1.

Step 3:

Directions are made from the top four frequency keywords. Hence, the INMUL prototype system displays four directions. For each direction, data are ranked by relevance points. The relevance points for data in direction 1 are calculated by the degree of matching between the keywords of the data and both KW1 (keywords input by a user) and the top frequency keyword (which is called Representative keyword 1). The relevance points for the data in the other directions are calculated in the same way. Then, using these relevance points for each direction, the data items are displayed in four directions in ranked orders.

Step 4:

The user clicks one of the displayed data for interactive search. Here, the keyword set indexed to the clicked data is called KW2 (in Fig.1 shown as KWsetD1). Again the



Representative keywords
(eg. kw3,kw1,kw2,kw4)

Keyword	Total frequency in the collection of documents
kw3	3
kw1	2
kw2	2
kw4	2
kw5	1
kw6	1
kw7	1

Fig.2 Extraction of representative keywords from a collection of documents

Table 1: The result of the experiment

ID No. of a query in BMIR-J1	The number of the relevant items in BMIR-J1	The number of the recalled items by the query	The increase/decrease number of the recalled items		
			By the first interaction	By the second interaction	By the third interaction
2	13	4	-1	+1(New)	—
3	5	2	0	0	—
4	4	2	0	—	—
5	14	4	-1	—	—
6	7	2	-1	—	—
7	5	4	0	-2	+1(New)
10	12	4	+1(New)	0	—
22	13	3	0	—	—

"—" means no trial.

relevance points are calculated using KW2 and KW1 and each representative keyword for each direction. Then the data are displayed in four directions on the screen. If each representative keyword is in KW2, the next top representative keyword is taken for that representative keyword and the relevance points are calculated again. So, the new directions will appear instead of the old ones.

Step 5:

Then, the user repeats step 4 recursively until he reaches the desired data items.

7.An experiment

An experiment was made using INMUL. A test collection used for the experiment was BMIR-J1[Keshi et al., 1996] that consists of 600 newspaper articles, 60 queries and relevancy judgements for each query. Eight queries were selected from the 60 queries. In case a query is a sentence, only nouns were selected from the query sentence, and these nouns were used as a query instead of a query sentence. This is because our NLP parsing program needs some modifications for the experiment. The result of the experiment is shown in Table 1 and an evaluation of this experiment is now under consideration.

8.Related issues and further studies

In the next prototype system, the functions described in Section 5 for the next system will be implemented and evaluated.

There is not yet an established evaluation method for interactive Information Retrieval. Appropriate evaluation methods for the proposed system would be minimization of the numbers of interactions before finding the desired item (or getting to the region of the desired item), or the maximization of the amount of data the user can see in a given time. More study is needed on evaluating the interactive system.

References

[Hemmje et al., 1994]M. Hemmje, C. Kunkel, A. Willett, LyberWorld - a visualization user interface supporting fulltext retrieval, in Proceedings of the annual international ACM SIGIR conference on research and development in information retrieval, pages 249-259, 1994.

[Keshi et al., 1996]I. Keshi, H. Kimoto, T. Tanaka, T. Ishikawa, Y. Masunaga, Y. Ogawa, J. Toyoura, T. Fukushima, T. Miyauchi, S. Miike, K. Matsui and T. Kitani, Overview of benchmark for Japanese IR system Ver.1.0(BMIR-J1)(in Japanese), IPSJ SIG Notes, Vol.96, No.11, pages139-146, 1996.

[Lin, 1991]X. Lin, A self-organizing semantic map for information retrieval, in Proceedings of the annual international ACM SIGIR conference on research and development in information retrieval, pages 262-269, 1991.

[Plaisant, 1994]C. Plaisant, Dynamic queries and visual information seeking interfaces, in Proceedings of the international symposium "Info-Tech '94", pages 26-32, 1994.

[Rao et al., 1995]R. Rao, J.O. Pedersen, M.A. Hearst, J.D. Mackinlay, S.K. Card, L. Masinter, P. Halvorsen and G.G. Robertson, Rich interaction in the digital library, Communications of the ACM, Vol.38, No.4, pages29-39, 1995.

[Williamson and Shneiderman, 1992]C. Williamson and B. Shneiderman, The dynamic HomeFinder: Evaluating dynamic queries in a real-estate information exploration system, in Proceedings of the annual international ACM SIGIR conference on research and development in information retrieval, pages 338-346, 1992.