The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0022-0418.htm

Searching for digital images on the web

Bernard J. Jansen
College of Information Sciences and Technology, The Pennsylvania State University, University Park, Pennsylvania, USA

Received 13 February 2006
Revised 18 August 2007
Accepted 20 September 2007

Abstract

Purpose – The purpose of this paper is to examine the way in which end user searching on the web has become the primary method of locating digital images for many people. This paper seeks to investigate how users structure these image queries.

Design/methodology/approach – This study investigates the structure and formation of image queries on the web by mapping a sample of web queries to three known query classification schemes for image searching (i.e. Enser and McGregor, Jörgensen, and Chen).

Findings – The results indicate that the features and attributes of web image queries differ relative to image queries utilized on other information retrieval systems and by other user populations. This research points to the need for five additional attributes (i.e. collections, pornography, presentation, URL, and cost) in order to classify web image queries, which were not present in any of the three prior classification schemes.

Research limitations/implications – Patterns in web searching for image content do emerge that inform the design of web-based multimedia systems, namely, that there is a high interest in locating image collections by web searchers. Objects and people images are the predominant interest for web searchers. Cost is a factor for web searching. This knowledge of the structure of web image queries has implications for the design of image information retrieval systems and repositories, especially in the area of automatic tagging of images with metadata.

Originality/value – This is the first research that examines whether or not one can apply image query classifications schemes to web image queries.

Keywords Worldwide web, Search engines, Information retrieval

Paper type Research paper

1. Introduction

The tremendous growth in the quantity of digital images is driving the need for more effective methods of storing, searching, and retrieving image data. With the development of such technologies, image indexing and retrieval are receiving increased attention. However, there has been limited research concerning image searching on the web, especially concerning how web users search for images in the web environment. Given the pervasiveness of the web, the manner in which users conduct image searches may also be good indications of how they will conduct future searches in other environments. Such research would improve the design of a variety of multimedia retrieval systems (Lim and Jin, 2002; Wu and Zhang, 2002).

Current image retrieval approaches are either concept-based, content-based, or a concept-content combination. In the concept-based approach, image retrieval research focuses on the retrieval of images utilizing indexed image collections (Goodrum and Kim, 1998). Researchers and practitioners have created various thesauri for visual

Journal of Documentation
Vol. 64 No. 1, 2008
pp. 81-101

0022-0418
DOI 10.1108/00220410810844169

information (see, the Library of Congress Thesaurus, Greenberg, 1993). Image system and indexers then use these thesauri to index images within a collection.

Many web search engines use textual clues within web documents to automate this concept-based approach. Marsh and White (2003) researched how an image relates to the surrounding text. These web search engines use text surrounding the multimedia object, along with other clues such as files names of multimedia content. This approach leverages the assumption that these textual clues relate to the object. Although many times valid, this assumption does not always hold. For example, software programs for desktop computers and digital cameras automatically generate file names that are random or temporal character sequences. However, this approach using surrounding textual clues has proven generally effective.

In the content-based approach, the focus has been on indexing images at the pixel level and the implementation of search features using pixel comparison (Wang, 2000). Content-based image information retrieval (IR) systems allow users to search image collections using color, texture, shape, and spatial similarity. These systems many times also provide text-based search functions for notations and text descriptions embedded within images. New technologies have placed the emphasis on content-based retrieval, with commercial systems such as the MediaSite.com (www.mediasite.com) system. On the web, content image retrieval systems such as WebSeek [3] and SingingFish (www.singingfish.com/) provide a variety of multimedia files to web searches.

The context-content hybrid approach generally uses some type of machine-learning method to teach the system to annotate images based on pixel-level information and to use computer-annotated words for retrieval. This potentially allows users to type in keywords for retrieval of un-annotated images in a manner very similar to the way users search for text. This third approach has the advantage of reducing the cognitive load of learning additional searching methods or jargon. Researchers are investigating the feasibility of this approach (Barnard *et al.*, 2003; Li and Wang, 2003). Some commercial systems already use this approach for general document meta-tagging (see, Verity K-2 Architecture).

However, it is not clear how the retrieval functionality of any of these approaches correlates with the image information needs of real users. Do searchers, especially web searchers, understand the concept-based indexing schemes? Do these schemes match the characteristics in which searchers articulate their queries? Does the pixel representation used by the content-based approach correlate with searchers' expression of their information needs? Are there trends in image searching that even permit automated meta-tagging of images?

There appears to be a lack of collaboration between the practitioner community who are concerned with manual indexing methods, and the research communities in computer science and computer vision who focus on automatic methods (Jörgensen, 2003). With the context approach, previous studies indicate that experts who classify images are not a good source of terms that are preferred by real users (Rasmussen, 1997). Additionally, it appears that users seldom search using content characteristics (Chen, 2001). Are these same characteristics observable in web searching?

Although user studies on image searching within specific domains have been conducted, research has shown that web users differ in their interaction with IR systems relative to users in other environments (Jansen and Pooch, 2001). In this

respect, studies concerning web users' image needs and how they express these needs in queries would be beneficial for the effective and efficient design of image IR systems that use the context, content, or hybrid approaches.

The research results presented in this article focus on the information needs of web users as represented by their image queries. Using image queries submitted to a major web search engine, we classify user queries using three published classification schemes for image queries and compare the characteristics to those identified in previous studies. We report on the sufficiency of these classification schemes in relation to actual queries for images by web searchers. We provide suggestions for improving image indexing and retrieval systems to meet the information needs of real users.

2. Review of literature

2.1. Relationship between user information needs and image attributes

Previous researchers (Ellis, 1984; Fidel, 1997; Jörgensen and Jörgensen, 2005) have noted that image retrieval research has ignored several fundamental user issues, including the differences between text and image retrieval, the image attributes important for retrieval, and the characteristics of users image queries. These user issues are especially germane with respect to image searching on the web, given the limited number of studies in this area.

For example, Greisdorf and O'Connor (2002) posits that the relevance of retrieved images for searchers may arise from both concept and content-based elements that are not even present in an image. Additionally, emotion-based query terms appeared to be an important descriptive category for image retrieval. Hertzum (2003) analyzed a year of e-mail requests to a large film archive to study what types of information needs real users have and how these needs are expressed. The findings include that the requesters make use of a broad range of need attributes in specifying their information needs. As much as 43 percent of the image requests contain no information about the context that gives rise to the request.

Choi and Rasmussen (2003) collected queries from 48 scholars working in the Library of Congress's American Memory photo archive. The researchers classified the queries into four categories, namely specific, general, abstract, and subjective. General queries accounted for 60.5 percent, and specific 26.3 percent. Analyzing transaction logs from professional image searchers, Jörgensen and Jörgensen (2005) report heavy use of Boolean operators and considerable query modification.

2.2. Web user information needs and image queries

These findings have significant implications for image retrieval on the web. The web is an immense repository of image information (Lesk, 1997), with an estimated 180 million images on the publicly indexed web and 3Tb of image data (Lawrence and Giles, 1999), not including other types of multimedia files, such as audio and video. Prior research has noted that constructing effective search strategies for images is difficult (Large *et al.*, 1998), although Fukumoto (2006) reports that searchers for images tend to follow a simple search strategy.

However, there has been limited research on image searching in web studies (Goodrum and Spink, 2001; Jansen *et al.*, 2000). Goodrum and Spink (2001) analyzed image queries and Jansen *et al.* (2000) analyzed audio, image, and video sessions from

the Excite web search engine over multiple years. On average, there were 3.36 image queries per user containing an average of 3.74 terms per query. Image queries contained a large number of unique terms. The most frequently occurring image related terms appeared less than 10 percent of the time, with most terms occurring only once (Goodrum and Spink, 2001).

Table I compares audio, video and image queries for the 1997, 1999 and 2001 Excite data sets.

Table I shows that by 2001, audio queries were more predominant than video and image queries, with 52.7 percent of multimedia queries being audio queries, 21.9 percent were video queries and 25.4 percent were image queries. This may be due to the development of mp3 technology and the file sharing software that allowed the free-exchange of audio files. Audio queries were not included in the 1999 dataset analysis.

The researchers noted that both sessions and queries are generally longer for multimedia than general web searching, which may indicate an increased cognitive load for multimedia searching. The number of multimedia sessions originating from the general search box decreased from 1997 to 2001 as a proportion of general queries, due to the introduction of multimedia buttons near the query box. Overall, we see multimedia web searching undergoing major changes as web content and searching interfaces evolve (Özmultu *et al.*, 2003).

All three studies (Goodrum and Spink, 2001; Jansen *et al.*, 2000; Özmultu *et al.*, 2003) were on general web searching and did not examine the effect of separate multimedia collections on web searching. Jansen *et al.* (2003) did examine searching on federated multimedia collections on AltaVista In their analysis, Jansen *et al.* (2003) compared general web, audio, image, and video web searching. The aggregate results are presented Table II.

In comparing the four types of searching (general, audio, image, and video) presented in Table II, we see that the mean terms per query for image searching was notably larger (four terms per query) than the other categories of searching, which were all less than three terms. The session lengths for image searchers were longer than any other type of searching, although video sessions were also relatively lengthy. The session lengths of image searches when combined with the longer queries may indicate that image searching is a more difficult cognitive task than other types of searching. Another indicator of the complexity of image searching is Boolean usage,

	Audio queries			Video queries			Image queries		
	1997	1999	2001	1997	1999	2001	1997	1999	2001
	3,810		9,655	7,630	17,148	4,011	27,144	22,190	4,651
Percentage of data set	0.37	–	0.9	0.7	1.6	0.3	2.6	2.1	0.4
Percentage of Multimedia Queries	9.80	–	52.7	19.8	43.5	21.9	70.4	56.4	25.4
Mean queries per session	2.4	–	2.6	2.9	3	2.6	3.2	3.4	2.8
Median queries per session	2	–	1	2	–	1	2	–	1
Maximum queries per session	51	–	119	70	–	59	267	–	102
Standard deviation for queries per session	2.9	–	4.8	3.8	–	3.8	5.4	–	5

Source: Özmultu *et al.* (2003)

Table I.
Comparison of audio and video queries from Excite studies

	General	Audio	Image	Video
Sessions	369,350	3,181	26,720	5,789
Queries	1,073,388	7,513	127,614	24,265
Terms				
Unique	297,528 (9.5%)	6,199 (33.4%)	71,873 (14.1%)	8,914 (19.1%)
Total	3,132,106 (100%)	18,544 (100%)	510,807 (100%)	46,708 (100%)
Mean terms per query	2.91 (sd = 4.77)	2.47 (sd = 1.62)	4.00 (sd = 3.21)	1.92 (sd = 1.09)
Mean queries per user	2.91 (sd = 4.77)	2.36 (sd = 3.85)	4.78 (sd = 10.44)	4.19 (sd = 6.14)
Users modifying queries	193,468 (52%)	1,496 (47%)	14,838 (56%)	3,350 (58%)

Source: Jansen *et al.* (2005)

Table II.
Comparison of general, audio, image, and video searching in 2002

which was 28 percent. This is more than four times the next highest category of general web searching. Jansen *et al.* (2007) report that 19 percent of queries submitted on the Dogpile search engine were executed on the image collection.

These studies (Goodrum and Spink, 2001; Jansen *et al.*, 2005; Özmultu *et al.*, 2003; Spink and Jansen, 2004) provide insight into some of the session and query characteristics of image searching by web searchers. However, none of these studies classified web image queries using known classification schemes. Without such research, one cannot determine whether the web user population's image information needs conform to currently accepted image searching approaches. More importantly, an understanding of how web users currently search for images may lead to better design of future IR systems.

2.3. Image retrieval systems on the web

In general, web users search for multimedia information as they would search for textual information. The advantage of this approach is that image searching is performed in an identical manner to text searching (Schauble, 1997). No additional burden is placed on the searcher. If the searcher desires an image document, the searcher just enters the query as in general web searching. The disadvantage of this approach is that it places more contextual knowledge burden on the searchers, who may not be familiar with image formats. Necessitating that users translate a non-textual information need into a textual query may challenge a user's cognitive load, creating a semantic gap (Gudivada and Raghavan, 1995).

In an enhancement to the query box, some web IR systems specialize in image collections or provide mechanisms for users searching for images (e.g. radio boxes or media specific search syntax). AltaVista (www.altavista.com) and Google (www.google.com) searchers can narrow a query to specifically search for an image. ClipArt Searcher (www.Webplaces.com/search/) and WebSeek (www.ctr.columbia.edu/Webseek/) are examples of image specific retrieval applications allowing users to search by terms or select from general categories of images and video. Most web search engines return thumbnail images and file names in the document result list. WebSeek also provides tools for content-based searching for images and videos using color histograms generated from the visual scenes. FaganFinder (www.finderfinder.com/img) provides a comprehensive listing of image search engines.

Most of the general web search engines use some sort of tabs or radio boxes for narrowing aspects of multimedia searching, an approach shown to reduce the cognitive load of web searchers (Jansen *et al.*, 2003). The use of tabs attempts to address the semantic gap (Gudivada and Raghavan, 1995) that occurs between the textual expression of a multimedia information need and the actual multimedia content.

Figure 1 shows the A9 search engine (www.a9.com) interface with web results and image results frames. The A9 search engine uses a radio button approach. The results are in response to the query drivers.

Figure 2 shows the Yahoo! search engine (www.yahoo.com) with multimedia tabs for audio, image and video searching. The results page is displaying image thumbnails in response to the query cobra. Notice the variety of domain results for this query term. The search engine also offers a variety of other searching options via hyperlinks.

In terms of image indexing, most web IR systems follow a relative straightforward approach. The simplest image search algorithm locates images files by searching for file extensions, matching the filename to terms in the query (Witten *et al.*, 1994). This may not be a good approach, especially given the move to automatic generation of image file names by applications (e.g. converting Power Point © presentations to web pages or transferring files from a digital camera). In these cases, the file name has little to nothing to do with the image. web IR systems may also retrieve documents that are primarily textual but that contain embedded images. In these cases, the image filename

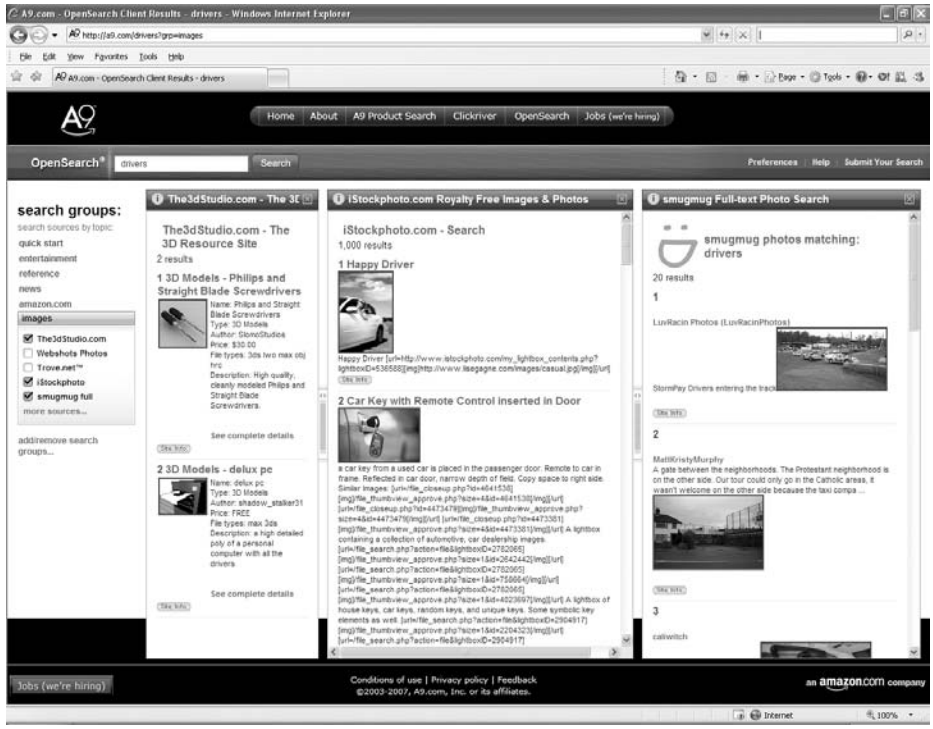


Figure 1.
A9 Search engine with web results and image frames

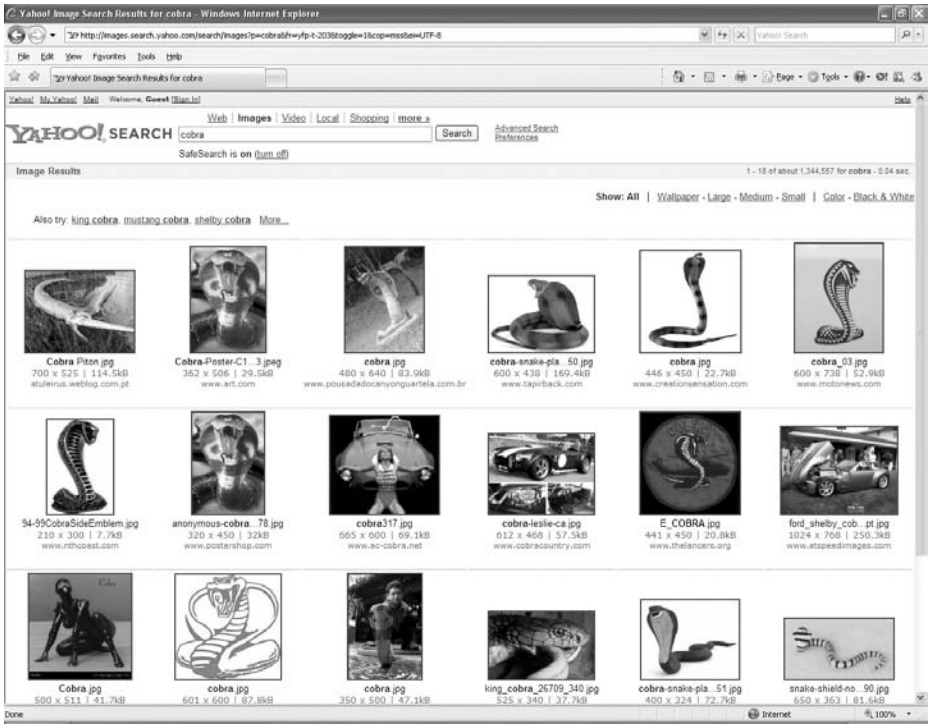


Figure 2.
Yahoo! Search image tab with thumbnails displayed

may not match the query terms, but the web documents may contain text that does match the query terms. The assumption being that the image relates to the text.

There has been some research assessing the effectiveness of approaches taken by web search engines. Automatic assignment of textual attributes has been conducted using captions from still images, transcripts, closed captioning, and verbal descriptions for the blind that accompany videos (Turner, 1998). Swain (1999) views text cues extracted from HTML pages and multimedia document headers, supplemented by off-line analysis, to be the primary sources of information for indexing multimedia web documents. However, Lawrence and Giles (1999) report that the use of web metadata tags is not widespread.

Although most web search engines pursue automated classification of images based on surrounding textual clues, there has been little reported effort on implementing classification of users' image queries to aid in image retrieval.

3. Classification schemes

Outside of the web domain, there is a variety of classification schemes. Jörgensen (2003) and Smeaton (2004) review a number of unique methodologies for image classification. Other researchers have examined classification schemes for image queries. A review of three previously published image query classification methodologies follows.

3.1. Enser and McGregor classification categories

The researchers (Enser and McGregor, 1992a) analyzed 2,722 image queries requests on an image system. The combined size of the collection was about 10 million images. The researchers classified these 2,722 requests in four categories:

(1) unique;
(2) unique with refiners;
(3) non-unique; and
(4) non-unique with refiners.

The property of uniqueness is a request for the visual representation of an entity where the desired occurrence can be differentiated from every other occurrence of the same entity type. Using refiners, one can narrow these queries to some degree of specificity. The researchers noted that almost 70 percent (1,905) of the queries were for a unique person, object or event, and 34 percent (925) of the queries were refined, mostly by time. As examples, Bill Clinton would be unique. Bill Clinton 1980 would be unique with refiner. The query middle aged man would be non-unique. The query middle aged man sitting by desk would be non-unique with refiners.

3.2. Jörgensen 12 image attribute classes

The researcher (Jörgensen, 1998) conducted a study with 107 first semester library science students with a variety of undergraduate majors. The study participants viewed six projected color images that were randomly selected from *The Society of Illustrators 25th Annual of American Illustration*. The participants described the six projected images, and Jörgensen then classified the descriptions into the 12 classes of image attributes. Jörgensen's (1998) 12 attribute classes are:

(1) abstract;
(2) color;
(3) content;
(4) description;
(5) historical;
(6) location;
(7) object;
(8) people;
(9) people-related;
(10) relationship;
(11) response; and
(12) visual.

3.3. Chen's categories and attributes modifications

Chen (2001) collected 534 queries from 29 college students majoring in art history, using pre-search and post-search questionnaires. The art course students utilized an image collection containing over 75,000 volumes and subscriptions to over 200 journals in relevant fields. The collection also contained significant holding in several art

related fields and facsimile collections of medieval manuscripts and original scrolls. Chen did not focus on whether or not the queries were utilized on an image retrieval system. Chen mapped the queries to previous query classification methods, including Enser and McGregor (1992b) categories, combining some of Jörgensen (1998) attribute classes. Based on Chen's (2001) analysis, the researcher proposed modifications to Enser and McGregor (1992b) categories – as shown below:

(1) location;
(2) literal object (unique names);
(3) art historical information;
(4) people;
(5) people-related;
(6) literal object (non-unique names);
(7) color;
(8) visual elements;
(9) description;
(10) abstract concepts;
(11) content/story;
(12) external relationships; and
(13) viewer response.

3.4. Applicability of these classification schemes to web image searching

Little research has examined the relative effectiveness of these various approaches to image indexing or retrieval using web search engines. Given the inadequate research concerning user image needs on the web and a lack of theoretical background for the design and evaluation of web image databases (Chen, 2001), it is logical to examine both web image queries and current image query classification schemes to discover the utility of these models for the design of web-based image retrieval systems. This is the motivation for this research.

The next section of the paper describes the research design used in this study.

4. Research design

This research reports the analysis of image queries submitted by real users with real information needs to a major web search engine. The objective was to map these image queries using known image query classification features, evaluating both the queries and the image classification schemes.

4.1. Research questions

The following research question is addressed: Can web image queries be effectively mapped using the previously identified image classification schemes by Enser and McGregor, Jörgensen, and Chen?

4.2. Research procedure

The images queries we obtained for this research had been submitted to Excite, a major web search engine at the time of data collection. Excite provided a transaction

log holding a large and varied set of queries. The transaction log spanned 24-hours of user searching on 30 April 2001 (Monday, midnight to midnight) and contained nearly one million queries. Excite was the fifth most popular in 2001 as measured by number of unique visitors (*Cyber Atlas*, 2000).

Each record within the transaction log contains three fields:

(1) *Time of day*. Measured in hours, minutes, and seconds from midnight of each day as logged by the web server.

(2) *User identification.* An anonymous user code assigned by the Excite server

(3) *Query terms*. Terms exactly as entered by the user.

With these three fields, we located a user's initial query and recreated the chronological series of actions by each user in a session, from which one could calculate standard web searching metrics (Jansen and Pooch, 2001).

Using a modified snowball technique (Patton, 1990), we extracted image queries from the entire data set using the image terms presented in (Jansen *et al.*, 2000), resulting in approximately 4,500 image queries. We automatically assigned a unique identifier to each query record, and then randomly selected 587 queries from this data set in order to provide a robust sample for analysis.

4.3. Content analysis

Three independent evaluators examined each one of the 587 image queries, identifying characteristics of the query using Enser and McGregor's categories and Chen refiners as a starting point, adding new refiners if the appropriate one did not exist. That is, the evaluators would evaluate a query, determining if the query was unique, unique with refiner, non-unique, or non-unique with refiners. If they evaluated the query as containing a refiner, they would then state what category that refiner belonged to using Chen's refiners as a starting point. Each evaluator would then rate the confidence of their rating from 1 (just guessing) to 7 (totally sure).

Each evaluator then evaluated the query using Jörgensen's attributes as a starting point, determining which if the 12 attributes the query contained and again adding new attributes if the appropriate one did not exist. The evaluators would then rate the confidence of their evaluation, again from 1 (just guessing) to 7 (totally sure).

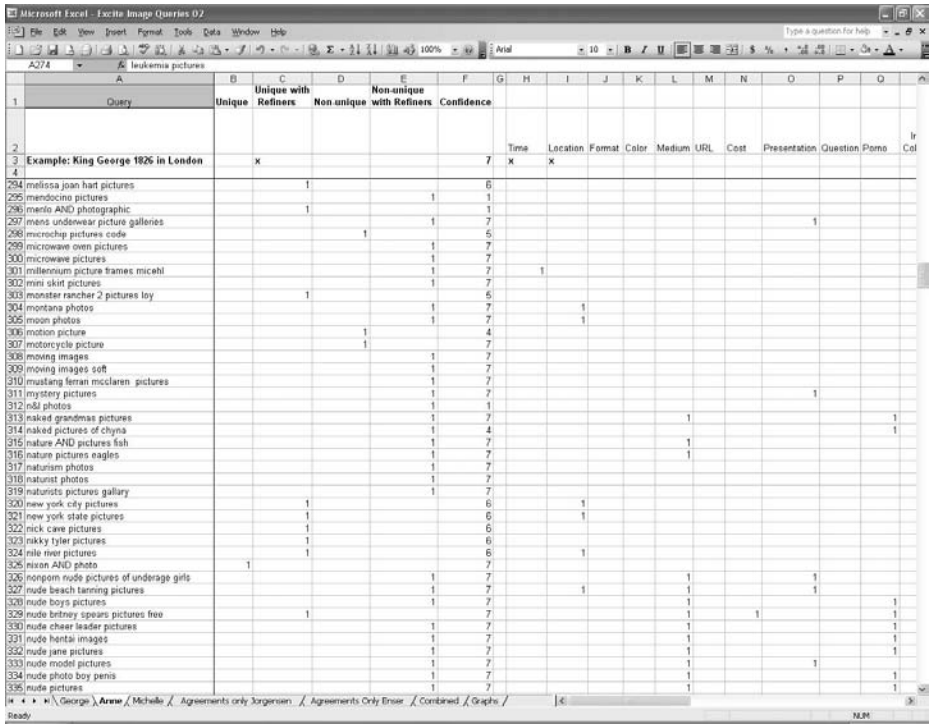
We provided each evaluator with training, walking through a query evaluation with them as they thought aloud. We also provide the evaluators with written instructions outlining the steps of query classification and the definitions of each of the refiners and attributes. Two or more of the reviewers had to agree on the classification of a query to be counted within a particular category, refiner, or attribute. A query could be mapped to more than one refiner or attribute. Therefore, the number of mapping incidents for refiners and attributes is greater than the number of queries. An example of an evaluation worksheet is shown in Figure 3.

We now present the findings of our research.

5. Research findings

5.1. Mapping results using Enser and McGregor's categories

The image queries were mapped to Enser and McGregor's image categories. From the set of 583 image queries, there were 73 (12 percent) not utilized in the Enser and McGregor comparison due to invalid reviewer agreements (i.e. all three reviewers rated



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Query	Unique	Unique with Refiners	Non-unique	Non-unique with Refiners	Confidence												
2								Time	Location	Format	Color	Medium	URL	Cost	Presentation	Question	Porno	Ir Col
3	Example: King George 1826 in London		x			7		x	x									
4																		
294	melissa joan hart pictures		1			6												
295	mendocino pictures				1	1												
296	menlo AND photographic		1			1												
297	mens underwear picture galleries				1	7									1			
298	microchip pictures code			1		5												
299	microwave oven pictures				1	7												
300	microwave pictures				1	7												
301	millennium picture frames micehl				1	7			1									
302	mini skirt pictures				1	7												
303	monster rancher 2 pictures loy		1			6												
304	montana photos				1	7				1								
305	moon photos				1	7				1								
306	motion picture			1		4												
307	motorcycle picture			1		7												
308	moving images				1	7												
309	moving images soft				1	7												
310	mustang ferrari mcclaren pictures				1	7												
311	mystery pictures				1	7									1			
312	n&l photos				1	1												
313	naked grandmas pictures				1	7						1					1	
314	naked pictures of chyna				1	4											1	
315	nature AND pictures fish				1	7						1						
316	nature pictures eagles				1	7						1						
317	naturism photos				1	7												
318	naturist photos				1	7												
319	naturists pictures gallery				1	7												
320	new york city pictures		1			6				1								
321	new york state pictures		1			6				1								
322	nick cave pictures		1			6												
323	nikky tyler pictures		1			6												
324	nile river pictures		1			6				1								
325	nixon AND photo	1				7												
326	nonporn nude pictures of underage girls				1	7						1				1		
327	nude beach tanning pictures				1	7				1		1				1		
328	nude boys pictures				1	7						1					1	
329	nude britney spears pictures free		1			7						1		1			1	
330	nude cheer leader pictures				1	7						1					1	
331	nude hentai images				1	7						1					1	
332	nude jane pictures				1	7						1					1	
333	nude model pictures				1	7						1				1		
334	nude photo boy penis				1	7						1					1	
335	nude pictures				1	7						1					1	

Figure 3.
Screenshot of evaluator worksheet with category and refiner judgments

the queries in different categories of Unique, Unique with refiner, Non-unique, or Non-unique with refiners. The overall results are displayed in Figure 4, and the refiners for these queries are displayed in Figure 5.

Of the 509 judgments for the Enser and McGregor comparison, one sees from Figure 4 that the highest percentage of requests for web images was non-unique with refiners (71.9 percent), substantially more than reported by Enser and McGregor (25 percent), and naturally, the percentage of Unique queries was lower (15.2 percent). Surprisingly, given the anecdotal comments concerning web users, the use of refiners was quite high with 87.1 percent of all image queries containing at least one refiner. For exact numbers of each category, see Table III.

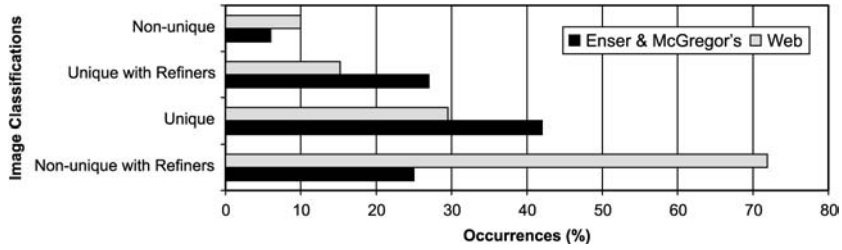


Figure 4.
Mapping using Enser and McGregor's four categories

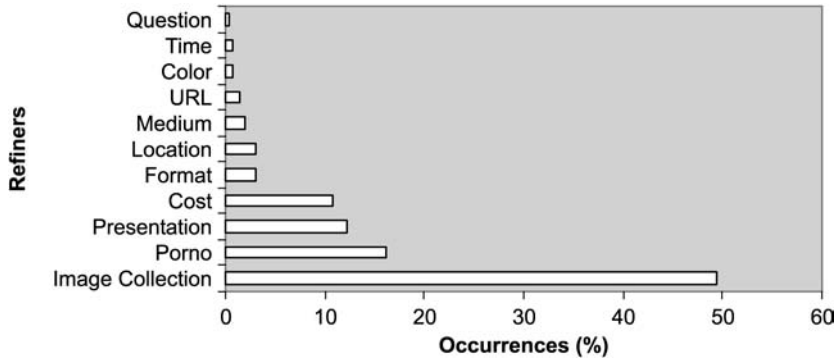


Figure 5.
Mapping using Enser and McGregor's refiners with evaluator modifications

Table III.
Comparing category finding between Enser and McGregor and web query evaluations

	Enser and McGregor	Web
Non-unique with refiners	25	71.9
Unique	42	29.5
Unique with refiners	27	15.2
Non-unique	6	10.0

The evaluators made 2,524 judgments on refiners for the 583 queries. We see from Figure 5 that Collections, Pornography, Presentation of the image, and Cost were all significant aspects of web image queries. For exact percentages, see Table IV.

Table V shows the confidence of the category evaluations, and we present inter-rater agreement for the refiners in Table VI. Refiners with significant agreement are bolded (i.e. objects, people, cost, URLs, color, and collections).

We see from Table V, that most evaluators were fairly confident of their ratings using Enser's four major classifiers. This is perhaps not too surprising given that the classifiers are rather broad. In Table VI, the agreements become much more diverse. In addition to objects, people, cost, URLs, color, and collections, there was also solid agreement with Location and Description. The refiners of People-related, Historical,

Table IV.
Mapping using Enser and McGregor's refiners with evaluator modifications

Refiner	Percent
Image collection	49.4
Porno	16.2
Presentation	12.3
Cost	10.8
Format	3.0
Location	3.0
Medium	2.0
URL	1.4
Color	0.8
Time	0.8
Question	0.4

Visual elements, Content/story, Image relationship, and Viewer response had little to no agreement among the evaluators.

5.2. Mapping results using Jörgensen's image attributes

We also classified the image queries to Jörgensen's image attributes. There were 14 (2 percent) queries not utilized in the Jorgenson comparison due to invalid reviewer comparisons (i.e. no two reviewers assign the query the same attribute). The overall results for the attributes are displayed in Figures 6 and 7, with a comparison to findings reported in (Jörgensen, 1998).

From Figures 6 and 7, we see that the three most frequently occurring attributes are Image collections (31.2 percent), Descriptions (19.0 percent) and People (18.8 percent). The actual percentages are shown in Table VII. The total number of judgments in the study of Jörgensen's image attributes was 1,404. Jörgensen's original 12 attributes did not effectively describe the attributes of the web image queries, so an additional three attributes were added, which were Cost, URL, and Image collections.

Concerning the three attributes added, it is apparent that cost is factor with web users, a fact that would not be apparent in lab studies using students, as Jörgensen and others have done. Given the unique hypertext environment, the location of the images is apparently a concern of web users given the high number of URL requests and searches for images collections. People and people-related queries seem to dominate the information needs of web users. Content-based attributes, such as color, do not appear to be utilized by web users even though web IR systems provide relevant results to these content-based attributes (e.g., using AltaVista, search for blue images).

We present inter-rater agreement for the refiners in Table VIII. Refiners of significant agreement are bolded (i.e. objects, people, cost, URLs, color, and collections).

We see from Table VIII, that image collection had the highest agreement (86 percent), which may be an indication of how prevalent and straightforward of an information need it may be in web image searching. Cost (69 percent) also has a high percentage of agreement. Cost refiners typically had the query term "free", making them easy to identify.

5.2.1. Mapping based on Chen's modifications. We classified the image queries to Chen image categories modifications. The overall results are displayed in Figure 8.

The total number of judgments for the study of Chen's image attributes modifications was 869. Similar to Jörgensen's scheme, we had to add attributes to account for differences in the web query. The results from the web mapping were much more diverse than that reported by Chen, as is apparent in Figure 8. Table IX presents the detailed results. The evaluator agreements are the same as reported in Table VI.

	Unique (%)	Unique with refiners (%)	Non-unique (%)	Non-unique with refiners (%)	Total
Total	15	77	51	365	508
Average confidence	5.7	6.4	5.3	6.0	6.0
St. Dev.	1.1	1.0	1.0	1.1	0.90

Table V.
Confidence in assigning attributes

	Object	People	People-related	Cost	URL	Historical	Color	Visual elements	Location	Description	Abstract concepts	Content/story	Image relationship	Viewer response	Collections	Total
Total	320	373	201	87	22	16	11	83	57	381	363	90	12	9	499	2,524
All	42	162	1	74	8	0	4	1	16	79	1	0	0	0	20	408
Two	147	102	28	8	3	0	4	6	19	188	68	3	0	2	418	996
Total	189	264	29	82	11	0	8	7	35	267	69	3	0	2	438	1,404
(%)	59	71	14	94	50	0	73	8	61	70	19	3	0	22	88	56
One	131	109	172	5	11	16	3	76	22	114	294	87	12	7	61	1,120
(%)	41	29	86	6	50	100	27	92	39	30	81	97	100	78	12	44

Table VI.
Inter-rater agreement for imaged query refiners

6. Discussion of results

In Enser and McGregor's (1992a) study, approximately 70 percent of the queries were for a unique person object or event and 34 percent of these requests were refined, primarily by time. In applying the Enser and McGregor's categories to web image queries, we find that the majority of queries were Non-unique, but almost all (87.1

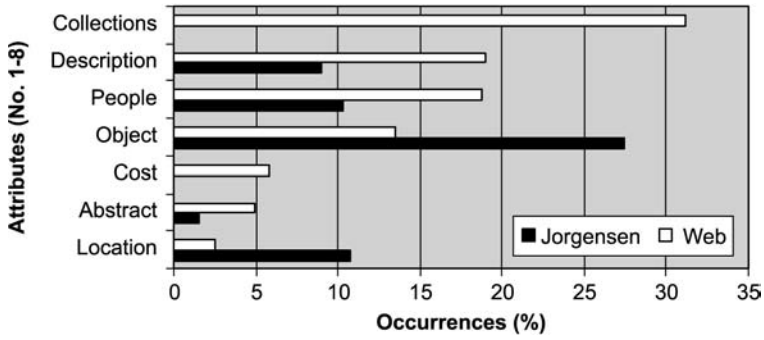


Figure 6.
Mapping using Jörgensen's image attributes with web modifications (1-8)

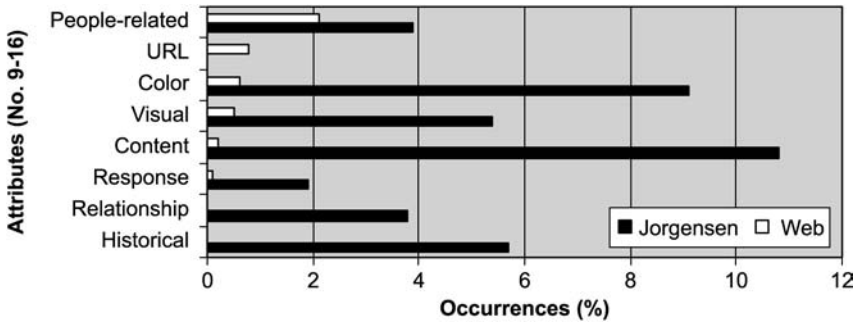


Figure 7.
Mapping using Jörgensen's image attributes with web modifications (9-16)

	Jörgensen	Web
Historical	5.7	0.0
Relationship	3.8	0.0
Response	1.9	0.1
Content	10.8	0.2
Visual	5.4	0.5
Color	9.1	0.6
URL	0	0.8
People-related	3.9	2.1
Location	10.7	2.5
Abstract	1.5	4.9
Cost	0	5.8
Object	27.4	13.5
People	10.3	18.8
Description	9	19.0
Collections	0	31.2

Table VII.
Comparing category findings between Jörgensen and web query evaluations

Table VIII.
Inter-rater agreement for image query attributes

	Time	Location	Format	Color	Medium	URL	Cost	Presentation	Question	Porno	Image collection	Other	Total
Total	41	72	69	41	166	38	108	355	34	189	436	0	1,549
All	4	7	14	0	0	0	68	6	0	50	15	0	164
Two	1	12	7	3	15	2	7	85	0	62	359	0	553
Total	5	19	21	3	15	2	75	91	0	112	374	0	717
(%)	12	26	30	7	9	5	69	26	0	59	86	0	46
One	9	26	18	5	129	3	3	250	1	53	58	0	555
(%)	22	36	26	12	78	8	3	70	3	28	13	0	36

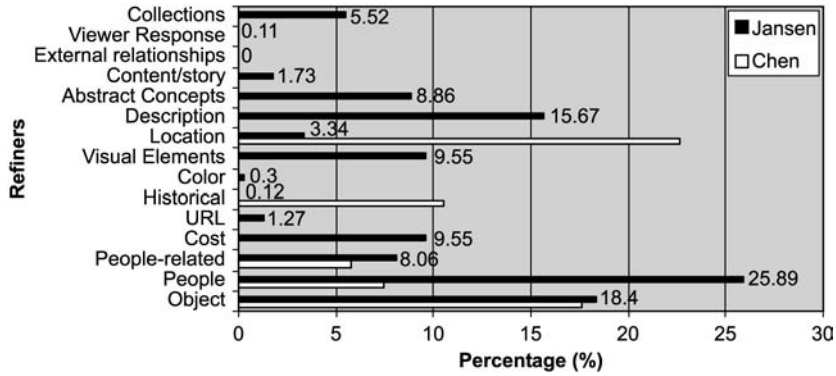


Figure 8.
Mapping using Chen image categories modifications

	Chen	Jansen
Object	17.65	18.4
People	7.46	25.89
People-related	5.8	8.06
Cost		9.55
URL		1.27
Historical	10.52	0.12
Color		0.3
Visual elements		9.55
Location	22.62	3.34
Description		15.67
Abstract concepts		8.86
Content/story		1.73
External relationships		0
Viewer response		0.11
Collections		5.52

Table IX.
Comparing category findings between Chen and web query evaluations

percent) of these were refined in some manner, with Image collection (49.4 percent), Pornographic (16.2 percent), Presentation (12.3 percent) and Cost (10.8 percent), refiners being the most common.

In the Jörgensen (1998) study, the Literal object attribute held the highest position (29.3 percent) in the searching task, followed by the People attribute (10.0 percent), and then the Color attribute (9.3 percent). In the analysis of web queries, the most frequency occurring attribute is Image collection (31.2 percent), Description (19.0 percent), People (18.8 percent), Object (13.5 percent) and then Cost (5.8 percent).

Concerning the Chen (2001) study, Location (23 percent) and Object (18 percent) were the frequently occurring refiners. These did not track with the most frequently occurring refiners in the web image query analysis.

Therefore, we see that none of these previously published image classifications schemas ported whole sale to the web image searching environment. In fact, some did not port well at all, with all methods showing a deficit in the attributes associated with web image requests. There are several possible reasons why these schemes not to

translate well to the web environment. Certainly, the different user communities and the associated types of image themes from each user group has to be a major consideration. For example, a major picture archive would expect to receive a large proportion of requests for highly specific ("unique") subject matter. A group of art history students might be expected to issue requests of a very different kind. The web, with its vast and heterogeneous content, would reasonably elicit different types of queries.

The results of this study point to the need for five additional attributes (i.e. collections, pornography, presentation, URL, and cost), which were not present in any of the three prior classification schemes. Results would indicate that web searchers expressed these attributes in response to the unique searching environment, the non-laboratory setting of the data collection, and the diverse image content collection.

Web searchers expressed the attribute "collection" in queries such as: stock photography, spanish armada pictures, and snow blizzard photos snowdrift snowstorms snowstorm shoveling. Presentation query examples are: tsurphu *and* monastery *and* photo and valentine's image backgrounds clipart. URL was the entering of a URL into the search box, with examples being www.angelfire.com/tn/sampson/warpics and www.b&hphoto.com. Cost generally referred to "free", such as: where can i get free pictures of gay policemen, all free pics, and amateur pictures free. One can imagine how the pornographic queries were structured.

7. Implications

It is apparent that cost is a factor with web users, an observation that would not be apparent in results from lab studies. This highlights the need for "real life" in evaluating users – system interaction. Cost was a refiner in 10.8 percent of the entire sample of queries. Given the unique hypertext environment, the location of the images is apparently also a concern of web users given the high number of URL requests and queries for Image collections. People and people-related queries seem to dominate the information needs of web users. Content-based attributes, such as Color, do not appear to be utilized by web users even though most web IR systems provide relevant results to these content-based attributes, albeit the indexing is done using a context approach.

How users search for images on the web is certain to have an impact on how users interaction with other knowledge bases. End user searching is now the primary method of locating images for many people. By mapping a sub-set of image queries to known image classification schemes, the results indicated that the features and attributes of web image queries differ relative to image queries utilized in other IR systems and user populations. The majority of image queries were requests for Non-unique entities and the majority of queries contained some refiner. Despite these differences, it is apparent that known image classification schemes can be modified to account for the characteristics of most web users.

8. Conclusion

In this paper, we compared three existing image query classification schemes to a set of web image queries. None of the three classification schemes captured the richness of web image searching. We present comparisons and additional classifiers for web image searching. From these results, one can derive a meaningful scheme for classifying web image queries via either manual or automatic meta-tagging. Certainly

one immediate implementation would be meta-tagging of queries to meet these new classifiers and to develop interfaces to support such searching in image collections. Understanding how real users search for digital images will benefit the design and development of effective image retrieval systems. In fact, web IR systems can use the results of this study for the design of image query classification schemes.

References

Barnard, K., Duygulu, P., Freitas, N.D., Forsyth, D.A., Blei, D.M. and Jordan, M.I. (2003), "Matching words and pictures", *Journal of Machine Learning Research*, No. 3, pp. 1107-35.

Chen, H. (2001), "An analysis of image queries of art history", *Journal of the American Society for Information Science and Technology*, Vol. 52 No. 3, pp. 260-73.

Choi, Y. and Rasmussen, E.M. (2003), "Searching for images: the analysis of users' queries for image retrieval in American history", *Journal of the American Society of Information Science and Technology*, Vol. 54 No. 6, pp. 498-511.

Cyber Atlas (2000), US Top 50 Internet Properties, May, At Home/Work Combined, available at: http://cyberatlas.internet.com (accessed 1 July 2001).

Ellis, D. (1984), "Theory and explanation in information retrieval research", *Journal of Information Science*, Vol. 8, pp. 25-38.

Enser, P.G.B. and McGregor, C.G. (1992a), *Analysis of Visual Information Retrieval Queries*, Report on Project G16412 to the British Library Research and Development Department, British Library, London.

Enser, P.G.B. and McGregor, C.G. (1992b), "Analysis of visual information retrieval queries", unpublished manuscript, London.

Fidel, R. (1997), "The image retrieval task: implications for the design and evaluation of image databases", *The New Review of Hypermedia and Multimedia*, Vol. 3 No. 1, pp. 181-99.

Fukumoto, T. (2006), "An analysis of image retrieval behavior for metadata type image database", *Information Processing & Management*, Vol. 42 No. 3, pp. 723-8.

Goodrum, A. and Kim, C. (1998), *Visualizing the History of Chemistry: Queries to the CHF Pictorial Collection*, Report to the Chemical Heritage Foundation Pictorial Collection, Chemical Heritage Foundation, Philadelphia, PA.

Goodrum, A. and Spink, A. (2001), "Image searching on the Excite search engine", *Information Processing & Management*, Vol. 37 No. 2, pp. 295-311.

Greenberg, J. (1993), "Intellectual control of visual archives: a comparison between the *Art and Architecture Thesaurus* and the *Library of Congress Thesaurus for Graphic and Materials*", *Cataloging and Classifications Quarterly*, Vol. 16 No. 1, pp. 85-101.

Greisdorf, H. and O'Connor, B. (2002), "Modeling what users see when they look at images: a cognitive viewpoint", *Journal of Documentation*, Vol. 58 No. 1, pp. 6-29.

Gudivada, V.V. and Raghavan, V.V. (1995), "Content-based image retrieval systems", *IEEE Computer*, Vol. 28 No. 9, pp. 18-22.

Hertzum, M. (2003), "Requests for information from a film archive: a case study of multimedia retrieval", *Journal of Documentation*, Vol. 59 No. 2, pp. 168-86.

Jansen, B.J. and Pooch, U. (2001), "Web user studies: a review and framework for future work", *Journal of the American Society of Information Science and Technology*, Vol. 52 No. 3, pp. 235-46.

Jansen, B.J., Goodrum, A. and Spink, A. (2000), "Searching for multimedia: video, audio, and image web queries", *World Wide Web Journal*, Vol. 3 No. 4, pp. 249-54.

Jansen, B.J., Spink, A. and Koshman, S. (2007), "Web searcher interaction with the Dogpile.com meta-search engine", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 5, pp. 744-55.

Jansen, B.J., Spink, A. and Pederson, J. (2003), "An analysis of multimedia searching on AltaVista", *5th ACM SIG Multimedia International Workshop on Multimedia Information Retrieval, Berkeley, CA*, pp. 186-92.

Jansen, B.J., Spink, A. and Pedersen, J. (2005), "The effect of specialized multimedia collections on web searching", *Journal of Web Engineering*, Vol. 3 Nos 3/4, pp. 182-99.

Jörgensen, C. (1998), "Attributes of images in describing tasks", *Information Processing and Management*, Vol. 34 Nos 2/3, pp. 161-74.

Jörgensen, C. (2003), *Image Retrieval: Theory and Research*, The Scarecrow Press, Lanham, MD.

Jörgensen, C. and Jörgensen, P. (2005), "Image querying by image professionals", *Journal of the American Society for Information Science and Technology*, Vol. 56 No. 12, pp. 1346-59.

Large, A., Beheshti, J. and Breuleux, A. (1998), "Information seeking in a multimedia environment by primary school students", *Library & Information Science Research*, Vol. 20 No. 4, pp. 343-76.

Lawrence, S. and Giles, C.L. (1999), "Accessibility of information on the web", *Nature*, Vol. 400, pp. 107-9.

Lesk, M. (1997), "Going digital", *Scientific American*, Vol. 276, pp. 58-60.

Li, J. and Wang, J.Z. (2003), "Automatic linguistic indexing of pictures by a statistical modeling approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25 No. 9, pp. 1075-88.

Lim, J.-H. and Jin, J.S. (2002), "Image indexing and retrieval using visual keyword histograms", *IEEE International Conference on Multimedia and Expo, Lausanne. Switzerland, 26-29 August*, pp. 213-16.

Marsh, E.E. and White, M.D. (2003), "A taxonomy of relationships between images and text", *Journal of Documentation*, Vol. 59 No. 6, pp. 647-72.

Özmultu, C., Spink, A. and Özmultu, Z. (2003), "Multimedia web Searching Trends: 1997-2001", *Information Processing & Management*, Vol. 39 No. 4, pp. 611-21.

Patton, M.Q. (1990), *Qualitative Evaluation and Research Methods*, 2nd ed., Sage Publications, Newbury Park, CA.

Rasmussen, E.M. (1997), "Indexing images", *Annual Review of Information Science and Technology*, Vol. 32, pp. 169-96.

Schauble, P. (1997), *Multimedia Information Retrieval*, Kluwer, Boston, MA.

Smeaton, A.F. (2004), "Indexing, browsing, and searching of digital video", in Cronin, B. (Ed.), *Annual Review of Information Sciences and Technology*, Vol. 38, pp. 371-407.

Spink, A. and Jansen, B.J. (2004), *Web Search: Public Searching of the Web*, Kluwer, New York, NY.

Swain, M.J. (1999), "Searching for multimedia on the world wide web" *IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, 7-11 June*, Vol. 1, pp. 9032-37.

Turner, J. (1998), "Words and pictures in information systems for moving images" *1998 Association of Moving Image Archivists Conference, Miami, Florida, USA, 7-12 December*, pp. 309-15.

Wang, J.Z. (2000), "SIMPLIcity: a region-based image retrieval system for picture libraries and biomedical image databases" *ACM Multimedia, Los Angeles, CA, October 2000*, pp. 483-4.

Witten, I., Moffat, A. and Bell, T.C. (1994), *Managing Gigabytes: Compressing and Indexing Documents and Images*, Van Nostrand Reinhold, New York, NY.

Wu, Y. and Zhang, A. (2002), "Category-based search using metadatabase in image retrieval" *IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland, 26-29 August*, pp. 197-200.

Corresponding author
Bernard J. Jansen can be contacted at: jjansen@acm.org