AUTHOR(S): BAXTER, G. and ANDERSON, D.

TITLE: Image indexing and retrieval: some problems and proposed solutions.

YEAR: 1995

Publisher citation: BAXTER, G. and ANDERSON, D. 1995. Image indexing and retrieval: some problems and proposed solutions. New library world [online], 96(6), pages 4-13. Available from: https://doi.org/10.1108/03074809510147470

OpenAIR citation: BAXTER, G. and ANDERSON, D. 1995. Image indexing and retrieval: some problems and proposed solutions. New library world, 96(6), pages 4-13. Held on OpenAIR [online]. Available from: https://openair.rgu.ac.uk/

Image indexing and retrieval: some problems and proposed solutions

Graeme Baxter and Douglas Anderson

Abstract
Image processing technologies are offering considerable potential for library and information units to extend their databases by the inclusion of images such as photographs, paintings, monograph title-pages and maps. Discusses problems and potential solutions in a structured fashion based on categories of thesauri (text and visual), hybrids, description language and automatic content analysis, with state-of-the-art examples.

Keywords: Computer networks, Image processing, Indexes, Internet

Introduction

In recent years, as image processing technologies have progressed, and as the costs of the appropriate hardware and software have fallen, the information profession has paid considerable attention to the introduction of image databases: collections of images – photographs, paintings, maps, diagrams, etc. – which are stored electronically and controlled in a database management system. As with any conventional database, a prime requisite of an image database is the provision of means by which its contents can be indexed and retrieved effectively. However, images have attributes which set them apart from text, and which make provision for access far from simple. It is the intention of this article to outline indexing and retrieval problems and examine various approaches which have been made in attempts to provide a solution.

In addition, as interest in the "information superhighway" continues to grow, this work will examine the implications such a diversity of approaches has for those wishing to gain access to remote image collections via electronic networks.

Problems of image indexing and retrieval

The most common method of handling images in a database for retrieval purposes has been to establish fields containing bibliographic data (artist, photographer, title, date, etc.) together with a field containing some descriptive text to support the image field(s). These fields can then be searched using conventional text-retrieval facilities, such as keywords and Boolean operators. Danziger [1] believes that such an approach is relatively straightforward, and that designing an indexing system for an image database is little different from designing any other indexing scheme. Indeed, she states that the indexer, when establishing the desired descriptive criteria, need "identify only those elements that will be used primarily to select pictures from the database" [1].

This, however, is a rather simplistic view, for, as Besser [2] points out, there are two interrelated factors which set images apart from textual material. First, many images are decidedly multidisciplinary in nature: they contain a variety of features, each of which may be of potential interest to researchers from somewhat diverse fields of study. Indeed, such images are often used for purposes far removed from those envisaged by their creators. Besser himself provides an ideal, illustrative example:

> *A set of photographs of a busy street scene a century ago might be useful to historians wanting a "snapshot" of the times, to architects looking at buildings, to urban planners looking at traffic patterns or building shadows, to cultural historians looking at changes in fashion, to medical researchers looking at female smoking habits, to sociologists looking at class distinctions, or to students looking at the use of certain photographic processes or techniques [2].*

Second, unlike books and other documents, images are not described explicitly by titles, prefaces, introductions or abstracts: they make no attempt to "tell us what they are about". As a result, viewers have to rely heavily on their own interpretation of an image and its subject content. This can be difficult particularly when dealing with images which are allegorical or symbolic in nature. Shatford Layne [3], for example, warns that an image of a man and a lion might, in fact, be about pride; and that an image of a person crying might actually be about sorrow. Indeed, the wide variety of meanings which an individual image can convey was highlighted in a small but notable experiment carried out by Sunderland [4] at the Courtauld Institute in London. Here, three people – an art historian, a layman and a 12-year-old child – were asked to describe Millais's Christ in the Carpenter's Shop, a painting which depicts a moment in Christ's childhood that is prophetic of the Crucifixion. This resulted in three diverse descriptions, each of which reflected the experience and the perception of the individual: the art historian concentrated on the symbolic features, the child gave a very descriptive account, while the layman's description lay somewhere between the two.

Given this reliance on subjective, conceptual factors, it is almost inevitable that the perspective of one indexer will differ from that of other indexers. Indeed, Markey [5], when studying the indexing of medieval works of art by a group of non-specialist indexers, found extremely low levels (i.e. 1 to 27 per cent) of inter-indexer consistency, particularly when the more abstract "expressional qualities" of the images were examined. Intra-indexer consistency, too, can be somewhat lacking when dealing with images. Seloff [6], for example, found that similar images in the NASA Johnson Space Centre's film repository, indexed by one individual but separated by time, varied greatly in their textual description. He concluded that the perspective of the indexer is apt to change from one week to the next, and that, furthermore, it is always different from that of the repository's users (i.e. the engineers and scientists). This latter point was echoed by Enser and McGregor [7] who, on analysing over 2,700 requests made by users of the Hulton Deutsch Picture Collection, found that the indexing arrangements "could only characterise the requests at a very general level, providing indexing terms the great majority of which were at some considerable conceptual remove from the focus of the query" [7].

It is apparent, therefore, that the indexing of images is by no means a simple operation; and that, in order to attempt to cater for the variety of access points from which users might try to retrieve images, depth indexing is required. This, though, can prove a rather laborious and costly business, for, as Besser points out, in an adaptation of an old maxim, "A picture may be worth a thousand words, but a thousand descriptors are probably not enough to describe the contents of an image database" [8].

Thesaurus-based indexing systems

In an attempt to introduce a measure of control to the vast number of descriptive words which might be used when indexing images, considerable attention has been paid to the development of thesaurus-based indexing systems. One of the more important of these is the Art and Architecture Thesaurus (AAT) which, as its name suggests, attempts to provide a controlled vocabulary for the visual arts and architecture. Originating at Rensselaer Polytechnic in Troy, New York in 1979, the first edition of the AAT was published in 1990 by Oxford University Press. It currently contains over 50,000 terms “hierarchically arranged according to a rigorously constructed, internally consistent structure” [9] modelled on that of the National Library of Medicine’s Medical Subject Headings. These hierarchies contain terminology describing physical attributes, styles and periods, agents, activities, materials and objects; and the terms themselves are derived from existing glossaries, subject lists, thesauri, reference works and scholarly monographs, as well as from the subject expertise of art historians and architects. Although the AAT has become increasingly widely used, it would appear that it can be a rather time-consuming system to operate. Indeed, Keefe [10] found that, even after training, it took an average of 40 minutes to index a slide and input the details into Rensselaer’s INFOTRAX database.

ICONCLASS, another major thesaurus-based system, was originally developed in the late 1940s and early 1950s by Henri van de Waal, Professor of Art History at the University of Leiden, who wished to devise a scheme that would provide a “consistent classification of all the subjects which mankind has succeeded to portray” [11]. Following van de Waal’s death in 1972, the project was continued by a group of his former students led by Couprie, and, subsequently, ICONCLASS (the name is derived from “iconography” and “classification”) was published in the period 1973 to 1985. It consists of 17 volumes of hierarchically arranged alphanumeric codes, each of which is associated with a textual description of a particular subject, theme or motif to be found in fine art. These codes are arranged within nine primary divisions:

(1) Religion and magic;
(2) Nature;
(3) Human being, man in general;
(4) Society, civilization, culture;
(5) Abstract ideas and concepts;
(6) History;
(7) Bible;
(8) Literature;
(9) Classical mythology and ancient history.

Described by Roberts [12] as “one of the most sophisticated and complex of the subject indexing systems geared to the fine arts”, ICONCLASS has received considerable attention, particularly in Europe and the United States. For example, it is used to index large collections at the Marburg Photographic Archive, the Courtauld Institute of Art in London, and the Harvard University Fine Arts Library.

Despite this widespread use, however, it would appear that a number of authors regard ICONCLASS as being a far from ideal solution to the problems of image indexing. Sunderland [4], while agreeing that ICONCLASS is well suited for dealing with the traditional iconography of art history, feels that it is less satisfactory when dealing with the description of common objects, such as tables, rivers or houses. He also believes that, if the system is to be effective, the indexing itself really has to be carried out by well-qualified art historians. Meanwhile, Pearman [13] criticizes

ICONCLASS for demonstrating a marked bias towards Western art after the Renaissance. Indeed, she claims that the apparent dismissal of non-Western art by both ICONCLASS and the AAT displays an ethnocentricity which has no place in today's global village. Furthermore, Cawkell [14] points out that ICONCLASS, like the AAT, can prove somewhat time-consuming to use. At the Courtauld Institute, for example, it takes around 30 minutes to index an image. Another established thesaurus-based system is TELCLASS. Originating in 1979, it has been developed by Evans at the BBC-TV Film and Videotape Library in Brentford, primarily for use with television broadcast material. TELCLASS is described as a "concept classification scheme for the control and retrieval of subjects derived from the whole field of human knowledge" [15]. It, too, consists of alphanumeric codes and associated terms, and these are arranged within six main groups:

(1) verbal;
(2) schematic;
(3) actuality;
(4) simulation;
(5) technical;
(6) formal.

Although TELCLASS was established by the BBC for in-house use, it has since been adopted by other organizations. McGraw-Hill, for example, have used three TELCLASS-type thesauri in their Multimedia Encyclopaedia of Mammalian Biology.

In addition to the three major thesaurus-based systems described above, countless other in-house schemes have been devised, each with its own distinct approach to image indexing. As a result, communication and cooperation between these collections is difficult, if not impossible. Ohlgren [16], in particular, is critical of these "mutually unintelligible Towers of Babel", and believes there is a definite need for standards for the classification and description of images. Hogan [17], however, has a more extreme view of thesauri:

> *In the case of images, the use of thesauri to control inconsistency is not effective due to the individual responses prevalent in human reactions to visual materials. That inconsistency is a reflection of creativity and diversity of human interests, situations, and context. If inconsistency is to be overcome, system designers will need to relinquish the idea of the utility of using words to index non-verbal understanding... We are looking for alternative ways of image retrieval, ways that are less dependent on familiarity with existing taxonomies and their assigned authorities.*

Visual thesauri and related hybrid systems

Hogan's comments are fuelled by a desire to see the introduction of wordless "visual thesauri" systems. This concept has gathered support in recent years, and has been the subject of some notable research, particularly in the United States; although it should be emphasized that, in the experimental systems developed so far, the use of words has not been eliminated completely.

Seloff [6] has described such a system at the NASA Johnson Space Centre's (JSC) film repository in Houston. Designed to manage an extensive collection of space-related images, it allows a more traditional textual approach as well as a prototype visual thesaurus approach. When considering the textual approach, however, it should be pointed out that this system differs from most image databases in that, when responding to a keyword search, it does not employ conventional Boolean logic. Instead, the system's data retrieval engine, Personal Librarian (which is an adaptation of Syracuse University's SIRE system [18]), uses a number of statistically-based search algorithms which rank the retrieved images in order of their similarity to the query. As a result, there is an increased probability that the user will find images relevant to his or her needs displayed towards the beginning of the set of hits, thereby avoiding a potentially time-consuming search through the rest of the retrieved images. This is in sharp contrast to normal, Boolean-based systems, where relevant images will be scattered throughout a set of hits which are presented in a somewhat meaningless sequence, such as physical storage order.

Meanwhile, the JSC system's prototype visual thesaurus is constructed around 200 "popular images" which, with the use of Hypercard software, are linked to associated terms in the JSC thesaurus – a hierarchical thesaurus of aerospace terminology which has been compiled in-house. Consequently, the selection of a descriptive term from the thesaurus retrieves its associated image, as well as broader, narrower and related terms together with their associated images. This prototype, however, has not been without its problems. Although a routine was written which, using the aforementioned statistical algorithms, would automatically select the most significant image to be linked to each thesaurus term, close examination of the chosen images revealed that the majority depicted subjects which were ancillary to the target terms. For example, a term describing a particular type of testing equipment was linked to an image which depicted the item to be tested, rather than the equipment itself. This led Seloff to conclude that linking appropriate images to the several thousand thesaurus terms would require considerable manual effort. Furthermore, subsequent analysis of the system by Seloff revealed that the visual thesaurus approach "is of more interest than utility to casual, or less experienced users, while the textual thesaurus approach is of more use to experienced users" [19]. Bearing this in mind, it is perhaps unsurprising that this prototype system has subsequently been "set aside"; although it should be pointed out that research into NASA's image retrieval problems is being continued at the University of Central Florida's Intelligent Multimedia Applications Laboratory [6].

Mindful of the difficulties encountered by NASA, Hogan *et al.* [17] at Syracuse University propose to take the visual thesaurus concept a stage further by concentrating on a more image-driven prototype. They are currently in the process of developing a system aimed at assisting in plant leaf identification, although they suggest that their model could also be applied in other subject areas, such as in the identification of bones, fossils, pollen spores, coins, weather patterns, etc. In this system, the user is presented initially with a range of images depicting basic leaf and leaflet shapes (i.e. narrow, elliptic, egg-shaped, oblong, etc.) and, on selecting one of these, is then presented with a representative image. Using a number of mechanisms which, Hogan points out, are common in computer-aided design, the user can then manipulate the properties of this leaf image so it then appears similar to the actual example on hand. For example, each image contains "control points" with "handles" which can be moved to adjust the shape of the leaf; while the system provides "palettes" of leaf margins, veins and stalks from which the user can choose those which match the physical specimen. This can be done for both sides of the leaf.

At any point during this process the user can choose to browse through those images in the database which resemble the modified query image. These are ranked in order of similarity and displayed as a matrix of leaves, with the closest match appearing in the centre of the screen and the others radiating outwards in order of decreasing similarity. The user can then select one of these images, either for further manipulation, or to retrieve associated images (of fruit, tree outlines and distribution maps) and textual information.

Although Hogan and his colleagues plan to retain traditional textual access in order to provide a degree of flexibility, they firmly believe that the “visual language” employed by their prototype has a central role to play in future image retrieval systems. Indeed, they openly invite others to undertake additional research in this area.

In some respects, the NASA and the Syracuse University prototypes can both be compared to the Book House database system developed at the Riso National Laboratory in Roskilde, Denmark [20]. The Book House is an icon-based retrieval system designed to assist inexperienced users in finding books, and it includes a “search by pictures” option among its search functions. Here, the user is presented with a series of small, specially-designed icons which attempt to express the subject content of the books in the database. For example, “crime” is represented by an image of an unshaven man in a striped, convict’s suit posing for a “mugshot”. Each image is linked to a number of associated index terms; therefore, by selecting one or more icons, the user effectively builds up a complex keyword search, or “search profile” and subsequently is presented with a set of appropriate book descriptions.

Of course, as has already been seen, the meaning of an image is very much dependent on the perspective of the individual. Therefore, in an attempt to obtain a consensus among prospective users, the Book House design process (much of which was based on the cognitive analysis of users’ information needs and search queries) included a number of word and picture association experiments; and the end results of these investigations (i.e. the icons) were tested by three groups of end-users – adults, children and librarians. Incidentally, it was discovered that the children’s interpretation of the images differed greatly from that of the adults (a finding which concurs with the result of Sunderland’s experiment described earlier in this article). Therefore, the meanings of the icons in the Book House’s children’s database are different from those in the database for adults.

The Book House has received an “overwhelmingly positive response” from its users and, in fact, has become widely regarded as an information retrieval landmark. Bearing this in mind, it is disappointing to note that, despite further work having been planned (including the construction of an “iconic thesaurus” [21]), research into developing the system has been abandoned by the Riso National Laboratory, although development will hopefully be implemented by Apple Computers who have shown interest in the package [22]. Indeed, given its success so far, it is not unreasonable to suggest that this icon-based approach should be the subject of further attention, not only in applying it to larger, more comprehensive, bibliographic databases, but also in investigating its potential as a means of accessing image collections.

Another advocate of a move away from text-only indexing is Besser. He states that:

> *Even extensive text-based descriptions of the images are seldom sufficiently descriptive for the researcher to determine which images are likely to be relevant to his or her needs. Even an enormous amount of descriptive text cannot adequately substitute for the viewing of the image itself* [2, p. 788].

With this in mind, Besser describes a prototype system at the University of California at Berkeley, which, he believes, provides a solution to the problem. Using IMAGEQUERY software, the system allows the user to undertake an initial text-based search; it then presents a retrieved set of simultaneously displayed miniature or “thumbnail” images, through which the user can browse in order to select those which are most appropriate to his or her needs. In addition, any one of the thumbnail images may be enlarged for closer inspection. This type of hybrid model is becoming increasingly popular; indeed, similar browsing facilities now feature in a number of other image retrieval systems. The Hulton Deutsch Collection, for example, provides sets of expandable thumbnail hits in its “People and decades” CD-ROMs [23]; while at the Queen’s University of Belfast, Philip *et al.* [24] have developed a database of photographs covering “Ulster life” which, with the aid of a network of four transputers, provides rapid retrieval and enlargement of sets of miniature “stub” images.

The provision of a browsing tool can certainly prove advantageous, for, as Cawkell [25] explains, “browsing through a succession of rapidly presented blocks of thumbnails immediately brings to bear the most efficient selective system by far – the human eye/brain”. Cawkell also points out that, for small collections at least, the provision of such a facility reduces the need for depth indexing. Larger collections, however, would still have to undertake relatively exhaustive indexing so that the number of images retrieved for browsing is kept at a manageable level.

Picture description languages

Some researchers have devised alternative methods of manual indexing which, they believe, are more suitable for subsequent, automated storage and retrieval. Generally, these take the form of special languages which allow the description of an image to be expressed as a machine-readable shorthand or code.

At the University of London, for example, Leung *et al.* [26] have experimented with a “Picture description language”, based on the Entity-Attribute-Relationship data model. They maintain that this model (which, incidentally, might be likened to the key system, part/property and action/effect categories used in Austin’s PRECIS) corresponds closely to the noun, adjective and verb which generally constitute a simple description of an image. Using a set of prescribed rules and a special syntax (which, again, is reminiscent of that used in PRECIS, as well as that used in its antecedent, Farradane’s relational indexing), this language allows image descriptions to be structured in a uniform manner, and also permits the characteristics of, and relationships between, objects in an image to be expressed and used as retrieval criteria. For example, the description of an image of a black cat sitting on a chair would be expressed as follows: sitting (black cat, chair).

It has been suggested, however, that methods such as this are really only practicable in relatively small collections containing relatively simple images [14, p. 183]. As both the index entries and the users’ enquiries have to be coded and input manually, their application in large, complex collections may prove time-consuming and unwieldy. This is evident when examining other picture description languages. For example, the code devised by Bordogna *et al.* [27] describes the silhouettes of digitized astronomical images in terms of their co-ordinates:

> ELLIPTIC NUCLEUS WITH ENDPOINTS IN THE POINTS (9,24) (17,25) (23,15) (13,14)
> ARM FROM (23,13) TO (23,18) WITH ENDPOINT IN (29,5)
> ARM FROM (9,19) TO (9,23) WITH ENDPOINT IN (3,25)

On the other hand, Chang and Wu [28] have developed an indexing mechanism which uses character strings to represent spatial relationships between different objects in an image. For the example in Figure 1 the character string would be:

T= {(A,B,7),(A,C,8),(A,D,8),(A,E,1), (B,C,8),(B,D,8),(B,E,2),(C,D,1), (C,E,3),(D,E,4)}

where
1 = north of reference object;
2 = north-west of reference object;
3 = west of reference object;
4 = south-west of reference object;
5 = south of reference object;
6 = south-east of reference object;
7 = east of reference object;
8 = north-east of reference object;
9 = at the same location as reference object.

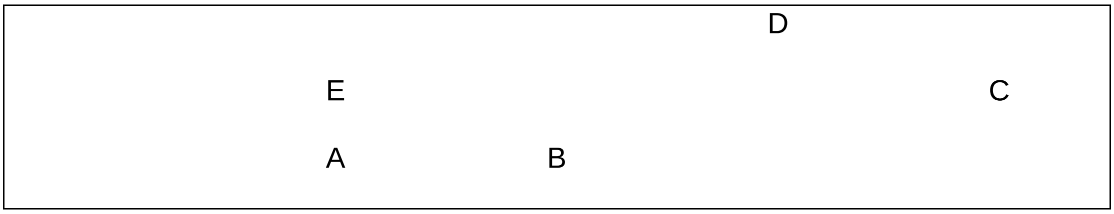

Figure 1: An image containing five distinct objects

Such techniques clearly necessitate considerable intellectual input from the indexer and illustrate a current problem of realistic manpower and time availability for the production of precise quality information retrieval devices.

Indexing and retrieval of images by content

There is currently a great deal of research into the development of systems which, rather than relying on descriptive text, retrieve an image by automatically analysing its content and matching it with a user's query. Much of this work is based on the mathematical analysis of the pixel arrays which constitute a digitized image.

Cawkell explains that in these content-based image retrieval systems, the user's query has to take one of two forms: an image with a structure resembling that of the required image has to be submitted to the system for matching purposes, or a question describing the structure of the desired image has to be asked [14, p. 184]. Among the systems which allow the former approach is that developed by Rickman and Stonham at Brunel University [14, pp. 189-90]. Having its origins in Wilkie, Stonham and Aleksander's WISARD project of the early 1980s [29], this system works by pattern matching an input image of a human face against a collection of black and white images in a facial image database. This it does using an eight-node neural network which, it might be said, attempts to simulate the actions of the human brain's processing units. Successive presentations of the submitted image are made to the system, during which each inter-connected node in the network, by means of a consolidation or rejection learning process, "locks on" to a particular feature or “principal component” of the image. Once this learning process is complete, each node will then generate output data whenever it is presented with that particular feature; therefore, when a number of features representing the whole image are presented to the system, output data will be produced by a corresponding number of nodes. As those images in the database which are similar to the input image will produce similar outputs, image selection by matching can then be carried out.

Cawkell points out that Rickman and Stonham’s work is widely regarded as being of great potential and that, given the equipment and expertise accumulated at Brunel over the years, it is one project particularly worthy of further support. However, he suggests that the method of inputting an image for matching purposes is somewhat limited in scope, and that, really, future work on the system should include an investigation into ways of asking it questions [14, pp. 190-1].

One system which does allow a form of questioning is the Query By Image Content (QBIC) prototype at the IBM Research Division in San Jose. Developed by Niblack *et al.* [30], QBIC is a semi-automatic system which uses the colour, texture and shape of images as the basis for indexing and retrieval. Indeed, when an image is added to the database, the system, using a variety of algebraic formulae, computes automatically the quantitative properties of the colour, texture (i.e. “coarseness”, “contrast” and “directionality”) and shape of the image, and these are stored in the system for use in subsequent queries. In addition, a degree of “segmentation” is allowed. Here, the indexer, using a drawing tool, roughly outlines those objects or elements within an image which are deemed significant enough to be supplied as future hits; the computer, meanwhile, refines or “shrink-wraps” these outlines using Snakes software.

When a query is to be performed, the QBIC user interface allows the searcher to specify the colour, texture or shape of the desired image by using visual examples. For colour selection, the system provides a “colour picker” screen containing a set of RGB sliders which allow the user to choose from 256 different colours; for texture selection, the user is presented with “patches” of synthetically generated textures; while for shape selection, the interface provides a blackboard drawing area in which the user can draw the shape of objects within the required image. These approaches can either be used individually or in any weighted combination. For example, the user searching for a seaside scene can ask for an image which is 25 per cent white (for the sand) and 50 per cent blue (for the sky and water). The query is then matched with the contents of the database using algorithmic techniques.

Alternatively, the user can adopt a “full scene” or “query by sketch” approach. Here, the user roughly sketches the whole of the desired image and the system, again using fuzzy matching procedures, compares it to automatically preconstructed, reduced resolution edge maps of each image in the database.

The QBIC system, with the aid of distance metrics, subsequently ranks the retrieved images in order of similarity to the query, and displays them as a series of thumbnails. Indeed, Niblack is at pains to point out that QBIC does not provide exact matches; instead, it serves as an “information filter” which allows the number of images in a large database to be reduced to a level suitable for browsing.

Niblack also suggests that, for the best results, QBIC should be used in conjunction with conventional text-based searches. This advice appears to have been heeded at the University of California at Davis, where, as part of a project in which QBIC is being applied to a portion of a large slide collection, the Getty Art History Thesaurus is currently being adapted in order to provide a complementary textual approach.

Incidentally, prior to textual indexing being added, QBIC search results at UC Davis have proved somewhat variable, although it is worth noting that searches by colour and by shape have produced some accurate results.

Another content-based image retrieval system which provides drawing tools for constructing a query is the Shape Analysis For Automatic Retrieval of Images (SAFARI) system developed by Eakins at Northumbria University. Here, the searcher uses commands for co-ordinates, direction, length, arc angle, etc. in order to build up a query shape which the computer, using automated content analysis, then compares with the contents of the database. Like QBIC, SAFARI then uses a distance metric to rank the retrieved images.

Since the SAFARI project began in the mid-1980s it has largely been applied to comparatively simple images: first, two-dimensional engineering drawings and then, more recently, trademark images. In fact, Eakins suggests that most of the current, experimental retrieval-by-content systems would prove inadequate if applied to more complex images:

> *The experimental systems... have concentrated on deliberately simplified situations... they are unlikely to be able to cope with general collections of images such as newspaper photographs most of which are incapable of any but the most trivial interpretation without the aid of external cues* [19, p. 22].

These comments are echoed by Cawkell who points out that it would be extremely difficult to construct a “query picture” for images portraying, say, “Smog in London” [25, p. 418]. Indeed, he forecasts that progress in automatic recognition techniques will prove rather slow; therefore, for the time being at least, the contents of most image databases will still require to be indexed by words [31].

Conclusions

It has been seen that the proposed solutions to the image indexing and retrieval problem are many and varied; and that each method has its own inherent advantages and disadvantages. Unfortunately, there is a decided lack of unanimity among the information profession as to which is the ideal approach. This problem is exacerbated by the fact that many imaging technology researchers choose to ignore indexing provision completely. As Cawkell observes:

> *It is surprising that more work has not been done on indexing. In fact there appears to be a conscious effort to avoid it. A “negative search” of the literature can be interesting on occasions. The lack of information about image indexing is such a case* [19, p. 9].

This situation is far from ideal, and it is apparent that image retrieval should receive greater attention. In particular, there is a need for a move away from the unco-ordinated production of individualized, stand-alone systems. Instead, the identification and development of internationally accepted image retrieval mechanisms should be a priority. This is particularly significant at a time when there is an overwhelming interest in electronic networks; for as new image databases (each with differing image file formats and retrieval protocols) are added continually to the Internet, the problem of their incompatibility is compounded. Indeed, if the Internet’s potential in providing rapid, convenient access to previously-untapped, remote, visual resources is to be realized, then a common approach to image retrieval really has to be adopted. Certainly, there have been some moves in this direction, but these have almost entirely concentrated on the fine arts. For example, the European Commission's Libraries Programme has funded the Van Eyck (Visual Arts Network for the Exchange of Cultural Knowledge) project which aims to develop standards for the exchange of images and related texts in art photographic databases [32]; while the Getty Art History Information Programme (Ahip) has established a scheme which aims to "address the development of standards to improve access to images of arts and cultural objects over electronic networks" [33].

Bearing this in mind, there is a distinct need for a similar project which will examine image indexing and retrieval standards from a more general, multidisciplinary viewpoint. However, the Internet operates under what might be described as a decentralized environment: no single organization is in overall control. Therefore, if such a scheme is to take place, it will rely heavily on an interested body taking the initiative. In considering whether or not this is likely, it is perhaps worthwhile mentioning the National Science Foundation which, in 1993, took responsibility for the introduction of the Internet Network Information Centre (InterNIC): a project aimed at improving registration services, directory and database services, and information services in the Internet. While it may be somewhat optimistic to expect the NSF to also become involved with image indexing and retrieval standards, it is to be hoped that their example will be emulated; if it is not, then the current, multifarious nature of image retrieval research may continue indefinitely.

References

1. Danziger, P.N., "Picture databases: a practical approach to picture retrieval", *Database*, Vol. 13 No. 4, 1990, pp. 13-17.

2. Besser, H., "Visual access to visual images: the UC Berkeley image database project", *Library Trends*, Vol. 38 No. 4, 1990, pp. 787-98.

3. Shatford Layne, S., "Some issues in the indexing of images", *Journal of the American Society for Information Science*, Vol. 45 No. 8, 1994, pp. 583-8.

4. Sunderland, J., "Image collections: librarians, users and their needs", *Art Libraries Journal*, Vol. 7 No. 2, Summer 1982, pp. 41-9.

5. Markey, K., "Interindexer consistency tests: a literature review and report of a test of consistency in indexing visual materials", *Library and Information Science Research*, Vol. 6 No. 2, 1984, pp. 155-77.

6. Seloff, G.A., "Automated access to the NASA-JSC image archives", *Library Trends*, Vol. 38 No. 4, 1990, pp. 682-96.

7. Enser, P.G.B. and McGregor, C.G., "Analysis of visual information retrieval queries", *British Library R & D Report 6104*, British Library Board, London, 1993.

8. Jul, E., "Howard Besser explores the development of image databases", *OCLC Newsletter*, Vol. 190, March/April 1991, pp. 16-17.

9. Peterson, T., "Developing a new thesaurus for art and architecture", *Library Trends*, Vol. 38 No. 4, 1990, pp. 644-58.

10. Keefe, J.M., "The image as document: descriptive programs at Rensselaer", *Library Trends*, Vol. 38 No. 4, 1990, pp. 659-81.

11. Sherman, C.R., “ICONCLASS: a historical perspective”, *Visual Resources*, Vol. IV No. 3, 1987, pp. 237

12. Roberts, H., “‘Do you have any pictures of .....?’: subject access to works of art in visual collections and book reproductions”, *Art Documentation*, Vol. 7 No. 3, 1988, pp. 87-90.

13. Pearman, S.J., “An opinion: mumblings of a slide librarian”, *Art Documentation*, Vol. 7 No. 4, 1988, pp. 145-46.

14. Cawkell, A.E., “Selected aspects of image processing and management: review and future prospects”, *Journal of Information Science*, Vol. 18 No. 3, 1992, pp. 179-92.

15. Evans, A., “TELCLASS: a structural approach to TV classification”, *Audiovisual Librarian*, Vol. 13 No. 4, 1987, pp. 215-16.

16. Ohlgren, T.H., “Image analysis and indexing in North America: a survey”, *Art Libraries Journal*, Vol. 7 No. 2, 1982, pp. 51-60.

17. Hogan, M., Jorgensen, C. and Jorgensen, P., “The visual thesaurus in a hypermedia environment: a preliminary exploration of conceptual issues and applications”, *International Conference on Hypermedia and Interactivity in Museums, Pittsburgh, October 1991*, Archives and Museum Informatics, Pittsburgh, PA, 1991, pp. 202-21.

18. Noreault, T., Koll, M. and McGill, M.J., “Automatic ranked output from Boolean searches in SIRE”, *Journal of the American Society for Information Science*, Vol. 28 No. 6, 1977, pp. 333-9.

19. Cawkell, A.E., “Indexing collections of electronic images: a review”, *British Library Research Review*, British Library Board, London, 1993, p. 28.

20. Pejterson, A.M., “A new approach to design of document retrieval and indexing systems for OPAC users”, in Raitt, D.I. and Jeapes, B. (Eds), *Online Information 93; Proceedings of the 17th International Online Information Meeting*, London, 7-9 December 1993, Learned Information (Europe), Oxford, 1993, pp. 273-90.

21. Fresko, M. and Brindley, L., *Optical Disc Technology and European Libraries: A Study of User and Technical Requirements*, Bowker-Saur, London, 1992, p. 219.

22. Personal communication with Annelise Mark Pejtersen, December 1994.

23. Enser, P.G.B. and Orr, H.P., “Creative picture research using the Hulton Deutsch CD collection”, *Program*, Vol. 27 No. 4, 1993, pp. 389-98.

24. Philip, G., Crookes, D., Morrow, P.J., Juhasz, Z. and Quinn, A., “Use of a transputer network in the development of an online photographic database”, in Raitt, D.I. (Ed.), *Online Information 91: Proceedings of the 15th International Online Information Meeting*, London, 10-12 December 1991, Learned Information (Europe), Oxford, 1991, pp. 315-26.

25. Cawkell, A.E., “Picture-queries and picture databases”, *Journal of Information Science*, Vol. 19 No. 6, 1993, pp. 409-23.

26. Leung, C.H.C., Hibler, D. and Mwara, N., “Picture retrieval by content description”, *Journal of Information Science*, Vol. 18 No. 2, 1992, pp. 111-19.

27. Bordogna, G. Carrara, P., Gagliardi, I., Merelli, D., Mussio, P., Naldi, F. and Padula, M., “Pictorial indexing for an integrated pictorial and textual environment”, *Journal of Information Science*, Vol. 16 No. 3, 1990, pp. 165-73.

28. Chang, C.-C. and Wu, T.-C., “Retrieving the most similar symbolic pictures from pictorial databases”, *Information Processing and Management*, Vol. 28 No. 5, 1992, pp. 581-8.

29. Aleksander, I., “Modern pattern recognition and the classification of works of art”, *Art Libraries Journal*, Vol. 7 No. 2, 1982, pp. 61-6.

30. Niblack, W. *et al.*, “The QBIC project: querying images by content using colour, texture, and shape”, in Niblack, W. (Ed.), *Storage and Retrieval for Image and Video Databases: Proceedings of the SPIE*, Vol. 1908, San Jose, California, February 1993, SPIE, Bellingham, WA, 1993, pp. 173-87.

31. Cawkell, A.E., “An introduction to image processing and picture management”, *Journal of Document and Text Management*, Vol. 1 No. 1, 1993, pp. 53-63.

32. *Library Association Record*, Vol. 94 No. 11, 1992, p. 692.

33. *Library Association Record*, Vol. 96 No. 10, 1994, p. 538.

Further reading

Agger, S., and Jensen, H., *The Book House: Visual Design*, Riso National Laboratory, Roskilde, 1989.

Barnett, P.J., ”The Art and Architecture Thesaurus as a faceted ARC format”, *Visual Resources*, Vol. IV No. 3, 1987, pp. 247-59.

Batty, D. and Stevens, P., “Automated retrieval systems for photo-image collections: problems and a solution”, in Petrarca, A.E.,Taylor, C.I. and Kohn, R.S. (Eds), *Information Interaction: Proceedings of the 45th ASIS Annual Meeting*, Vol. 19, Columbus, Ohio, 1721 October 1982, Knowledge Industry Publications, White Plains, NY, 1982, pp.23-5.

Cawkell, A.E, “Imaging systems and picture collection management: a review”, *Information Services and Use*, Vol. 12 No. 4, 1992, pp. 301-25.

Couprie, L.D., “ICONCLASS: an iconographic classification system”, *Art Libraries Journal*, Vol. 8 No. 2, 1983, pp. 32-49.

Cox, J. and Taleb, M., “Images on the Internet: enhanced user access”, *Database*, Vol. 17 No. 4, 1994, pp. 18-26.

Dowe, J. III, “Content-based retrieval in multimedia imaging”, in Niblack, W. (Ed.), *Storage and Retrieval for Image and Video Databases: Proceedings of the SPIE*, Vol. 1908, San Jose, California, February 1993, SPIE, Bellingham, WA, 1993, pp.164-7.

Dyson, M.C., “How do you describe a symbol? The problems involved in retrieving symbols from a database”, *Information Services and Use*, Vol. 12 No. 1, 1992, pp. 65-76.

Enser, P.G.B., “Query analysis in a visual information retrieval context”, *Journal of Document and Text Management*, Vol. 1 No. 1, 1993, pp. 25-52.

Gordon, C., “Report of the ICONCLASS workshop, 2-4 November 1987”, *Visual Resources*, Vol. V No. 3, 1988, pp. 205-58.

Harrison, H.P. (Ed.), *Picture Librarianship*, Library Association, London, 1981.

Leung, C.H.C., “Architecture of an image database system”, *Information Services and Use*, Vol. 10 No. 6, 1990, pp. 391-7.

Leung, C.H.C. Mannock, K.L., Mwara, N.K. and Page, S., “Retrieval by content in an image database”, *ITS News*, Vol. 22, August 1990, pp. 41-6.

Lunin, L.F., “The descriptive challenges of fiber art”, *Library Trends*, Vol. 38 No. 4, 1990, pp. 697-716.

Ornager, S., “The image as an information resource”, in Adler, K., Helmer, E. and Holm, H.I. (Eds), *Technology and Competence: Proceedings of the 8th Nordic Conference on Information and Documentation*, Helsingborg, 19-21 May 1992, Tekniska Litteratursallskapet, Stockholm, 1992, pp. 17-21.

Pearson, D. (Ed.), *Image Processing*, McGraw-Hill, Maidenhead, 1991.

Pejterson, A.M., “New model for multimedia interfaces to online public access catalogues”, *The Electronic Library*, Vol. 10 No. 6, 1992, pp. 359-66.

Rorvig, M.E., “Introduction”, *Library Trends*, Vol. 38 No. 4, 1990, pp. 639-43.

Svenonious, E., “Access to nonbook materials: the limits of subject indexing for visual and aural languages”, *Journal of the American Society for Information Science*, Vol. 45 No. 8, 1994, pp. 600-6.

Tibbo, H.R., “Indexing for the humanities”, *Journal of the American Society for Information Science*, Vol. 45 No. 8, 1994, pp. 607-19.

Verguson, C.A., “A library in the media”, *Audiovisual Librarian*, Vol. 12 No. 2, 1986, pp. 70-5.