COVER SHEET

Information Processing and Management (2001), *37*(2), 295-312.

IMAGE SEARCHING ON THE EXCITE WEB SEARCH ENGINE

Abby Goodrum

College of Information Science and Technology, Drexel University

3141 Chestnut Street

Philadelphia, PA 19104-2875

USA

Amanda Spink

School of Information Sciences and Technology

The Pennsylvania State University

State College, PA 16801

USA

Abstract

A growing body of research is beginning to explore the information-seeking behavior of Web users. The vast majority of these studies have concentrated on the area of textual information retrieval (IR). Little research has examined how people search for non-textual information on the Internet, and few large-scale studies has investigated visual information-seeking behavior with general-purpose Web search engines. This study examined visual information needs as expressed in users' Web image queries. The data set examined consisted of 1,025,908 sequential queries from 211,058 users of Excite, a major Internet search service. Twenty-eight terms were used to identify queries for both still and moving images, resulting in a subset of 33,149 image queries by 9855 users. We provide data on: (1) *image queries* – the number of queries and the number of search terms per user, (2) *image search sessions* – the number of queries per user, modifications made to subsequent queries in a session, and (3) *image terms* – their rank/frequency distribution and the most highly used search terms. 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% of the time, with most terms occurring only once. We contrast this to earlier work by P.G.B. Enser, Journal of Documentation 51 (2) (1995) 126–170, who examined written queries for pictorial information in a non-digital environment. Implications for the development of models for visual information retrieval, and for the design of Web search engines are discussed.

1. Introduction

The Web is a complex and unique source of multimedia information. Digitized text, audio, images and video link to each other on the Web of information bearing objects. In this paper, we focus on users' searching for images. It is estimated that there are 180 million images on the publicly indexable Web and a total amount of image data of about 3 Tb (Lawrence & Giles, 1999). Over one million digital images are being added to the Web every day (Jain, 1993). The need to find an image from within a specific digital collection on the Web is important to many user groups, including journalists, engineers, historians, designers, teachers, artists, and advertising agencies. Image needs and uses across these groups of users may vary considerably however. Users may need access to images based on such primitive features such as color, texture or shape, or access to images based on abstract concepts and symbolic imagery (Fidel, 1997). The technology to access images is also changing rapidly and currently surpasses our understanding of how users interact with visual information.

As more and more people and organizations load images onto the Web, the searching and retrieval of images has become a major challenge for researchers and users alike. How users search for multimedia, and specifically image content materials on the Web, and the design of more effective Web image retrieval systems, is a growing area of research. Few previous studies of information seeking on the Web have specifically examined textual queries by users seeking image information.

2. Related research

2.1. The current state of text-based image retrieval

Most existing information retrieval (IR) systems are text-based, but the problem remains that images frequently have little or no accompanying textual information. Historically, the solution has been to develop text-based ontologies and classification schemes for image description. Text-based indexing has much strength, including the ability to represent both specific and general instantiations of an object at varying levels of complexity. Textual metadata can provide information not available from the image, e.g., names of people shown in the picture, or geographical location of the shot. Text-based image retrieval literature is reviewed by Rasmussen; Lancaster, 1998; Lunin and Cawkell, 1993.

Long before images could be digitized, access to image collections was provided by librarians, curators, and archivists via text descriptors or classification codes. These indexing schemes were often developed in-house and reflect the unique characteristics of a particular collection or clientele. This is still common practice. Recent advances include research by Zheng (1999), and Goodrum and Martin (1999), who reported the hybridization of multiple schemes for classifying collections of historic costume collections. Hourihane (1989) and Jorgensen (1995) also review a number of unique systems for image classification. However, little research has examined the relative effectiveness of these various approaches to image indexing in electronic environments.

Attempts to provide general systems for image indexing include Getty's art and architecture thesaurus (http://shiva.pub.getty.edu/aat_browser/) that consists of over 120,000 terms for the description of art, art history, architecture, and other cultural objects, and the library of congress thesaurus of graphic materials (http://lcweb.loc.gov/rr/print/tgm1/). The art and architecture thesaurus provides access to 33 hierarchical categories of image description using seven broad facets (associated concepts, physical attributes, styles and periods, agents, activities, materials, and objects). The approach in many collections, particularly general library environments, has been to apply an existing classification system like the Dewey decimal classification system to image description using the library of congress thesaurus of graphic materials.

However, assignment of terms to describe images is not solved entirely by the use of controlled vocabularies or classification schemes. The textual representation of images is problematic because images convey information relating to what is actually depicted in the image as well as what the image is about. Shatford (1986) discusses this issue within a framework based on Panofsky's (1967) approach to analyzing levels of meaning in images. For example, an image may be *of* a cathedral, but be *about* the power of the church. Shatford-Layne (1994) extended this discussion by proposing a theoretical model for analyzing the subject of an image and suggests that it may be necessary to determine which attributes will result in useful groupings of images and which attributes should be left to the user to identify. Turner and Turner also conducted research determining the subject content of both still and moving images within this framework with the aim of discovering appropriate ways to index moving images.

2.2. Metadata

An area of image retrieval research for which very little has been published is retrieval by associated metadata. Metadata includes attributes such as image creator, image format, date of creation, and simple object descriptions taken from titles or captions. The Dublin core metadata set, has been adopted for the description of Web documents. Similarly, the World Wide Web Consortium (W3C) is developing a resource description framework (RDF) to provide a shell for expressing a very large number of metadata types such as textual descriptions of the resource and color histogram or other numeric representations of image content. However, a recent study by Lawrence and Giles (1999) shows that the use of Web metadata tags is still not widespread.

Unfortunately, human assignment of textual attributes is both time-consuming and costly. Human indexing also suffers from low term agreement across indexers (Markey, 1984), and between indexers and user queries (Enser and McGregor, 1993 and Seloff). Automatic assignment of textual attributes has been conducted using captions from still images, transcripts, closed captioning, and verbal description for the blind accompanying videos (Turner, 1999). While these approaches greatly reduce the labor involved in manual assignment of keywords, many images are without accompanying text. Furthermore, users' image needs may occur at a primitive level that taps directly into the visual attributes of an image. These attributes may best be represented by image exemplars and retrieved by systems performing pattern matches based on color, texture, shape, and other visual features.

2.3. Content-based image retrieval

Problems with text-based access to images have prompted increasing interest in the development of image-based solutions. This is most often referred to as content-based image retrieval (CBIR). CBIR relies on the characterization of primitive features such as color, shape, and texture that are automatically extracted from the images themselves. Commercial CBIR systems in use include IBM’s query by image content (QBIC) described first by Flickner et al. (1995), Virage's VIR image engine (Bach et al., 1996), and Excalibur's image retrieval ware. On the Web, CBIR image retrieval systems include WebSEEK (Smith & Chang, 1997), Informedia, and Photobook among others. The search services Altavista and Yahoo! also provide image search using CBIR.

Queries to CBIR systems are most often expressed as visual exemplars of the type of image or image attribute being sought. For example, users may submit a sketch, click on a texture palette, or select a shape. The system then identifies those stored images with a high degree of similarity to the requested feature. Idris and Panchanathan (1997) discuss in detail the various technologies for image indexing and retrieval based on shape, color, texture, and spatial location. They also examine issues related to the retrieval of moving images, including shot detection and video segmentation. Aigrain, Zhang and Petkovic (1996) provide an overview of approaches to image similarity matching for database retrieval and discuss the difficulty of expressing high-level image needs to low-level image features.

2.3.1. Color

Retrieving images based on color similarity is achieved by computing a color histogram for each image that identifies the proportion of pixels within an image holding specific values (that humans express as colors). Current research is attempting to segment color proportion by region and by spatial relationship among several color regions (Stricker and Orengo, 1995 and Carson et al., 1997).

2.3.2. Texture

Texture is a difficult concept to represent. The identification of specific textures in an image is achieved primarily by modeling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality may be estimated (Tamura, Mori & Yamawaki, 1978). However, the problem is in identifying patterns of co-pixel variation and associating them with particular classes of textures such as “silky”, or “rough”. Ma and Manjanath (1998) extended work in this area through the development of a texture thesaurus that matches texture regions in images to words representing texture attributes.

2.3.3. Shape

Queries for shapes are generally achieved by selecting an example image provided by the system or by having the user sketch a shape. The primary mechanisms used for shape

retrieval include identification of features such as lines, boundaries, aspect ratio, and circularity, and by identifying areas of change or stability via region growing and edge detection. Of particular concern is the problem of dealing with images having overlapping or touching shapes.

2.4. User interaction with images

Users seeking images come from variety of domains, including law enforcement, journalism, education, entertainment, medicine, architecture, engineering, publishing, advertising, and art. Most of the published research in this area focuses on specific collections, or specific groups of users. For example, Goodrum (1998) explored users' perceptions of surrogates for environmental videos, Ornager (1997) examined the use of newspaper image archives, Keister (1994) analyzed queries submitted to the image archive at the national library of medicine, and Markey (1988) and Hastings (1995) explored the use of images by art historians.

Much of the research in visual information seeking behavior and use has been conducted in non-digitized collections with written or verbal queries. The seminal work by Enser (1995) analyzed nearly 3000 written requests from 1000 request forms at the Hulton Deutsch archive. Queries for visual materials exhibited a greater level of specificity than requests for textual materials, and the majority of requests were for specific instances of a generic category ("London Bridge" rather than the generic "Bridges"). Armitage and Enser (1997) extended this research by categorizing requests across seven picture archives that resulted in a framework for queries with four main categories (who, what, when, where) and three levels of abstraction (specific, generic, abstract). Similarly, Keister's (1994) analysis of query logs at the National library of medicine showed that most queries were structured using both abstract concepts as well as concrete image elements. She concluded that the aesthetic and emotional needs of the user are highly subjective and do not lend themselves to indexing. In contrast, Jorgensen's (1999) research has focused on how to provide access to the emotive content of images.

However, research examining users' interactions with electronic image retrieval systems is still quite sparse. Most of the research has examined written queries for pictorial information in a non-digital environment. We now turn our attention to the Web image retrieval research that forms the focus of the remainder of this paper.

2.5. Web image retrieval

Most of the major search sites have some provision for searching for images. For example, at the time of writing this paper, Altavista (www.altavista.com) lets users preface a query with the image operator *image:* and Hotbot (www.hotbot.com) provides check boxes to limit a search to certain media types. In general, search engines that provide image search tools find images by searching for image file extensions and matching the image's file name to terms in the query. Some search engines may also retrieve web pages containing an image file name not matching the query terms, but containing query terms in the title, the URL, or elsewhere in the page. These pages will generally be ranked lower in the retrieval list.

The disadvantage of this approach is that it places more of the contextual knowledge burden on the searchers, who may not be familiar with image formats and file extensions. Further, cognitive processing is required on the part of users to translate the visual need into a textual query. This creates what some authors refer to as a lack of representational congruity (Goodrum, in review), or as a *semantic gap* (Gudivada & Raghavan, 1995). This problem is exacerbated by the presentation of retrieved items as text-only lists rather than as thumbnail images, or video key frames.

2.5.1. Web image search services and directories

There are also specialized image search engines and image directories within some of the major search services. Of these, one is worth special note. Webseek allows users to search by term or select from general categories of images (both still and moving). Unlike standard text-based search engines, Webseek returns thumbnail images for users to view. The system also provides tools for content-based searching for images and videos using color histograms generated from the visual scenes.

When the data for this project was collected, Excite did not provide an explicit mechanism for locating images. Users looking for images specified this by creating their own image request terms to the query. For example, a user searching for images of bears might input "*bear gifs*", or "*photos of bears*". Image request terms are words that may indicate image files or content. For example, the file extensions: *.jpg*, *.tif*, *.avi* or the terms: *images*, *photographs*, *clips*, or *pics*. This approach is problematic for users not familiar with image file extensions. It can also be troublesome for users who do not know if their image query terms are used anywhere in the documents they hope to find.

2.5.2. Web image queries

Textual terms are the basic building blocks of IR queries. Queries are the primary means by which users' information needs are expressed to an IR system. Terms and queries are important variables in the exploration of information-seeking behavior in electronic environments. Query analysis has also formed the basis for the examination of VIR. Most studies of this nature focus on image seeking and use in indexed image collections (Enser; Goodrum and Kim, 1998; Hastings, 1995; Turner and OConnor). These studies indicate that the number of terms by a single user searching for an image may be low, but the pool of terms employed across all users searching for images is quite large.

This contrasts to the mean number of terms used to find textual documents in structured databases (Spink & Saracevic, 1997.) These textual studies identified a mean of 7–15 terms per query. Jansen, Spink and Saracevic (Jansen et al., 1998; Jansen et al., 1999 and Jansen) found that IR queries on the Web averaged only 2.35 search terms. As discussed in this paper, this closely approximates the mean number of search terms employed by users searching for images on the Web. Limited studies have examined users' behavior when seeking images from the Web. This is an important area of research for the development of models of image information seeking and more effective Web-based image retrieval tools.

This paper reports selected results from a major and ongoing study of searching behavior by Excite users (Jansen et al., 1998; Jansen et al., 1999 and Jansen). We examined a subset of transaction logs of searches conducted by users on Excite, a major Web search service. The objectives of the study include the analysis of (1) image queries, (2) image search sessions, and (3) image terms and term distribution.

3. Research design

3.1. Background on Excite data

Founded in 1994, Excite is a major Internet media public company that offers free Web searching and a variety of other services. The company and its services are described in more detail at its Web site (http://www.excite.com). Excite searches are based on the exact terms that a user enters in the query, however, capitalization is disregarded, with the exception of logical commands AND, OR, and AND NOT with no stemming. An online thesaurus and concept linking method called intelligent concept extraction (ICE) is used to find related terms in addition to terms entered. Search results are ranked by relevance.

3.2. Data collection

The data set consisted of 1,025,908 user queries. Each transaction record contained three fields that allowed us to recreate the chronological series of actions during each user session:

• *Time of day*: measured in hours, minutes, and seconds from midnight of 9 March 1997.

• *User identification*: an anonymous user code assigned by the Excite server.

• *Query terms*: exactly as entered by the given user.

We define sessions, queries, and terms in the following way:

• *Session*: the entire series of queries by a user over a number of minutes or hours. A session could be as short as one query or contain many queries.

• *Query*: one or more search terms, and possibly including logical operators and modifiers.

• *Term*: any unbroken string of characters (i.e., a series of characters containing no spaces). The characters in terms include letters, numbers, and symbols. Terms can be words, abbreviations, numbers, symbols, URLs, or any combination thereof. We counted logical operators in capitals as terms. However, in a separate analysis we isolated them as commands, not terms.

In large part, the data was taken as is, i.e., we did not 'clean' the data in any way – these queries represent real searches by real users. The only normalization we undertook in one of the counts (unique terms – not case sensitive) was to disregard capitalization, because Excite disregards it as well (i.e., TOPIC, topic and Topic retrieve the same answers). Excite does not offer automatic stemming, thus *topic* and *topics* count as two unique terms, and '?' or '*' as stemming commands at the end of terms are not recognized by Excite and are counted as separate terms.

3.3. Image query parsing

Parsing the original 1,025,908 query data set for 28 terms, identified image queries. These terms represent a combination of common image file extensions and ordinary image terms. The selected terms included both singular and plural forms, as Excite does not automatically

truncate. No claim is made, however, that these terms comprise a comprehensive set of all possible image terms used. These terms are presented in alphabetical order in Table 1.

Term	Term
Pictures	Jpeg
Pics	Films
Photos	Gifs
Video	Jpg
Movies	Movie
Picture	Mpeg
Photo	Clips
Images	Photography
Pic	Avi
Videos	Mov
Film	Mjpeg
Image	Pcx
Photographs	Tif
Gif	Png

Table 1. Image request terms

Analysis of queries, sessions, and terms was conducted on the subset of queries containing at least one of these terms. Section 4 of the paper discusses the results of this analysis.

4. Results

Some general statistics about the data corpus are presented in Table 2.

Table 2. Number of queries, users and terms for the original and image data corpus

	Original data corpus	Image data corpus	% of original data
Queries	1,025,908	33,149	4
Users	211,058	9,855	3
Terms	2,146,154	124,058	12

Although image queries constitute a relatively small percentage of queries overall, it must be noted that nearly 10,000 users with over 33,000 requests for image information represents a considerable volume of activity. It should also be noted that the lack of explicit mechanisms for image searching by Excite at the time of the study might account for a lower percentage of image queries compared to other search engines.

4.1. Image terms

All queries had to contain at least one image request term in order to be included in this study. Some users, however, input more than one image request term. Table 3 shows the number of image terms used and their frequency in the data set.

Table 3. Image request terms and their frequency

Term	Frequency	Term	Frequency
Pictures	10,749	Jpeg	340
Pics	6,407	Films	330
Photos	3,769	Gifs	283
Video	2,351	Jpg	266
Movies	1,792	Movie	225
Picture	1,720	Mpeg	148
Photo	1,291	Clips	147
Images	1,176	Photography	130
Pic	1,111	Avi	128
Videos	1,015	Mov	54
Film	630	Mjpeg	5
Image	550	Pcx	1
Photographs	479	Tif	1
Gif	459	Png	1

There was 35,558 occurrences of image terms in the data. The most frequently occurring image term appears in less than 10% of all queries. Common terms for images such as *pictures* and *movies*, occur with greater frequency than file extensions.

4.2. Types of queries

Image filenames are often ambiguous and it can be difficult to represent image content semantically for a heterogeneous user population. Little is known about image query modification and iterative query behavior. We therefore examined sequential queries to learn if users modified their terms after an initial query. In Table 4 unique queries represent the first-image query made by a user in a search. Modified queries are the subsequent queries by the same user that were modified by the addition and/or subtraction of terms.

Type	Total	%
Unique queries	13,385	40.4
Modified queries	19,764	59.6

Table 4. Types of image queries

Since this data set was parsed out of the larger data set by the appearance of image request terms, subsequent queries by the same user not containing image terms do not appear. In contrast to the Web searching studies by Jansen et al., 1998; Jansen et al., 1999 and Jansen, modification of image queries is high. Since we do not have users' relevance data, we cannot tell if query modification is a form of query reformulation, or reflects separate image information needs expressed sequentially by the same user. The high degree of query modification may represent a lack of retrieval precision for image queries, but it is impossible to do this without feedback from users.

4.3. Image sessions

Without access to users or their relevance judgments it was also difficult to determine if multiple interactions represented unsuccessful searches (Table 5).

Statistic	Number
Mean (queries per user)	3.36
Mode (queries per user discounting same queries)	2

Table 5. Mean statistics for image sessions

As can be seen in Table 5, users input an average of approximately three queries during an image searching session. This is a slightly greater number of interactions on average than was identified in previous studies by Jansen et al. Multiple search interactions have been well documented in text searching, but have not been documented extensively in image search behavior. Although users conducted multiple interactions with Excite, we do not know if each interaction was related to the same image information need or represented new topics for the same user. More importantly, it is not possible to know with certainty whether serial queries from the same machine represented queries made by a single person.

4.4. Terms per query

Terms per query were 3.74 (Table 6). When compared to queries for textual information from structured databases such as Dialog, image queries on the Web display very few terms. This appears to be a fairly strong trend in Web searching behavior overall and not a reflection of

the visual nature of the information need. Research conducted by Jansen et al., 1998 and Jansen identified a mean of two terms per query. Although this is lower than the 3.74 found in this study, we must consider that at least one term in each query was used as an image request term and that some queries contained more than one of these terms.

Statistic	Number
Mean (terms/query)	3.74
Mode (terms/query)	3

Table 6. Statistics on terms per queries

There were 124,058 terms that occurred in the image query data set. This count does not distinguish between upper and lower case characters, or terms used as Boolean operators. Over half of the terms were used only once, and the most frequently occurring term is used less than 10% of the time. If we subtract the image request terms *pictures* and *pics*, the top term occurs less than five percent of the time. This corresponds somewhat to earlier image retrieval research that demonstrated that users lack a level of term agreement in describing images (Turner; Goodrum and Kim, 1998 and OConnor). Although the variety and number of terms used to describe images in these studies is often quite high, the top terms in these experiments tend to display a high level of agreement, sometimes as much as 100%. Such was not the case in this study, but it must be noted that we are examining terms used in a large-scale retrieval environment.

4.5. Term frequency

From the complete rank-frequency table, the most frequently occurring 100 terms have been culled and are presented in Table 7. Asterisks appear where expletives have been deleted.

Table 7. Term frequency

Term	Number	%	Term	Number	%	Term	Number	%
Pictures	10,749	8.66	Jpeg	340	0,27	In	174	0,14
Pics	6,407	5.24	Men	336	0,27	Collection	173	0,14
Free	5,642	4.55	Films	330	0,27	and	166	0,13
Photos	3,769	3.04	Celebrities	329	0,27	Anal	164	0,13
Nude	3,145	2.54	x-rated	319	0,26	Death	164	0,13
Sex	2,751	2.22	Male	300	0,24	Cum	163	0,13
Video	2,351	1.90	Older	300	0,24	Big	160	0,13
Of	2,256	1.82	Young	291	0,23	With	160	0,13
Movies	1,792	1.44	Hot	290	0,23	Index	155	0,12
Picture	1,720	1.39	Crash	296	0,23	Art	150	0,12
Adult	1,402	1.13	Gallery	285	0,23	Download	149	0,12
Photo	1,291	1.04	Gifs	283	0,23	Mpeg	148	0,12
Xxx	1,286	1.04	Black	283	0,23	Clips	147	0,12
Women	1,192	0.96	Hardcore	268	0,22	Parties	144	0,12
Images	1,176	0.95	F**k	266	0,21	Star	139	0,11
Pic	1,111	0.90	Jpg	266	0,21	Boys	139	0,11
Videos	1,015	0.82	Nudity	257	0,21	Animal	137	0,11
And	892	0.72	Bondage	236	0,19	On	134	0,11
AND	883	0.71	Celebrity	235	0,19	Sexual	131	0,11
Gay	738	0.59	Teens	231	0,19	Photography	130	0,10
Nudes	666	0.54	Mature	230	0,19	Avi	128	0,10
Teen	662	0.53	Lesbian	227	0,18	Archive	127	0,10
Porn	631	0.51	Movie	225	0,18	Fat	125	0,10
Film	630	0.51	Amateur	214	0,17	For	127	0,10
Diana	614	0.49	Erotic	213	0,17	Stars	126	0,10
Girls	601	0.48	Sexy	197	0,16	Accident	123	0,10
Image	550	0.44	F***ing	196	0,16	Tits	121	0,10
Photographs	479	0.39	Female	195	0,16	Models	121	0,10
Gif	459	0.37	Girl	194	0,16	Nasty	121	0,10
Princess	422	0.34	Entertainment	190	0,15	Celebs	120	0,10
The	357	0.29	Animated	190	0,15	Poses	117	0,09
Porno	349	0.28	Beautiful	183	0,15	Cards	113	0,09
P***y	344	0.28	Asian	180	0,15	Rape	112	0,09
						Transvestite	112	0,09

These 100 terms have a combined frequency of 71,209 appearances as search terms in the set of image queries, and account for 57% of image query terms. Of these 100 terms, nine may be eliminated as not carrying content directly (and, AND, of, the, in, for, on, to, or, and a). The image request terms that appear in this list may also be eliminated, leaving us with 71 content bearing terms. It is somewhat amusing to note that the term occurring with the greatest frequency after these removals is the term: *Free*. In an attempt to understand what images people are looking for on the Web, we sought a rough categorization of the top 100 terms.

Although 25 of these terms are clearly dealing with sexual content, many of the terms are ambiguous when separated from the context of their queries. For example, what is being sought with the terms *women*, *men*, *or teens*? It cannot be assumed that the images sought

are necessarily sexually explicit. Nor can one make gross assumptions concerning how images will be used based on queries alone. The remaining terms ranged from the innocent *cards* to the possibly less virtuous term *hot*. The classification of these terms into broad subject categories is necessarily arbitrary, but a rough picture of the largest categories of image searching emerged (Table 8).

Table 8. Classes of frequently occurring terms

Category	Terms with highest frequency *excluding of, and, the, for, &*	Frequency for category	% of all terms
Image terms	Pictures, pics, video, movies, picture, photo, images, pic, videos, film, image, gif, jpeg, films, gifs, jpg, movie, mpeg, clips, photography, avi	31,248	25.18
Modifiers	Nude, adult, nudes, teen, older, young, hot, crash, black, nudity, mature, amateur, animated, beautiful, Asian, big, with, on, fat, poses, teens	9,516	7.67
Sexual	Sex, xxx, gay, porn, porno, p***y, x-rated, hardcore, f*ck, lesbian, erotic, bondage, sexy, f*cking, anal, cum, sexual, tits, nasty, rape, transvestite	8,945	7.21
Cost	Free	5,642	4.54
Gender	Women, girls, men, male, female, girl, boys	2,957	2.38
Other	Crash, collection, death, index, download, panties, star, animal, archive, stars, accident, cards	1,846	1.48
People	Diana, princess, celebrities, celebrity, models, celebs	1,841	1.48
Arts and leisure	Gallery, art, entertainment	625	0.5

Words used to define the query as an image information need is the most frequently occurring class of terms. Modifiers, qualifiers, or terms representing specific attributes such as age, time, location, action, or color may represent what Enser (1995) calls “refiners”. Terms such as *young*, *fat*, *blonde*, and *beautiful* may serve to refine a general term such as *girl*, into a more specific visual request. Research examining image queries (Turner; Hastings, 1995; Goodrum and Kim, 1998 and OConnor) have consistently demonstrated the

occurrence of such terms in image queries and image descriptions in non-web based image systems.

While it is not feasible to list all unique terms used, Fig. 1 presents a summary of the distribution of these terms in a log rank-frequency graph.

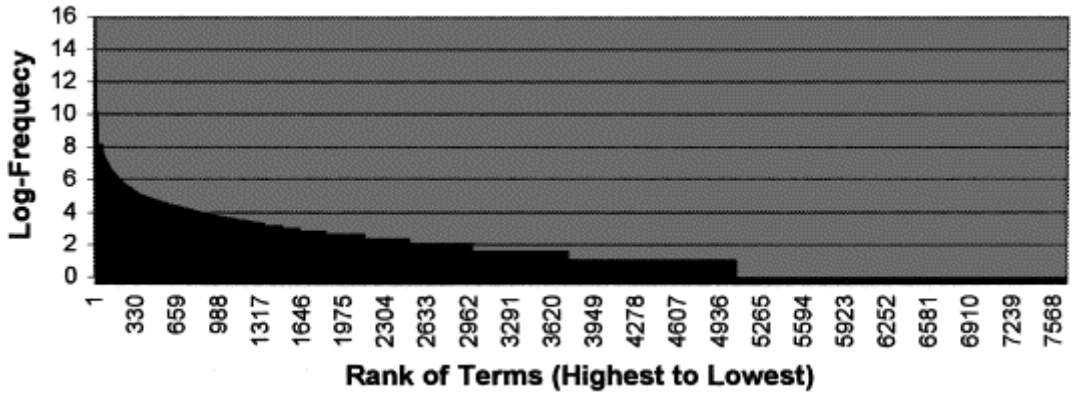

Fig. 1. Frequency distribution of log of all terms in queries.

The *x*-axis indicates the rank of terms and the *y*-axis indicates the logarithm (base 10) of the frequency of their occurrence. As can be seen, the distribution of highly occurring terms falls off swiftly with a long tail representing a large number of terms having a frequency of one.

5. Discussion

Our study findings reveal many interesting insights into image searching by Web users. First, the number of people using Excite to search for images is a clear indication that provisions for image searching need to be made by Web search services. Hotbot and Altavista already provide explicit mechanisms for image searching. It should be noted that in the short time since this project began, Excite has changed its service to provide tools for still and moving image searches as well.

Users input relatively few terms to specify their image information needs on the Web. Whether this is due to lack of experience, time constraints, or an inability to translate image information needs into textual queries is a topic for further study. Users are interacting iteratively during the course of a single session, but input relatively few queries overall. Whether this is due to a lack of support for query reformulation or is a result of low precision in image retrieval is as yet unknown. Most terms are used infrequently, with the top term occurring in less than 9% of queries. There was a high degree of variability across terms,

with over half of the terms used only once. Terms indicating sexual content materials appear frequently. They represented a quarter of the 100 most frequently occurring terms, but were a small percentage of the total number of terms overall.

If we assume that modified queries represent query reformulation, then what can we make of the large number of modifications made by users? Is this indicative of some level of mismatch between a query language that is essentially textual and an information need that is essentially visual? Would the provision of thumbnail images as part of the retrieved set list make browsing easier and lessen query modification? Without data on user satisfaction or relevance judgments, it is difficult to tell.

Our analysis of image queries on Excite suggests that image searching on the web is somewhat different from textual information seeking behavior on the web. Although some patterns of visual information seeking behavior in structured databases are also found in image seeking behavior on the web, further research is needed to model visual information seeking behavior in electronic environments.

General search engines cannot distinguish between an image of a flower and an image of a truck. They must rely on surrounding text and file names for the retrieval of non-textual information-bearing objects. The use of textual systems for the retrieval of images results in an increase in the contextual load placed on the user, as is evidenced by the number of terms and the number of queries needed to retrieve multimedia objects on the Web. Although it may not be possible at this time to provide users with non-textual mechanisms for querying a search service's database, it is not difficult to provide tools to assist users in specifying an image information need and retrieving information with media file extensions. What is more challenging is the provision of image surrogates in the retrieved item list. Appropriate provision of extracted thumbnails and video key frames to support relevance judgments and query reformulation is an area of urgently needed future research. While search engines such as Excite are accepted tools for textual IR, it is reasonable to assume that specialized mechanisms for image searching will become accepted as well (Smeulders and Jain, 1997).

6. Conclusion and further research

A number of important but unresolved issues continue to hamper progress in image retrieval. Among them are fundamental questions surrounding how users represent non-textual

information needs, how they search for images, and how they interact with search engines to obtain non-textual materials. If we are to succeed in delivering maximum benefit from image retrieval efforts, we must combine theoretical and practical expertise across the broader information systems and collection management communities.

Although primitives such as shape, color, and texture are undoubtedly important features for image representation, there is little understanding of how best to implement these attributes for actual image retrieval. The focus to date has been primarily on the use of features that can be computationally acquired, but more research must be done to identify the visual attributes needed by users for various tasks and collections. Problems arise when either documents or information needs cannot be expressed in a manner that will provide congruence between the representation and its referent. In the case of multimedia searching, there are problems in representing image information needs with textual queries, and with representing retrieved images as short textual abstracts. Conversely, there exists a considerable gap between the primitive image features such as color, texture, lines, edges, and angles, and the higher-level cognition necessary to equate these features with terms that occur to human beings in the course of a search.

Our analysis of these Web image queries indicates that users engaged in image searching may be challenged by this lack of representational congruity. Our findings also highlight several key aspects of image searching on the Web. First, the number of users searching for images suggest a need to provide Web mechanisms to facilitate this searching and possibly for viewing results. Second, image IR sessions and queries are still short compared to traditional IR system searching, but longer relative to general Web sessions and queries. This may suggest either a problem with the Web IR system or that the precision of the Web IR system has satisfied the searcher's information need. Third, there appear to be a small number of image terms that occur frequently and a large number of terms that occur very infrequently. Web IR systems should capitalize on the frequently occurring terms and offer thesaurus-type assistance for infrequently occurring query terms.

This paper has provided results from a large study of image searching by Excite users. Our data analysis will continue and further results will be reported in the future. We are currently analyzing a data set of 1.7 million queries from excite users, focusing on image query sessions. This data will be compared with the results reported in this paper.

Acknowledgements

The authors gratefully acknowledge the assistance of Major Jim Jansen from the United States Military Academy (West Point), Lew Hassell of Drexel University, and Jack Xu, Graham Spencer, Doug Cutting, Amy Smith and Catherine Yip of Excite, Inc. in providing the data for this research. Without the generous sharing of data by Excite, Inc. this research would not have been possible. We also acknowledge the generous support of our institutions.

References

Aigrain, P., Zhang, H.J. and Petkovic, D., 1996. Content-based representation and retrieval of visual media: a state-of-the-art review. *Multimedia Tools and Applications* **3** 3, pp. 179–202.

Armitage, L. and Enser, P.G.B., 1997. Analysis of user need in image archives. *Journal of Information Science* **23** 4, pp. 287–299.

Bach, J., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R. and Jain, R., 1996. The Virage image search engine: an open framework for image management. *Storage and Retrieval for Still Image and Video Databases, Proceedings of SPIE* **2670**, pp. 76–87.

Carson, C., et al. (1997). Region-based image querying. *Proceedings of IEEE Workshop on Content-Based Access to Image and Video Libraries*, (pp. 42–49). San Juan, Puerto Rico.

Cawkell, A. (1993). Indexing collections of electronic images: a review. *British Library Research Review, 15*.

Enser, P.G.B., 1995. Progress in documentation: pictorial information retrieval. *Journal of Documentation* **51** 2, pp. 126–170.

Enser, P., & McGregor, C. (1993). Analysis of visual information retrieval queries. *British Library Research and Development Report, 6104*.

Fidel, R., 1997. The image retrieval task: implications for the design and evaluation of image databases. *The New Review of Hypermedia and Multimedia* **3**, pp. 181–199.

Flickner, M., et al., (1995). Query by image and video content: the QBIC system. *IEEE Computer, 28* (9), 23–32.

Goodrum, A. (1998). A cognitive approach to representing moving image documents. In W. Mustafa el Hadi, J. Maniez, & S. Politt, *Advances in knowledge organization: Vol. 6. Structures and relations in knowledge organization. Proceedings of the Fifth International Society for Knowledge Organization Conference. August*, *1998 Lille*, *France* (pp. 256–263). Wurzburg: Ergon.

Goodrum, A., & Kim (1998). *Visualizing the History of Chemistry*: *Queries to the CHF Pictorial Collection*. Report to the Chemical Heritage Foundation Pictorial Collection. http://www.chemheritage.org/Publications/ChemHeritage/Goodrum/goodrum.htm.

Goodrum, A. and Martin, K., 1999. Bringing fashion out of the closet: classification structure for the Drexel historic costume collection. *Bulletin of the American Society for Information Science* **25** 6, pp. 21–23.

Gudivada, V.N. and Raghavan, V.V., 1995. Content-based image retrieval systems. *IEEE Computer* **28** 9, pp. 18–22.

Hastings, S. K. (1995). Query categories in a study of intellectual access to digitized art images. *Proceedings of the American Society for Information Science Annual Meeting*, *3*, 3–8.

Hourihane, C. (1989). A selective survey of systems of subject classification. *Computers and the History of Art*, 117–129.

Idris, F. and Panchanathan, S., 1997. Review of image and video indexing techniques. *Journal of Visual Communication and Image Representation* **8** 2, pp. 146–166.

Jain, R. (1993). Workshop report: NSF workshop on visual information management systems. In Niblack, W. R., & Jain, R. C., *Storage and Retrieval for Image and Video Databases*, *Proceedings of SPIE*, *1908*, 198–218.

Jansen, B. J., Spink, A., & Saracevic, T. (1998). Failure analysis in query construction: data and analysis from a large sample of web queries. *Proceedings of the Third ACM Conference on Digital Libraries*, Pittsburgh, PA, 289–290.

Jansen, B., Spink, A., & Saracevic, T. (1999). Relevance Feedback*. Proceedings of WebNet 99*: *The World Conference of the World Wide Web*, *Internet*, *and Intranet*, Hawaii. October 1999.

Jansen, B.J., Spink, A. and Saracevic, T., 2000. Real life, real users and real needs: a study and analysis of user queries on the web. *Information Processing and Management* **36** 2, pp. 207–227.

Jorgensen, C., 1995. The applicability of existing classification systems to image attributes: a selected review. *Knowledge Organization and Change* **5**, pp. 189–197.

Jorgensen, C. (1999). Retrieving the unretrievable in electronic imaging systems emotions, themes, and stories. *Proceedings of SPIE – The International Society for Optical Engineering*, 3644, Jan 25–28, 1999.

Keister, L. (1994). User types and queries: impact on image access systems. In *Fidel*, *R.*, *Challenges in Indexing Electronic Text and Images*, *Proceedings of the American Society for Information Science Annual Meeting* 1994, 7–22.

Lancaster, F. (1998). *Indexing and abstracting in theory and practice* (2nd ed.). London: Library Association.

Lawrence, S. and Giles, C.L., 1999. Accessibility of information on the web. *Nature* **400**, pp. 107–109.

Lunin, L., 1987. Electronic image information. *Annual Review of Information Science and Technology* **22**, pp. 179–224.

Ma, W. and Manjanath, B., 1998. Netra: a toolbox for navigating large image databases. *Proceedings of IEEE International Conference on Image Processing (ICIP97)* **1**, pp. 568–571.

Markey, K., 1984. Indexer consistency tests: a literature review and report of a test of consistency in indexing visual materials. *Library and Information Science Research* **6**, pp. 155–177.

Markey, K., 1988. Access to iconographical research collections. *Library Trends* **37** 2, pp. 154–174.

O'Connor, B., O'Connor, M. and Abbas, J., 1999. Functional descriptors of image documents: user-generated captions and response statements. *Journal of the American Society for Information Science* **50** 8, pp. 681–697.

Ornager, S., 1997. Image retrieval: theoretical and empirical user studies on accessing information in images. *Proceedings of the 60th Annual Meeting of the American Society for Information Science* **34**, pp. 202–211.

Panofsky, E., 1967. *Studies in iconology*, Harper and Row, New York.

Rasmussen, E., 1997. Indexing images. *Annual Review of Information Science and Technology* **32**, pp. 169–196.

Seloff, G.J., 1990. Automated access to the NASA-JSC image archive. *Library Trends* **38** 4, pp. 682–696.

Shatford, S., 1986. Analyzing the subject of a picture: a theoretical approach. *Cataloging and Classification Quarterly* **6** 3, pp. 39–62.

Shatford-Layne, S., 1994. Some issues in the indexing of images. *Journal of the American Society of Information Science* **45** 8, pp. 583–588.

Smeulders, A. and Jain, R., 1997. *Image databases and multimedia search*, World Scientific, Singapore.

Smith, J., & Chang, S. (1997). An image and video search engine for the world wide web, In Sethi, I., & Jain, R., *Storage and Retrieval for Image and Video Databases V*, *Proceedings of SPIE, 3022*, 84–95.

Spink, A. and Saracevic, T., 1997. Interactive information retrieval sources and effectiveness of search terms during mediated online searching. *Journal of the American Society for Information Science* **48** 8, pp. 741–761.

Stricker, M., & Orengo, M. (1995). Similarity of color images. In Niblack, W. R., & Jain, R. C., *Storage and Retrieval for Image and Video Databases III*, *Proceedings of SPIE*, *2420*, pp. 381–392.

Tamura, H., Mori, S. and Yamawaki, T., 1978. Textual features corresponding to visual perception. *IEEE Transactions on Systems, Man and Cybernetics* **8** 6, pp. 460–472.

Turner, J., 1990. Representing and accessing information in the stock shot database at the national film board of Canada. *The Canadian Journal of Information Science* **15**, pp. 1–22.

Turner, J., 1994. Indexing ordinary pictures for storage and retrieval. *Visual Resources: An International Journal of Documentation* **10**, pp. 265–273.

Turner, J. (1999). Words and pictures in information systems for moving images. *Proceedings of the 1998 AMIA Conference*, 7–12 December 1998, Fontainbleau Hilton, Miami. Beverly Hills, CA: Association of Moving Image Archivists, 309–315.

Zheng, M., 1999. Metadata elements for object description and representation: a case report from a digitized historical fashion collection project. *Journal of the American Society for Information Science* **50** 13, pp. 1193–1208.