THE IMAGE DATABASE
A NEED FOR INNOVATIVE INDEXING AND RETRIEVAL

Susanne Ornager, Royal School of Librarianship, Birketinget 6, 2300 Copenhagen S, Denmark

Abstract: This paper touches upon the problems arising in connection with indexing and retrieval for effective searches of digitized images. Different conceptions of what subject indexing means are described as a basis for defining an operational subject indexing strategy for images. Fields of application discussed include the messages in an image, linking between information running from text, image to object, user group requirements, and development of user interfaces.

1. Introduction

To state that the image is an information resource is not a recent phenomenon. Images have been collected, stored and retrieved for centuries, but images have been associated with what could be called “special collections” i.e. collections for users with special interests. As the images have belonged to “collection minorities” they have not technically been considered as interesting as document collections. The research carried out in the information area has focused on functions such as database description, data models, retrieval methods etc. for printed material.

The situation is changing. This change is due to:
- the change in society in general i.e. the change to a more visual oriented community
- the technological development i. e. faster and cheaper computers and storage medias, scanners, hypermedias, communication devices, etc.

As the technology allows the image collection to be stored and retrieved by the use of computers, the image can to day be seen as an information resource like any other resource stored in a database.

The term “image database” has been used in different ways in the literature e.g., as a new compression method for storage of images, as a textual (non-image) database containing descriptions of images rather than the images themselves, as a database containing full text documents scanned into the computer, as an interface using icons, about computer graphics, and so on.

There seems also little consensus on how image databases might be generalized to allow the development of data models, such as have emerged for non-image data. Significant image database concepts have yet to be established because there are far too few systems that can be called true image databases. This implies that such a database would have symbolized image features which would be managed uniformly with non-image information (Tamura, 1984). Technically one can refer to the concept of a picture database as a record and field database, with a special field to store an image (Raskin, 1987).

Image databases as defined above have appeared on the market, but they vary widely in how much they can do with an image. Just as database management systems for non-image data covers a wide spread of products from the simplest to the powerful, the same seems to be true of image databases. We find systems where the images and the databases that describe them have been kept separately, and systems on videodiscs containing analogue stored text and images. Lately digital processing has been claimed to be the solution to the problem of marrying image and text, although videodisc held some promises for merging image and text databases, it has shown that subject access was a problem as the images were still divorced from the descriptive text.

We know how to capture, store, and transmit images, however, while many of the technological aspects of image processing appear to be hardware dependant and solvable today, innovative indexing and retrieval for effective retrieval of digitized images still seem to be in the realm of frontierland. The fascination of technology seems to have outshined the reasons for storing the images which is RETRIEVAL.

In this paper the focus will therefor be on criteria and concepts for image retrieval.

2. Indexing and retrieval of images

Talking about indexing we talk about two types of subject information:
- The information which is explicit i.e. information which is expressed in the terminology applied by the author of the document.
- The information which is implicit i.e. information which is not directly expressed by the author, but which is readily understood by a (human) reader of a document.

The methodology generally recommended as good practice for subject analysis and indexing is a two step exercise:
- One analyses a subject matter and expresses the conceived information in a concrete statement, for instance in the form of an index term.
- One translates the indexing term to a controlled vocabulary of indexing terms for instance a thesaurus. The preferred term acts as a surrogate for the concept.

The content-oriented conception of subject indexing relying on both explicit and implicit information in documents for determining adequate document representations is the most common approach to subject indexing.

The true image database (Tamura, 1984) contains a visual representation of the image in connection with a text. The primary purpose of a the supplementary verbal description is to retrieve a number of pictures - on the basis of different subject criteria - from the total volume of images in the base. The description must include more elements than those offered by the descriptive cataloguing. The elements to be emphasized must be determined by means of a subject analysis of the material.

3. User requirements

A user group's requirements must be well defined (known) to enable the indexer to include them in the indexing, and hence, it will be necessary to perform empirical user studies, which will often be an extensive task, even in restricted domains. Another major problem is that the immediate requirements of a particular user group may be exaggerated at the expense of future users. Being aware of these problems it is important to emphasize that the subject analysis based on the documents themselves must always be included in the indexing.

A high proportion of the working systems described in the literature of the last ten years have been developed for art-historical purposes (for an archive, gallery, research project or similar function). That these systems have been developed for a specific discipline with its own repertoire of subject-matter (however broad that may be) and associated themes, has usually moderated the difficulty of designing and building them. At the very least it has made it possible to define clearly the area of the application and the interests of the users.

Developing a database for pictures of a wider, more general subject-matter does not normally allow such definitions to be made, and it may look as an impossible task to extrapolate from the particular to the general for creation of such a databases. However, we decided to do an empirical study involving a number of 25 newspaper archives in order to determine rules for image representations by detecting the components focused on during the content analysis (when starting with an image why do we have that specific description of its content), to categorize the user population according to their ways of asking questions, and to analyze the questions asked.
When indexing the photos many archives focus on background information, specific names and/or events. The story associated with the photo actually label the surface found in the image. Although the supporting story is important for the content analysis the survey also showed that elements such as feelings and emotions are essential for describing the photo.

The users of the newspaper archives are mostly journalists. Although they have the same educational background their requirements vary because of the different topics they cover. The analysis of the survey shows that the users can be categorized within five groups:

1. The ones who ask very "narrow" because they have a photo in their mind
2. The ones who ask very "broad" because they want to choose for themselves.
3. The ones who tell about "the story" and are open to suggestions.
4. The ones who give "the story" to the archive and want them to find the photo(s) because "the archive knows best".
5. The ones who only care about the size of the photo in order to fill in an empty space on the page.

Each of the groups will focus on certain aspects of the photo and to cater for their queries it is necessary to take these facets into account when analyzing the picture, e.g. type 1 will need the collection to be indexed in a very exhaustive and specific way, while type 5 only needs a superficial content description containing the size of the photo.

The questions asked vary as can be seen from the grouping of the users. In most of the archives 50% of the questions are about named persons while the other half includes subjects, of these 20% are so called "difficult" queries where an in-depth interview is needed in order to find the photo(s). Examples of the latter kind of questions are: Loneliness, depressed people due to unemployment, morning atmosphere, depopulation of small Norwegian villages, impotence, jealousy etc.

Not a great deal is known about the kind of questions that users are likely to put to an image collection. There is virtually only one analysis of the kind of questions asked and the success of the system in responding to those questions in the published papers. The one exception is Enser's (Enser, 1993) work on the Hulton picture collection, and although this collection aims at a different user group it actually shows the same sampling of questions as was seen in the newspaper archives - for example "great personalities of the 20th century", or "street scenes in 1850".
4. Term association clusters

The collections chosen for the survey hold documentary and press photos (Gecsei, 1989; Strunck, 1991). It may be asserted that what characterizes such photos is that they show reality in the form of events or objects which may be verbalized. Furthermore, such photographs are a subjective interpretation of reality and a very small part at that. It may present some difficulty verbalizing the space surrounding the photograph. A simplified definition is that the photograph has a literal message and a symbolic message, and whereas the literal message has a descriptive function - and may be described - the symbolic message is partly subjective (Ornager, 1992). The user interprets the picture through his individual (subjective) and cultural (objective to a certain degree) experience and knowledge. As we deal with the problem of separating a number of photographs from the total number of photographs in a database according to subject, the starting point indicated by the postulated characteristics must be to concentrate on a verbal description of that which the photograph conveys on the literal level. The symbolic message is conveyed through the photograph “itself” on the visual level.

How can one verbalize the literal level? It is known that a thesaurus like the Art & Architecture Thesaurus is performing well as a retrieval tool for text databases, it can be assumed that it will do the same for image databases. Although one can question the role of a thesaurus, a collection of words, when the image is right there, available in the catalogue or database record, it must be emphasized that the thesaurus is a powerful link between the various forms of information in the field of images that run along a continuum from text, image to object.

If the above statement is true the indexing of images is necessary in order to separate them in connection with a subject query, the subject demarcation must be expanded by taking into account the intended use of the image. The selection = the indexing should be made on the basis of the explicit information i.e. the descriptive cataloguing and the picture itself, the implicit information i.e. the assigned keywords from a thesaurus and keywords expressing the intended use of the image.

This points to a series of user experiments in which the image search is improved by the introduction of an (intelligent?) user interface. The starting point of the experiment is an image database with digitized photographs, analyzed according to subject and indexed in keeping with the above-mentioned principles. The thesaurus utilized for the indexing must be on-line. To improve the search facilities a “user interface vocabulary” is constructed serving as a supplement to the thesaurus. Clusters of terms may function as such a supplement. A cluster of terms is here taken to mean a number of words associated with a term. The words may have semantic relations to the term, but the words need not be semantically re-
lated to form part of the cluster.

The cluster of terms annexed to the image base is constructed from user-defined, associative terms. A number of persons are involved in a test period prior to the construction of the clusters. These persons are introduced to a number of terms which must comply with the following conditions:

- they must be descriptors in the thesaurus
- they must have been used for the indexing of photographs.

The thesaurus terms are the stimulus words and the associations evoked with the participating persons by the thesaurus terms are the response words. In the experiment, the persons state an x number of associations, expressed as single words or short phrases. In the analytic phase, the stimulus and response words are compared. Each stimulus word is coupled with the group of response words indicated by the persons participating in the experiment, and the response words constitute the associative cluster of terms relating to the stimulus word (Pejtersen, 1987).

In addition to the thesaurus terms, the persons are confronted with a number of photographs. The persons state an x number of single words or brief phrases to describe each photograph. The words derived from the picture are compared with the thesaurus words applied for the indexing of the picture. A picture-associative cluster of terms is assigned to each thesaurus word. The cluster of terms comprises the words derived from the picture that can be annexed to the thesaurus term.

The two types of clusters - response words and picture-derived words are merged providing the thesaurus terms with associative clusters of terms consisting of both response words and picture-derived terms. It should be noted that the cluster words may be words which also occur as descriptors in the thesaurus.

The associative terms of cluster may form part of the total search vocabulary for the image base, or act as a hidden user interface which is called when a user has retrieved a number of photos and want either to extend the number with associated pictures, or retrieve pictures which represent other aspects on the basis of a picture from the retrieved volume. On the face of it, it seems most logical to let the clusters have both functions. The users’ chances to find what they are looking for increase with the number of approaches possible in the search situation. Thus, it will be possible to search the image base in the initial position both by means of thesaurus terms and cluster terms.
5. The future of image databases

The development of a user interface for image database retrieval has been described in the previous chapter. Will this be the future way of accessing image databases? It is hard to predict the development trends as most of the research in this area has focused on the technical aspects of image databases. One can ask: Why not foretell the future from the technical point of view? There is a vast amount of information on experiments about queries by query-pictures and searching by “thumbnail” (miniature representations of each picture) pictures (Cawkell, 1993).

It can be claimed that due to the nature of visual information the essence of an image can be quickly grasped as compared to textual documents. Browsing access to visual information is therefore a future way to retrieve information, in image databases. If the thumbnail pictures are used a number of 25 to 30 can be accommodated on a screen and still contain sufficient legible detail to represent parent pictures unambiguously. If it is true that a human can discriminate between pictures in this manner, the need to improve an elaborated indexing scheme for discrimination purposes is reduced. However, to start the browsing one needs a template or a picture thesaurus to find the pattern to begin the retrieval - such a pattern search will probably need words.

New advances in image-query systems are shown in the edge-sketch method developed by Niblack & Barber (Niblack, 1993) and in the sketch method developed by Hirata & Kato (Hirata, 1993). In the edge-sketch query method, the user draws a shape, however crude, on the screen, and the machine obliges by presenting those thumbnails which most nearly resemble the sketch in an attempt to reduce the noise retrieved by reason of the sketch's inadequacy. The system is used for approximate matching between an input query-picture and a database picture. The inventors of the sketch method claim that the user has only to draw a rough sketch to retrieve similar images from a picture database. Similarities between the sketch and images in the database are automatically evaluated and the most similar candidates are shown to the user. The system was tested using a database of 205 paintings of full colour landscapes and portraits.

Although the idea of inputting a query-picture has much to recommend it, query-pictures do not always replace the descriptive power of words which may be better for some abstract concepts. For instance it is hard to see how a query-picture could be devised for e.g. “depopulation of small Norwegian villages” or “jealousy”.

We do assert that there will never be a wordless society, however, the future will head towards an integrated image information system, where one can browse through pictures or words, select an image and either via this as a pattern, or a cluster of words based on the indexing terms of the picture navigate through the
system. The image database can be enhanced with a knowledge system, where the clusters of words can be organized in a semantic network, this again could be linked with information about users - a user model built during the navigation, or through an interviewing session.

6. Conclusion

In the paper it was concluded that significant image database concepts needed to be established in order to define the true image database. It was further claimed that although many of the technological aspects of image processing appeared solvable today, the research is still mainly focused in this area, and not in the area of innovative indexing and retrieval. The fascination of technology is still overwhelming.

This contradicts the authors view which is that searchers will continue to access databases, even image databases, by means of words. The focus in research must change from mainly technological aspects to that of building interfaces for browsing and searching. Traditional subject indexing paired with an associative term clustering system was suggested as the basic for a search vocabulary for the image database.

Artificial intelligence and knowledge based systems are central to image management developments. In the future we will probably see more of these technologies as well as the involvement of neural networks in the development and use of image databases. However tempting it is to be innovative one must not forget that we still need rules to be established for the description of images.

References


