



Searching Digital Image Libraries

Julie J.S. Parker
Information Infrastructure Laboratory
HP Laboratories Bristol
HPL-2002-184
July 8th, 2002*

E-mail: julie_parker@hpl.hp.com

web, digital
image library,
image searches,
keyword
search,
collaborative
problem
solving, human
interaction

Publications such as consumer magazines rely heavily on image libraries as sources for the images they publish in their issues. Traditionally, magazine staff discuss their image requirements over the telephone with library staff and the library conducts the search. Many libraries have now developed Web sites allowing their customers to search for images themselves. However this survey found that these digital libraries are not popular with magazine staff, who continue to prefer the traditional means of contact. The reasons for this preference, and the reasons it is unlikely to change are explored.

* Internal Accession Date Only

Joint Conference on Digital Libraries: Workshop on "Usability for Digital Libraries", 18 July, 2002 Portland, Oregon

© Copyright Hewlett-Packard Company 2002

Approved for External Publication

Searching Digital Image Libraries

Julie J S Parker

Hewlett Packard Research Laboratories
Filton Road, Bristol, BS34 8QZ, England
+44 (0)117 312 8393
julie_parker@hpl.hp.com

ABSTRACT

Publications such as consumer magazines rely heavily on image libraries as sources for the images they publish in their issues. Traditionally, magazine staff discuss their image requirements over the telephone with library staff and the library conducts the search. Many libraries have now developed Web sites allowing their customers to search for images themselves. However this survey found that these digital libraries are not popular with magazine staff, who continue to prefer the traditional means of contact. The reasons for this preference, and the reasons it is unlikely to change are explored.

Keywords

Web, digital image library, image searches, keyword search, collaborative problem solving, human interaction.

INTRODUCTION

This report results from a survey of commercial users of digital and conventional image libraries. The users studied were consumer magazines, chosen because they are heavy image library users. According to BAPLA, the British Association of Picture Libraries and Agencies, magazines represent the second largest buyers of images, purchasing the use of almost a quarter of all library images [2]. Fourteen staff from fourteen magazines in England took part in face-to-face and telephone semi-structured interviews. The magazines' domains ranged from parenthood, health and fitness, gardening and other leisure activities, music, current affairs and science. Using a recent issue of their magazine as a prompt, participants were questioned upon their requirements of images, how these influenced the ways in which they search, and where they search. They were also asked about their use of web-based digital image libraries, how these compare with the traditional ways in which they deal with image libraries, and with their uses of digital and conventional images.

The sale of images (primarily photography but also art work) is a large and growing industry. The most recent statistics from BAPLA gave estimates of over 350 million images held over 600 image libraries in Britain [2].

These libraries earned more than £110 million from the sales of the use of their images in 1999.

However, while most image libraries had either their own Web site, or had access to one, it was also estimated that only 4% of the images libraries sold were through digital transactions. Most of the Web sites were not set up to support commercial transactions. CEPIC (the European Union's equivalent of BAPLA, whose membership includes British image libraries) also carried out a survey of its members in 2001 [3], based on the BAPLA survey. They too found that while a majority of image libraries have their own Web sites, only a minority use these for digital transactions, and that most pictures are still dispatched to their customers using couriers and mail, rather than electronically.

In the following sections, magazines' requirements of images, and the significance of images in magazines, are explored. The importance and complexity of their requirements combine to make the process of locating appropriate pictures a time-consuming and challenging task for the specialist picture researchers whose job it is. How images are found through conventional and digital image libraries, and the advantages and problems of each are described. The current generation of digital image libraries have not necessarily made the work of finding images easier, nor have digital images eliminated the need to see the "real" pictures. Digital image libraries have, rather, changed the problems picture researchers face.

These difficulties, and the ways in which digital libraries might need to develop in response are explored.

THE USE OF IMAGES IN MAGAZINES

Images are used intensively in consumer magazines. A single issue of an average size monthly magazine may contain several hundred images ranging from the full colour “double page spread” for a feature article, to small “filler” pictures for news items. The requirement for an image generally starts with a ‘brief’ from a magazine’s editor to the picture researcher and/or picture editor. The brief is a verbal description of an image or set of images required to illustrate an article. The brief may be very precise (‘pictures of a specific village on a misty day at dawn in the winter’) or very general (‘pictures of whales’). Roles vary from magazine to magazine, but the picture editor is generally responsible for the final choice of the images, subject to the magazine editor’s approval, and for the layout and arrangement of images on the page. The picture researcher finds the images, and carries out the administration associated with handling the pictures received and paying for those whose use is purchased.

The most basic requirements of images are to meet the brief and to be of outstanding or high technical quality. However, images must meet other criteria as well, and serve several functions. The most basic is to provide structure and interest to a page. Images help to break up a page, and attract the viewer’s attention to the article. Images also inform and ‘bring to life’, supporting and enriching the textual content of an article.

Images also promote emotional responses to the content such as excitement, aesthetic appreciation or shock. Considerable time and care is spent in editorial sessions upon the selection of images that will help to elicit the type and degree of emotional response sought. Images also have aesthetic and motivating functions. The beauty of an image may inspire the reader: beyond attracting her to look at the pictures and read the article, the editors may hope to encourage a particular behaviour, whether that be to help young parents adopt safer techniques when bathing their newborn baby, or to inspire people to take up cycle touring.

Magazines must survive in a very competitive marketplace. They have to achieve differentiation from their competitors. To do this, they develop an approach to their subject, or ethos, which guides the type of articles their issues contain. Along with their ethos, they have a house style or ‘look’, and of course, images are fundamental to the look, and a vital part of a magazine’s communication with its readers. The look reinforces the ethos and says to potential and existing readers: ‘this is our style and these are our values: we are aimed at people like you’.

CONSTRAINTS UPON IMAGE SELECTION

If images had only to satisfy these requirements, the editorial team’s job would be relatively easy. However, there are several factors which constrain their choices and make selection more complex. Pictures have to be consistent with the magazine’s ethos. The magazine aiming to teach young parents the correct way of bathing a baby would only consider pictures where the child was being correctly held and supported. The picture has to be what it claims to be. A natural history magazine would not use a picture of an extremely rare mammal taken in a zoo, even if the picture were highly aesthetically pleasing and alternatives were hard to find, if the article is about the animal in the wild.

The look of the magazine is not determined just by the editorial staff of the magazine. Reader reaction is a strong force in constraining the range of possibilities the magazine can consider. For example, prior experience of reader aversion to distressing images of war will limit subsequent choices of photographs illustrating conflicts .

On a practical level, an image also has to fit well in the article as a whole. It has to balance in colour and style with other images on the page and to contain material which complements them. The same is true if the article spans several pages. A beautiful and appropriate image may be rejected if it cannot be made to fit with a group of others.

The process of determining which pictures will be used in an issue is, therefore, highly complex. The criteria are many and varied, subjective and objective. There are also additional constraints which result from experience in the field, such as: ‘this is poor quality but it’s the only picture of this

endangered animal’, or ‘our main competitor used this picture last year.’

THE TRADITIONAL IMAGE LIBRARY

Magazines locate images from a number of sources. The libraries are a main source because of the vast repository of images available through them. Libraries range from the tiny specialists to large multi-subject libraries. While images used from libraries may also be published anywhere else (from books to advertising brochures), although with some constraints, they are an economic way of meeting the need. Commissioned photography guarantees unique pictures but is too costly to be used other than for cover pages and feature articles.

When the magazine picture researcher has been given a set of briefs, he will contact the library picture researcher. The same searches are often requested at several libraries to maximise the chance of finding the right images. The telephone is the preferred mode of contact, because unlike fax or email, a discussion about the requirements is immediately possible. Some briefs are straightforward, but more usually, for the reasons given in the previous section, qualification and clarification of the brief is needed.

When requirements are not straightforward (e.g. a picture which ‘conveys some of the legal complexities surrounding divorce’), there will often be a discussion between them. As three magazine researchers said:

“I prefer using the phone because I can emphasise what’s important to me and we can develop it.”

“Phoning them is quicker and you can make yourself understood. You can have a continuing conversation. They will add suggestions too which could help: ‘We’ve got a photographer who’s done this’. You give them the bones of the story and they’ll make suggestions.”

“The human contact is incredibly important, because they’re going to know their collection.”

The discussion draws on the expertise of both participants. The library researcher knows the library’s collection of pictures; the magazine researcher understands the magazine’s look, ethos and set of constraints. The magazine researcher describes the brief and talks around it, while the library researcher explores with her possible solutions. “Have you thought about this type of

picture?” or “We have pictures on X – would they be any good?”. In other words, they work together in a collaborative problem solving exercise, pooling their different knowledge and experience. Magazine researchers emphasised the high value they place upon this team work. They particularly appreciate it when they start off without any well-formed ideas for suitable images for a given brief. It was also clear that magazine researchers like working with libraries where they can develop long term relationships with their researchers. Not only is there pleasure in working socially; it assists the process of finding good images.

“Personal relationships with people in image libraries makes the process of finding images so much easier. Over the years the person comes to understand your requirements. Their match gets better, but it’s never good enough.”

“You build a relationship with a picture library over the years. Some only employ one or two people. They are nice and easy to talk with and they’ll help me, even if I call and say ‘I don’t know exactly what I need - can you help me?’”

PROBLEMS WITH THE TRADITIONAL PROCESS

One of the most frequently mentioned problems with the traditional process described above is that the researchers have to make do with verbal descriptions of the ideas and images they are discussing. Given the high degree of subjectivity in the process, not only is this inefficient, but often yields results that are disappointing to the magazine team. A picture may fit the objective aspects of a brief but may fall short in other ways. There may be subsidiary aspects of the picture that do not fit the requirements of the magazine’s look or ethos.

Image libraries send magazines selections of images for each brief, as negatives, via the mail or courier. They often prefer to err on the side of caution, and so also send images which may only be of marginal relevance to the brief. The magazine researcher is responsible for the physical processing of images: receiving them from the libraries, storing and tracking them to make sure none get lost, and sending them back. Since each issue might involve searches at up to twenty libraries and receiving thousands of images, the administration is not a trivial task and can occupy several days of the researcher’s time. Then there are

the costs involved: the library's search fees and the transport costs. Disappointing selections are not just a setback in locating an appropriate image; they are also costly in time, money and effort.

Many libraries produce glossy catalogues as guides to the sorts of pictures contained in their collections. These are not very popular with magazine researchers for two reasons. First, they contain only subsets of the library's pictures, and second, it is often difficult to find in which section pictures of interest have been 'filed'.

DIGITAL IMAGE LIBRARIES

Most image libraries already developed their own web-based image libraries. These consist of digitised copies of some of the prints or slides in their collection. Search is usually by key word. Key words are intended to capture the objective and subjective aspects of each picture, and there are an arbitrary number of them associated with each. Viewers can browse images and order analogue copies or download digital copies from the site (though as the BAPLA [2] and CEPIC [3] surveys found, most sites are not yet equipped for e-commerce). The key feature of digital image libraries is that collaborative problem solving between the picture researchers is no longer possible, as the library researcher has been removed from the process. This is good from the library's perspective, as it frees the researcher for other work. However, the magazine researcher is now expected to do the search on the Web site for herself.

PROBLEMS WITH DIGITAL IMAGE LIBRARIES

The picture researchers in this study showed antipathy towards searching in digital libraries. Their principal reasons included the considerable time pressures under which editorial staff already work.

"They [libraries] don't understand about the economy of time. We work under enormous time pressures."

"You never feel you have more time – it's just time spent differently."

"Beautiful Web sites, but you just don't have the luxury of time to look at them. You want a picture and you want it *now*. Speed is of the essence."

"It takes too long to get into some Web sites. You have to enter passwords, and if you haven't got

one you have to ring them up and organise it. ... I still prefer phoning up and getting someone else to do it for me."

While magazine researchers enjoyed problem solving with library researchers, reporting it to be one of the most enjoyable parts of their work, search on Web sites is largely stressful and unfulfilling. They are stressful for several reasons. First, doing the searching themselves requires that magazine researchers find new time in which to do a job formerly done by the library researcher.

The inefficiency and frequent lack of results from the process of searching are the other main reasons web searches are disliked. Image libraries have independent schemes of classification, each of which has to be learned. One magazine researcher reported that she regularly called across the office to ask her colleagues for help with synonyms, having had no success with obvious words. Key word searches are too hit and miss:

"...if I ask to see Series 3 it will show me 180,000 images and it's too much to scroll through. If I make it more specific then I get nothing at all. It's a difficult balance – you can get too many or miss the one you want by missing out one piece of information."

"Sometimes you come across a picture and you think 'why haven't I seen that before? It should have been available under different key words too.'"

Key word searches proceed by trial and error. On top of this, there are the 'dead times' picture researchers must spend as a result of system and network delays:

"Searching is another thing which consumes our time. It's so busy, and the internet slows us down, waiting around."

"Occasionally I use the Web, but I don't have the time to search – loading up, trawling through and searching takes too much time. It's easier for me to ring up and get someone else to do it for me."

Not only is the process inefficient; it has to be repeated multiple times at different libraries, for the same reasons traditional searches are requested at many libraries – this is more wasted time. There is

also no way of knowing, without calling the library, whether search failures result from true negatives (no available picture), or false ones (the right keywords have not been used).

The researcher needs to contact the library anyway, for a different reason. As few libraries can afford to digitise their entire collection, they must make judgements about which images are most likely to yield a return on investment. So those they deem most likely to sell are those they digitise. The requirements of magazines, by contrast, are more likely to be for the unusual or highly specific. Since the digital image library is not complete, the magazine researcher still has to telephone the library to check whether there are other pictures. Why then bother with an unnecessary intermediate and time-costly step if the phone call has to be made anyway?

Interestingly, the reasons given for disliking Web and catalogue searches were similar. Both are subsets of collections and are therefore not reliable guides to the collection as a whole. In Web searches, the problem of finding images changes from 'where in the catalogue?' to 'what keywords?' but evokes similar problems.

For all these reasons, magazine researchers generally prefer to have traditional searches conducted for them. At the moment, they seem to have inherited the worst of both worlds. Overall, the Web consumes extra time in the searching process rather than saving it. Additionally, time savings have not been made on administration, as a result of editorial teams' preferences to work with transparencies and slides (principally so the quality and characteristics of the image can be seen) rather than digitised images. However, of the fourteen magazine staff interviewed, there were two who regularly used both digital and traditional processes to advantage. Typically they conduct an initial search at a given digital library to identify a subset of candidate pictures, and then ring the library researcher to discuss it and order a small number of transparencies.

One of the two magazines uses many fewer pictures per issue than most: typically, just one or two per article. It is therefore easier for her to put together a suitable set of pictures. Unusually, her magazine used many photos that were digital in origin, to illustrate

current news articles, which also saved her considerable administration.

The second - a picture editor - took search even further. He conducted initial Web site searches, identified a selection of promising pictures for each brief, and then telephoned the library researcher. They then pursued the search together over the phone, so that both could view and discuss search results on their browsers. This approach - overcoming the main disadvantages of search in both traditional and digital libraries - yielded the highest degree of reported satisfaction with the use of digital libraries.

Why did the other researchers not also adopt a technique like this? The answer would appear to be due largely to the pressure of time. For many reasons, not least of all a budget and a picture research team both larger than average, work in this particular magazine was felt to be busy rather than constantly pressured.

DISCUSSION AND CONCLUSIONS

There is a great deal of research being conducted today into methods of searching for images based upon some aspect of image content (for recent research, see [1]).

However, at least for magazine researchers and others like them who work under considerable time pressure, it seems unlikely that improved and content-based search methods would answer their needs, since searching is still a significant consumer of time. Given that picture researchers reported how highly they valued joint problem solving with the library researchers, it seems that enhancing their capability to do this might be one advance that could offer them more productivity gains. Web, or IP (Internet Protocol) telephony [5] (phone calls through the Internet between multimedia PCs and integrated with Web sites) may offer magazine researchers the best of both worlds when the technology reaches maturity [4]. It offers the magazine researcher the ability still to delegate the search to the library researcher, but to benefit from the interactive visual feedback that the traditional process lacks. This may conflict however, with the libraries' aims for returns on their investment in digital libraries, which include freeing their researchers for other work.

The value of digital libraries may be more apparent to magazine researchers through other service developments to inform them and save them time. The CEPIC study [3] found that fewer than a fifth of digital libraries were members of a portal site. Portals could reduce the time spent in repeating searches, although they do not overcome the 'false negative' dilemma. Other examples include more use of alerts via email ('we have new pictures in your area of interest'), and more information, such as: 'the history of use of this picture is...' and 'we have additional pictures on this subject that are not in our digital collection.'

- Pacific Rim Conference on Multimedia, October 2001.
2. British Association of Picture Libraries and Agencies (BAPLA) "*What is the Value of a Picture?*" 1999 Press Release, available by request from BAPLA.
<http://www.bapla.org.UK/industry/index.asp>.
3. CEPIC. Transcript of the European Industry Survey presentation by Peter Remmersen, CEPIC Congress, Amsterdam, 25 May 2001.
<http://www.cepic.co.org/>
4. IDC Bulletin. *Evaluating New Telephone Systems: Strategic Benefits Are Coming, But Are Products Mature?* Paul Strauss, August 2001.
5. IP Telephony Basics.
www.intel.com/network/csp/solutions/ipt/index.htm

REFERENCES

1. *Advances in Multimedia Information Processing* – Proceedings of the 2nd IEEE