Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information Processing and Management 42 (2006) 723–728
INFORMATION PROCESSING & MANAGEMENT
www.elsevier.com/locate/infoproman

An analysis of image retrieval behavior for metadata type image database

Toru Fukumoto

National Institute for Educational Policy Research, 6-5-22 Shimomeguro, Tokyo 153-8681, Japan

Received 11 November 2004; received in revised form 28 December 2004; accepted 5 January 2005
Available online 2 March 2005

Abstract

The aim of this paper was to analyze users' behavior during image retrieval exercises. Results revealed that users tend to follow a set search strategy: firstly they input one or two keyword search terms one after another and view the images generated by their initial search and after they navigate their way around the web by using the 'back to home' or 'previous page' buttons. These results are consistent with existing Web research. Many of the actions recorded revealed that subjects behavior differed depending on if the task set was presented as a closed or open task. In contrast no differences were found for the time subjects took to perform a single action or their use of the AND operator.

Keywords: Image retrieval; Metadata; Keyword; Searching behavior

1. Introduction

The number of images being stored in the Internet is increasing due to an increase in affordable digital recording devices, such as digital cameras and scanners. In order to effectively manage these digital images, an image album or an image filing system has become the subject of study. As the Internet grows so the task of image retrieval becomes more complicated. As a result the need to devise an effective method for retrieving images from a database becomes more important.

There are three kinds of image database: feature type, sensitiveness type, and metadata type. Feature type (Myron et al., 1995) is based on the colors or shapes of object in the images. When the retriever puts in color or shape, the system starts searching directly the database with color histograms or shape. In the

E-mail address: fukumoto@mail.club.ne.jp

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2005.01.008

case of sensitive type (Kimoto, 1999) the user inputs sensitive words and the system exchanges them for color information and searches the database. In the case of metadata type (Kiyoki, Kaneko, & Kitagawa, 1996) each image in the database is given a metadata tag which defines its characteristics through text or digits. In order to search the database the user has to use the corresponding metadata tag.

For the purpose of this study we focused on metadata image databases. In order to create a metadata type a database creator first needs to define the structure or the framework of the metadata. After this point the database administrator needs to attach the metadata to images within the database. It is only at this point that a user hoping to retrieve images can do so by entering the appropriate terms. The database is searched using the users search term and cross-checked against the metadata issued by the administrator. Examples of metadata are keywords, text, classification items and so on.

There are now accepted ISO (International Standards Organization) regulations governing the structure contents description of metadata for multimedia known as MPEG-7 (MPEG, 2001). MPEG stands for ''Motion Picture Expert Group'' although MPEG-7 puts little emphasis on the description of contents in the text it can extract features from them (i.e. Video stream, Audio and schema itself which store them).

2. The purpose of our research

Over the past years a considerable number of studies have been conducted on information retrieval studies for information over the Internet. Many of these studies have concentrated on text retrieval or the analysis of logs. More recently there has been a new method proposed for assessing information searching behavior on WWW (Goodrum & Spink, 2001; Miura & Fujiwara, 2001) however this only deals with text information not image based information. The Internet is rapidly expanding and the number of stored images on sites has increased therefore we argue that effective retrieval methods aimed at retrieving stored images from a database have become significantly more important. However there have not been many studies that have attempted to assess user's behavior on image information retrieval tasks. The aim of this study was to analyze user's behavior on image retrieval tasks.

3. Method

3.1. Participants

Twenty undergraduate students participated in this study. All were accustomed to using search engines on the Internet such as Google and Yahoo!

3.2. Materials

We developed the image database based on previous research (Fukumoto & Akahori, 2000, 2001). The number of images in this database was 1700 with the contents being varied. The design of the system can be seen in Fig. 1 with Apache, PHP3, PostgreSQL and a retriever accesses it with web browser. The size of image is 200×133 (thumbnail) and 600×400 (original, it is showed when a retriever magnifies). There are 20 images per one page. A retriever can change $5, 10, 15, 20$ with pull down menu. Cache files of browser are deleted every task.

3.3. Procedure

Subjects were shown an image and asked to retrieve it so that they could become accustomed to the system. We repeated this procedure twice with a different image on each trial (see Fig. 2). These initial tasks

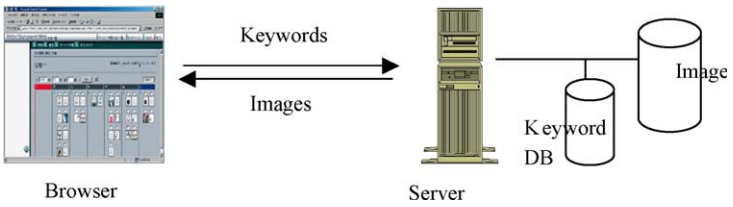

Fig. 1. System design.

Fig. 2. Images of Task1-1 and Task1-2.

were labeled Task 1-1 and Task 1-2, respectively. We then issued instructions Task 2-1 and Task 2-2 to the subjects asking them to search for images that would be suitable for a "Summer Greetings Card" and "Favorite Landscape" respectively and to retrieve an image/s from the database. We did not impose a time limit on this exercise nor did we restrict the number of images that subjects could retrieve, these tasks were labeled Task 2. During both tasks subject's behavior was recorded on video.

4. Results

4.1. Categorization of actions

Table 1 shows the ratio of subjects' actions. Subjects actions were classified using an existing classification method for Web Search behavior (Miura & Fujiwara, 2001), the main features being (1) Keyword operation, (2) Web page operation, (3) Browser operation, and (4) Other (please see Table 1). On all tasks Keyword operation occupied from 40% to 50% of the subjects time. In contrast during subjects Web search, Web page operations occupied about 40% of subjects time. Keyword operations on all Input tasks occupied 40% of subject's time. 'Back' operations occupied from about 10% to 30% of their time. There were a few add, delete or replace actions. Subjects tended to act as follows: first, to input one or two keywords one after another and allow the system shows the images as the result of their search and then they would operate the 'back to home' or 'previous page' button. These results concur with existing Web behavior research (Miura & Fujiwara, 2001).

4.2. Process of search

Table 2 shows the ratio of subjects' search processes, that is, page, action, time, and average time per action. We counted the averages for each process and examined the differences between all tasks.

The result of *t*-test found that three processes, Pages, Actions, and Time for Task2 were significantly more used than those on Task1-1 and Task1-2 [$df = (19)$, t-value = 3.27,4.83,4.90 $p = .01$]. But there are

Table 1
Categorization of subject's actions

		Task1-1 (%)	Task1-2 (%)	Task2-1 (%)	Task2-2 (%)
Keyword operation	Input	50.48	55.03	40.29	44.26
	Add	1.67	3.13	1.49	0.83
	Delete	7.60	2.50	1.11	2.56
	Replace	1.00	0.71	0	7.14
Web page operation	Magnify	7.75	10.28	8.78	9.49
	Select link	0	0.56	0.29	0
	Select pull-down	0	0.42	0.19	0
Browser operation	Forward	0	12.56	10.66	19.62
	Back	25.65	9.56	30.01	15.66
	Home	0	0.79	4.88	0
	Jump	0	0	0	0.43
	Subtotal	96.14	98.66	97.71	100
Other	Other	3.86	0.71	2.29	0

Table 2
Ratio of search processes

	Task1-1	Task1-2	Task2-1	Task2-2	Significance
Pages	1.00	2.80	5.55	4:20	$t = 3.27^{**}$
Actions	4.70	4.95	22.40	19.20	$t = 4.83^{**}$
Time (MM:SS)	01:20	01:30	06:30	05:45	$t = 4.90^{**}$
Time/actions	00:21	00:19	00:17	00:22	$t = 0.98$

** $p = .01$.

no significant difference between all processes of Task1-1 and those of Task1-2 [t-value = 0.98 $p > .1$]. These results illustrate the differences in the nature of the tasks. Task1-1 and Task1-2 were presented as closed task questions. That is that subjects had to find images that they had already been shown. In contrast, Task 2-1 and Task2-2 were based on open task questions and subjects were able to choose whatever images they felt best matched their interpretation of the question. For this reason we suspect that subjects took longer to search database for suitable images.

For set actions such as inputting a keyword or going back to a previous page no significant time differences in the execution of these tasks was found. [t-value = 0.98 $p > .1$]. We believe that this characteristic is not dependent on the task but is instead a personal characteristic.

4.3. Keywords of search

Table 3 shows the ratio of subjects' keyword operations, that is, the number of input keywords, the number of unique keywords, and the depth of the AND operation. We counted the averages for each process and looked to see if there were any significant differences between the tasks.

Table 3
Ratio of keyword operation

	Task1-1	Task1-2	Task2-1	Task2-2	Significance
Inputting keywords	2.60	1.85	9.30	4.70	$t = 4.35^{**}$
Unique keywords	2.50	1.80	7.65	4.30	$t = 4.06^{**}$
Depth of AND	1.45	1.25	1.25	1.40	$t = 1.25$

Two operations, inputting and uniqueness at Task2-1and Task2-2 are significantly more used than that of Task1-1 and Task1-2 at 1% level. And Task1-1 is no significant difference from Task1-2. But Task2-1 is significantly more used than that of Task2-2. These results suggest that the differences in the nature of the tasks (closed or open task question) significantly affected how subjects performed. And Task2-2 is easier to hit upon keywords and input them than Task2-1 for subjects in familiarity with this theme (Favorite Landscape). However there were no significant differences found for the fixed depth of the AND operation for either condition or between subjects.

5. Discussion

There are two kinds of processes described on Table 1. One is a search function that is keyword operations such as input, add, delete and replace; the other is browse, that being Web-page operations and browser operations such as magnify, forward, back and so on. According to Table 1, the ratio of these two processes for Task1-1 and Task1-2 is 6:4, and for Task2-1, 4:6 Task2-2, 5:5. Task1-1 and Task1-2 were closed task questions so the search strategy used was more important relative to browsing. In contrast, Task2-1 and Task2-2 were the reverse of this situation with browsing being more important that the search strategy.

A lot of processes, Pages, Actions, Time, Inputting keywords and Keyword uniqueness for Task 2 were found to differ significantly when compared to the same processes in Task1-1 and Task1-2. We believe that this result was due to the nature of the tasks. In Task 2 subjects were presented with an open task question, whereas in Task 1-1 and Task 1-2 subjects were presented with a closed task question that resulted in them having to retrieve specific images and not ones they thought best fit the experimenters instructions.

These findings suggest that subjects tend to change their strategy depending on a kind of task that was set. Subjects tended to follow a set pattern of behavior: first, to input one or two keywords one after another and view the results of their search and then operate the 'back to home' or 'previous page' button. Input actions occupied 40% of users time and back operations occupied from about 10% to 30%. It was also found that subjects took a fixed amount of time to execute a function regardless of the task and this was also true for their use of the AND function. We believe the findings generated by the single function and the AND operator were the result of the users own personal characteristics and not therefore not effected by the tasks.

This result (the usage of user's behavior on image retrieval tasks) suggests the following; that a system should be developed that has a fast-browse interface to allow users to search for images that users are already familiar with. An interface that allows the user to choose between options such the comparison interface found on on-line shopping carts. This type of comparison interface would allow users to view at a glance the images retrieved on the basis of their input-query and give them the opportunity of quickly viewing a selection of images before they decide if they wish to revise their search or continue browsing the images they have been presented with.

6. Conclusion

In this paper, we sought to analyze user's behaviors on image retrieval tasks. Users behavior tended to follow a set pattern that being; firstly to input one or two keywords one after another and to view the results generated by their search and then to operate the 'back to home' or 'previous page' button. These results concur with existing Web search research. More actions were executed when the task was presented as an open question as compared to a closed one and these differences were found to be significant. These results lead us to suggest that users require a 'quick-browse' interface or a 'choice and comparison' interface

dependent upon the aim of their search. No significant time differences were found in users executing an action nor were there any significant differences found in their use of the operation AND.

This research shows to need a fast-browsing or comparison system according to open or close task. My future works are as follows:

- With larger test set: participants and materials.
- Developing a fast-browsing or comparison system and evaluate them.

In the future, we hope that multimedia retrieval using metadata will be easier for all individuals. When viewed from an educational viewpoint we believe that image retrieval would aid teachers in allowing them to retrieve and use more accurate images in order to support their subjects' existing course materials and for students by allowing them to easily retrieve relevant images from the Internet.

References

Fukumoto, T., & Akahori, K. (2000). The criteria and evaluation of metadata/keywords in image retrieval. In *Proceedings of the ICCE/ICCAI 2000* (pp. 542–546).

Fukumoto, T., & Akahori, K. (2001). An analysis of image cognition structure for metadata type image retrieval. In *Proceedings of the ICCE/SchoolNet 2001* (pp. 367–371).

Goodrum, A., & Spink, A. (2001). Image searching on the excite web search engine. *Information Processing and Management, 37*(2), 295–311.

Kimoto, H. (1999). An image retrieval system using impressional words and the evaluation of the system. *Transaction of IPS Japan, 40*(3), 886–898.

Kiyoki, Y., Kaneko, Y., & Kitagawa, T. (1996). A semantic search method and its learning mechanism for image databases based on a mathematical model of meaning. *Transaction of IEICE Japan, J79-D-2*(4), 509–519.

Miura, A., & Fujiwara, N. (2001). Experimental study of information searching behavior on WWW—Categorization of behaviors and examination of difference between tasks. *Transaction of Japanese Society for Information and Systems in Education, 18*(1), 121–128.

MPEG Web Page (2001). URL: http://www.chiariglione.org/mpeg/.

Myron, F., Harpreet, S., Wayne, N., Jonathan, A., Qian, H., Byron, D., et al. (1995). Query by image and video content: The QBIC system. *IEEE Computer, 28*(9), 23–32.