12-1-1995

Content-based image retrieval and its benefits for the stock photography market

Romano Padeste

Content-Based Image Retrieval and its Benefits for the Stock Photography Market

by

Romano Padeste

A thesis project submitted in partial fulfillment of the requirements for the degree of Master of Science in the School of Printing Management and Sciences in the College of Imaging Arts and Sciences of the Rochester Institute of Technology

December 1995

Thesis Advisor: Professor Frank Cost
Co-advisor: Professor Sabine Süsstrunk

Certificate of Approval

Master's Thesis

This is to certify that the Master's Thesis of

Romano Padeste

With a major in Electronic Publishing
has been approved by the Thesis Committee as satisfactory
for the thesis requirements for the Master of Science degree
at the convocation of

1 February 1996
date

Thesis Committee:

Thesis Advisor

Graduate Program Coordinator

Director

ROCHESTER INSTITUTE OF TECHNOLOGY

This volume is the property of the Institute, but the literary rights of the author must be respected. Please refer to permission statement in this volume for denial or permission, by author, to reproduce. In addition, if the reader obtains any assistance from this volume, he must give proper credit in his own work.

This thesis has been used by the following persons, whose signatures attest to their acceptance of the above restrictions.

Name	Address	Date

Title of thesis: Content-Based Image Retrieval and its Benefits for the Stock Photography Market

December 1995 Romano Padeste

Acknowledgements

The author would like to acknowledge the following individuals for their contribution to the completion of this thesis project.

Professor Sabine Süsstrunk of the School of Photographic Arts and Sciences and Professor Frank Cost of the School of Printing Management and Sciences at RIT for their encouragement and advice throughout this project.

Dr. Franziska Frey of the Image Permanence Institute at RIT for providing valuable information on image processing.

Ms. Erin Hickey and Ms. Cindy McCombe for proof reading.

On a more personal note, the author would like to thank his family for their moral and financial support. Finally, the author would like to extend his thanks to his friends here in the United States and in Switzerland; without their encouragement, the completion of this thesis project would not have been possible.

Table of Contents

Page

List of Figures

Abstract

The development of powerful low-cost desktop computer systems has changed the pre-press business — where tight deadlines must be met — persistently. An increasing number of newspapers and magazines are acquiring, handling, and storing images digitally while the use of hardcopies and slides decreases.

Today's computers and high capacity storage-media enable stock photography agencies to build digital image databases, giving users fast access to large numbers of images. However, the transition from analog to digital image archives imposes new problems: with thousands of images at hand, the search for a particular image may turn into the search for the needle in a haystack.

The first image Database Management Systems (DBMSs) were extended text DBMSs, which stored the image data along with a set of manually entered descriptive keywords. The major problem with this approach is that there is no generally agreed-upon language to describe images. Even sophisticated DBMSs are unable to detect synonyms; hence, an image described with certain properties such as "curvy" may not be found if a user enters

"wavy" as a search criterion. Furthermore, some image properties are hard to describe with keywords. A search is likely to fail if properties were not described at the database population stage when images are added to the database. Finally, assigning a sufficient set of keywords to every image adds a tremendous amount of labor to the population stage.

Research at many scientific institutions and companies is geared towards overcoming the shortcomings of image DBMSs with keyword-based search engines. Pattern recognition — which allows for comparing images based on their *visual* content — is being introduced to image DBMSs, improving the accuracy of search engines. Sketches, sample images, and other means of describing the visual content of images may be used as search criteria in addition to keywords. This thesis project summarizes the basics of pattern recognition and its applications in image database management for content-based image retrieval.

The purpose of this thesis project is to determine the impact of content-based image retrieval on the stock photography market in the near future. In order to obtain the necessary information, two different questionnaires were sent out to a number of selected stock photography agencies, newspapers, and magazines. The evaluation of the replies was conducted for the three groups separately.

The replies from stock photography agencies showed a high interest in digital image archives. They also showed concerns about increased overhead with digital archives. The estimated amount of work required for categorizing images and assigning keywords ranged from fifty to ninety percent as compared to ten to fifty percent for scanning. All survey participants agreed that pattern recognition can improve the accuracy of keyword-based search engines. However, they all denied that this approach would reduce the need for assigning keywords.

Different needs could be determined for newspaper and magazines. Newspapers rely heavily on keywords since images are often chosen based upon the circumstances under which they were taken while their visual content may be secondary. Therefore, newspapers' profits from content-based image retrieval are minute. For magazines, the visual content of images seemed to have a higher priority and the appreciation for corresponding search capabilities was accordingly higher.

To summarize, users of digital image archives can profit from content-based image retrieval if the visual content is an important issue. For image providers, there are a number of reasons that delay the transition to content-based image retrieval. Currently, there is only one shrink-wrapped commercial product available that meets the needs of stock photography agencies. This product requires additional work for fully exhausting its capabilities. Finally, many companies have already built their image database and the transition to another system is time-consuming, expensive, and risky.

Chapter 1

Introduction

Statement of the Problem

Stock photography agencies which maintain large image archives have started to digitize their material in order to allow more efficient and faster access to their images. Most of them provide their catalogs on CD-ROM while others have even started to offer their customers on-line access to a digital image database. The major advantage of the on-line database is that updates can be made continuously, and the customers have immediate access to those updates. This is especially important for agencies which concentrate on the newspaper and news magazine market. Besides the simplified update procedures, on-line catalogs permit instant access to thousands of images. Nevertheless, retrieving an image may still be time consuming since most of the search engines are based on keywords. This thesis project analyzes existing stock photography databases from both the customer's and agency's point-of-view. It concentrates on the speed and accuracy of image retrieval based on keyword searches and evaluates further methods, i.e. pattern recognition techniques, to improve queries.

Background and Significance

Depending on the type of publication, an image editor will use stock photos more or less extensively. Image editors for newspapers will benefit from on-line access to image databases in terms of meeting deadlines. They will be able to put the latest images of international events in the paper if they have the possibility to acquire them on-line. Since these pictures are related to certain events, they may be found easily with queries that are based on keywords.

In the case of news and special interest magazines, a large variety of images will be made available to them from stock photography agencies. The images may be related to news as well as to background information and editorial articles. Layout and editorial considerations will have an impact on the selection of particular images. Therefore, it may be more difficult to find a satisfying result with language-based queries. Image editors for magazines often rely on professional picture researchers who work for stock photography agencies to find particular images for their customers. Using CD-ROMs and on-line databases, photography editors are able to do this research themselves.

Multimedia is an emerging segment of the publishing industry that has image acquisition needs similar to the magazine market. While multimedia publishers do not need high resolution images, they may request an even larger variety of images than magazine publishers. In addition, multimedia applications often use motion picture as well as still images, which may compound the difficulties of building an accurate query system. For reasons of simplicity, neither the multimedia nor the book publishing market were included in this survey.

Although this thesis project discusses the problem of image retrieval in large image databases focusing on stock photography, the results can likely be applied to other types of image databases. Besides commercial stock photography agencies, museums and scientific institutions with vast image archives will profit as well from the developments outlined in Chapter 3.

Reasons for Interest

After studying scientific photography, the author worked for three years as a photographer in a small family-owned company. One of his responsibilities was to maintain the company's photography-archive. This archive had two parts: a recent part with all the negatives and slides of the last twenty years, and an older part, where the material from the beginning of this century up until the seventies was stored. In the old archive, about ten thousand envelopes and boxes of documentary material were stored. During the three years of his employment, the author had a number of telephone calls from people who were looking for old pictures of the town where the company is located. Since the only information about the company's archive was a list containing the number of the envelope, the name of the customer, the date when the pictures were taken, and the price, it was quite a challenge to retrieve one particular image. There were two basic options: to ask the company's Chief Executive Officer whether he remembered a picture which would meet the requests of the customer, or to search through hundreds of envelopes in hopes that a corresponding picture could be found by accident. Even if the search was successful, the price for a copy of the image would not pay but a fraction of what the work power for retrieving the image cost the company. Thus, the maintenance of this archive was a loss for the company and many of the stored images were buried like treasures.

The author had a similar experience with his own photographic work that he did for this company. Eventually he could sell a previously made picture to a second customer due to the fact that he remembered this particular picture. He attempted to suggest that the information about stored images should be entered in a computer database but he realized that this would improve the probability to retrieve an image just in some cases. In many other cases it would not have made a difference because it is impossible to predict further use for an image. This is the inherent weakness of all language-based search engines for image databases, and, to the author's knowledge, there were no other types of database systems available at that time.

The confrontation with modern stock photography agencies and their huge image archives raised the following question again: how would one possibly retrieve one particular image without having exact data about the picture. From a photography editor's point-of-view, an image database should allow users to search for a picture in a more associative way based on a related article or guidelines given by the layout. Shape recognition seemed to be a feasible way to enhance language-based search engines and, as a matter of fact, IBM tried this with their newly released QBIC (Query By Image Content) search engine.[1] An evaluation of customers' experiences with existing systems will enable the development of suitable solutions for the future. Eventually, this technology will also be available for photographers to help them maintain their archives within a reasonable time frame.

Endnotes

Seybold Publications. *The Seybold Report on Desktop Publishing.* Vol. 9, No. 1 (September 12, 1994): 34

Chapter 2

Theoretical Basis

Pattern Recognition and Shape Recognition

Pattern recognition is a general term used in computer sciences for applications that compare a given organized system of information (referred to as object) to another system of information. Pattern recognition can be applied to any kind of digital data, such as written language, sound, or images. In image processing, for example, pattern recognition can be used to determine whether a part of an image has a particular texture. In order to enable the computer to recognize organized systems of information, these systems (or objects) must be mathematically described, most commonly as vectors or arrays of numbers. In addition, a set of categories or classes, such as "circle" or "square", have to be predefined. The process of pattern recognition is considered successful when a sample pattern can be associated with a class. Since a system may not perfectly fit in a class, most pattern recognition applications provide an estimate of confidence.[1]

Most pattern recognition systems need to be trained before they work efficiently. The process of pattern recognition requires knowledge about specific patterns and classifications. Thus, these systems are referred to as knowledge-based systems. The training process can either take place automatically — i.e., the system trains itself — or under the supervision of a human teacher.[2]

Shape recognition is a subclass of pattern recognition and is used to detect the geometric shapes of objects in an image. It is one of the approaches used to classify and subsequently *understand* the image. It is, however, used in combination with other techniques of pattern recognition.

Applications of Pattern Recognition

Early research on pattern recognition was done for military applications, such as target detection, and remote sensing. Today, the applications for pattern recognition are as varied as the use of computers. Pattern-recognition sub-routines can be used in word processing programs for spelling and grammar checking. In image processing programs, some data compression algorithms are based on pattern recognition techniques. Furthermore, pattern recognition is used for image manipulation and optical character recognition (OCR). Medical researchers use pattern recognition to analyze x-ray images and enhance the human diagnosis. Various industrial applications also exist, such as automated manufacturing and quality control.[3] Video cameras which are connected to pattern recognition systems enable robots to fulfill extremely complex tasks and to help prevent accidents caused by human interference.

Many applications were developed using specialized algorithms to fulfill specific needs. As a result, there are various approaches with different advantages and disadvantages. However, the increasing power and speed of computers enhanced the research on general purpose algorithms in the last few years.

Pattern Recognition in Image Processing

Today's sophisticated image processing software packages, such as Adobe Photoshop™, provide many filter functions with underlying pattern detection capabilities.

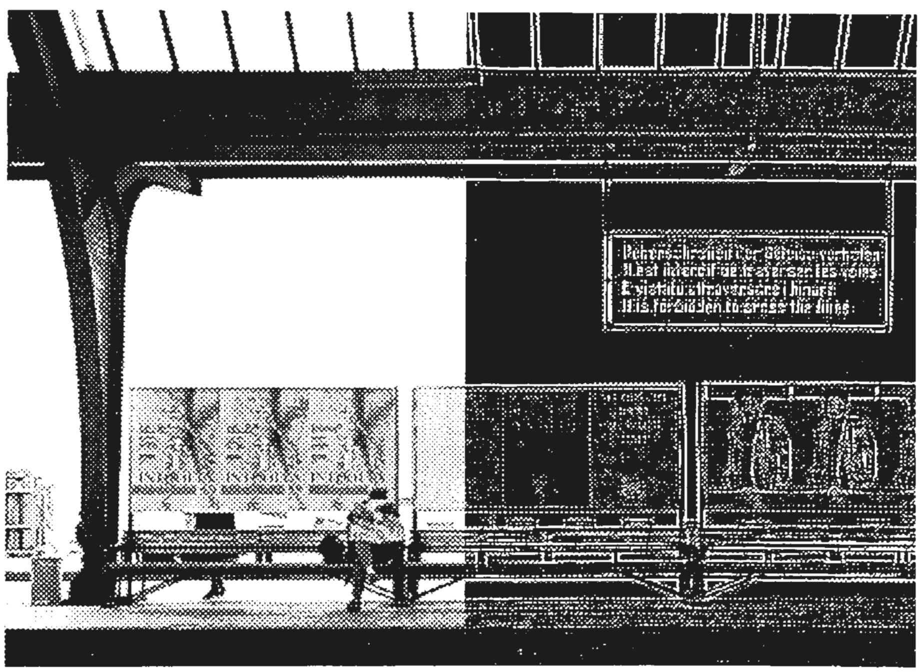

Figure 1 - Edge Detection Filter

Edge detection, which was applied to the right-hand half of this grayscale image, is one of the basic image processing techniques used to extract particular features of an image.

Photoshop's "Find Edges" filter, for example, has the capability to extract edges within an image. Also, some compression schemes for image data compression are based on pattern-recognition algorithms to detect redundant information that can be discarded when storing the image.[4]

Depending on the purpose of the software, various pattern recognition sub-routines may be used at different levels of image processing. In general, image processing can be sub-divided into low-, medium-, and high level image processing. Low-level processing involves image sensing and preprocessing. At this stage, pattern recognition algorithms are used to perform operations such as the previously mentioned edge detection function.[5]

Medium-level image processing involves image segmentation and description of the image data. In the segmentation process, a variety of criteria are used to classify each pixel based on some of its properties.

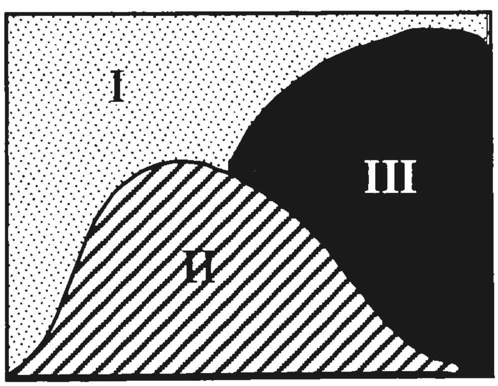

Figure 2 - Image Segmentation

Using image segmentation, three different areas have been detected based on textural differences in this sample image.

Segmentation can be based on gray-level thresholding, region growing and splitting, texture description, and other operations. The description of the image data can be subdivided in feature and shape description. These steps require powerful computers, preferably with parallel processing capabilities, which limited the applications to large computer systems for many years.[6] However, this restriction no longer holds true with the recent introduction of RISC (Reduced Instruction Set Computing) technology to the desktop computer market.

Finally, high-level image processing deals with the recognition of objects and the interpretation of the scene. At the recognition stage, the different objects of the image are classified. Depending on the application, this classification can be a very complex task. The final step of interpretation is strongly related to knowledge-based systems and artificial intelligence (AI).

Pattern Recognition in Image Classification and Understanding

A short description of the steps involved in human vision and image interpretation is necessary to understand the complexity of high-level image processing. These steps are the optical projection of the reality in the eye, the recording and encoding by the retina, and the interpretation of the data by the brain.

> The key problem for perception is how meaning is read from neural signals from the senses. Vision presents problems of special interest, as especially for man visual perception is so

> rich, and yet — unlike the senses of touch, taste and smell — the optical images in eyes have no intrinsic biological significance, for they cannot directly monitor the presence of food or poison, or hot or cold. Organisms cannot eat images, or be eaten by them: to be useful they must, somehow, convey properties of objects which are not present in the image, which physically is but shadows or patterns of light. Though we see that a table is hard and a glass brittle, the eye's images are not hard or brittle. Somehow a wealth of such non-optical characteristics of objects are read from neural signals from retinal images — no doubt through inherited and learned interactive experience with objects and what they do in various situations.[7]

In his article "How do we interpret images?"[8] Richard Gregory describes two main approaches to human image interpretation. The "passivistic" theory proposes that meaning is in the world itself and that this meaning is to be picked up by an observer who only selects what meaning he or she needs. The "activistic" theory, on the other hand, proposes that the observer creates rich meanings from limited and only indirectly related data available from the senses. This assumes that the observer has knowledge about the objects.

Gregory then introduces an *algovistic* theory that combines active and passive concepts. According to this algovistic theory, there is a distinction between two different levels of knowledge. Low-level knowledge consists of general rules, such as the likelihood of objects having particular shapes. For example, it is not very likely that a house is round. These general rules can be applied to many different objects and allow the observer to deal with objects that are partly hidden. Gregory demonstrates this with a number of illusory figures where the human perception adds missing contours and shapes to objects based on the likelihood of their shapes. Low-level knowledge applies to perception and may occur in the retina before the signals are transmitted to the brain. High-level knowledge enables conception (or understanding) because it is *specific* knowledge about *particular* objects. Gregory emphazises this division between perception and conception by referring to the illusory figures: even though human observers know that they are looking at illusory figures (conception), they cannot avoid to "see" the illusions (perception.)

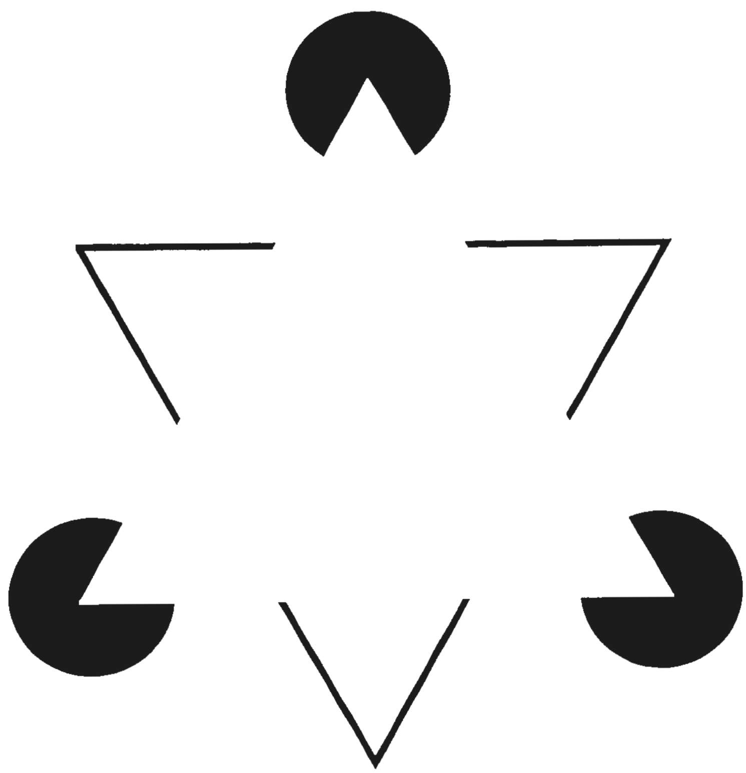

Figure 3 - Illusory Figure[9]

In this example, an illusory white triangle seems to cover parts of three black circles and a white triangle with a black outline.

Knowledge-based Computer Vision systems are designed similarly to the human vision system, as described by the algovistic theory. General rules, such as general geometric shapes, are applied to the image to isolate the different objects of a scene. After this perceptional step, more specific rules are necessary to classify objects, analyze the relationship between different objects, and finally understand a scene. Depending on the application, the understanding of the scene requires a more or less extensive knowledge base. The complexity of natural scenes, as they occur in image databases of stock photography agencies, has placed a considerable constraint on the application of pattern recognition techniques for image retrieval for many years. However, recent developments in computer technology have increased the viability of this approach as well as the demand for it. For content-based image retrieval, various techniques of low-, medium- and high-level image processing may be applied (see Chapter 3.)

Endnotes

[1] Rothman, Peter, Syntactic Pattern Recognition: Tutorial. *AI Expert*, Vol. 7, No. 10 (October 1992): 41 - 42

[2] Ibid.

[3] Hussain, Zahid, *Digital Image Processing: Practical Applications of Parallel Processing Techniques*. Ellis Horwood Limited, Chichester, UK, 1991: 228 - 335

[4] Fisher, Yuval and Lawrence, Albert F. "Fractal Image Compression for Mass Storage Applications," in *Image Storage and Retrieval Systems: Proceedings of The SPIE (The International Society for Optical Engineering) Congress Held in San Jose, CA , 13 - 14 February, 1992.* SPIE Volume 1662, 1992: 244

[5] Hussain 1991: 73 - 111

[6] Hussain 1991: 113 - 141

[7] Gregory, Richard, "How do we interpret images?" in *Images And Understanding*, Barlow, Horace, Blakemore, Collin, Weston-Smith, Miranda ed, Cambridge University Press, Cambridge, UK, 1990: 310

[8] Ibid.: 310 - 332

[9] Ibid.: 318

Chapter 3

Review of the Literature

Early applications of pattern recognition usually targeted very specific tasks with relatively limited variance of objects in an image. These specialized applications used specific techniques that met the requirements of the task. Consequently, different approaches were merged as the applications and their corresponding requirements became more general. As a result, a vast variety of slightly different approaches have been suggested and tested as well. The discussion of the literature related to this thesis project will focus on SPIE (The Society of Photo-Optical Instrumentation Engineers) Proceedings of the following congresses:

1. Image Storage and Retrieval Systems, February 13 - 14, 1992, San Jose, CA[1]
2. Storage and Retrieval for Image and Video Databases, February 2 - 3, 1993, San Jose, CA[2]
3. Storage and Retrieval for Image and Video Databases II, February 7 - 8, 1994, San Jose, CA[3]
4. Storage and Retrieval for Image and Video Databases III, February 9 - 10, 1995, San Jose, CA[4]

These proceedings document the tremendous efforts that were made over the past few years to improve image retrieval in large databases. The articles provide detailed information on different approaches to image indexing, content-based image retrieval, database architecture and management, mass storage for image databases, and digital watermarks. Furthermore, many articles deal with indexing and retrieval of digital video sequences. Due to the complexity of the subject, this review covers mostly the articles which describe general purpose and commercial implementations.

Image Retrieval Methods

At the SPIE congress on Storage and Retrieval for Image and Video Databases in 1993, IBM presented results of the QBIC (Query By Image Content) research project.[5] The project explores methods of retrieving images based on their visual appearance in order to enhance keyword-based queries. In the introduction, the authors point out the following problems of queries that are solely based on keywords:

> First, . . . if the current query refers to image properties that were not initially described, the search will most likely fail. Second, some visual properties are difficult or nearly impossible to describe with text such as certain textures or shapes. . . . Thirdly, even if all useful characteristics of an image are described with text, there is no commonly agreed-upon vocabulary for describing image properties, so that a "curvy" item may not match a "wavy" one.[6]

The QBIC project was designed for large image databases where browsing through the thumbnails of the entire database would be too time consuming.

Based on the assumption that users may be looking for images which have a visual appearance similar to some given image, query methods that use sample images as query criteria were developed. Additionally, query methods that use sketches, user-constructed query images, color and texture patterns, layout or structural description, and other image information were implemented as well. Due to this approach, approximate rather than exact matches are returned as search results. Therefore, an interactive application was designed which allows the user to visually evaluate the search results and, if necessary, refine the query.

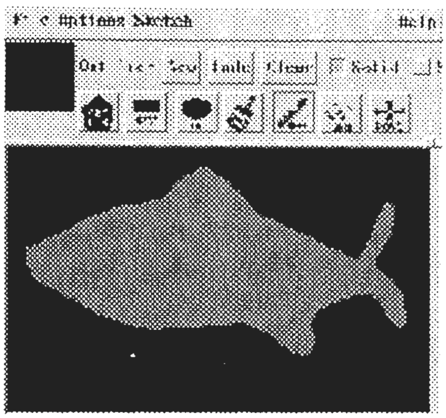

Figure 4 - The QBIC Sketch Drawing Window[7]

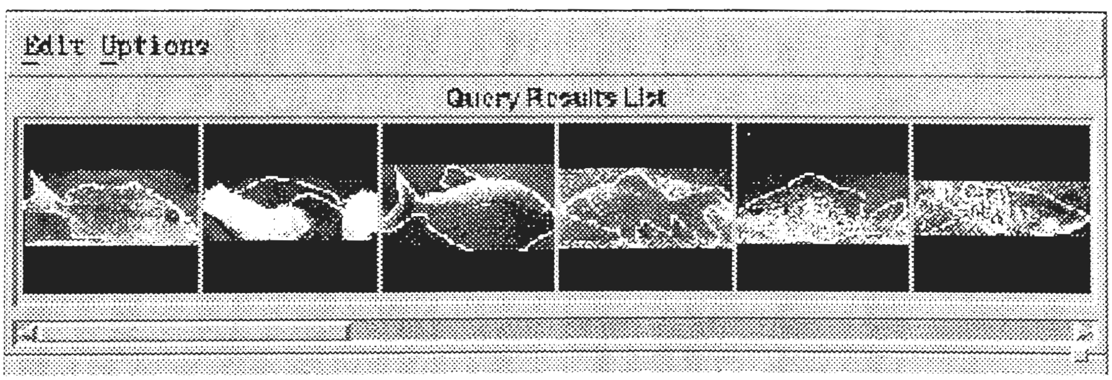

Figure 5 - QBIC Search Results[8]

These images were returned as results of the query displayed in Figure 4.

The QBIC application involves the following three logical steps: database population, feature calculation, and image query. First, the images are added to the database along with thumbnails and any available text information about the images. A second — optional — step allows the user to interactively or manually identify important objects within the images. This can be done using a "snake" tool to outline the object contour. An approximate outline drawn by the user is iteratively adjusted to the actual object outline by the software application. Another option for object identification is the "flood fill" where the user selects one pixel and the application adds all pixels of a similar color within a given threshold. For relatively uniform objects with a good distinction from the background, this method is more suitable than manual outlining. Additional text information can be entered to describe individual objects that were selected using either of the two selection methods.

Following the population step, a number of features are calculated for each image added to the database. Color features are calculated based on the average color coordinates of the image as well as each identified object within the image. Texture-feature calculation is based on criteria such as coarseness, contrast, and directionality. Shape features are calculated using the outlines of identified objects. Currently, it is assumed that all shapes are planar and non-occluded. Finally, sketch features are calculated using an edge detection operator and reducing the binary edge image to a size of sixty-four by sixty-four pixels.

After the feature calculation, queries can be performed using a number of different criteria. QBIC allows users to search for images based on all the previously calculated features. To submit a query, colors and textures can be selected using the corresponding picker window.

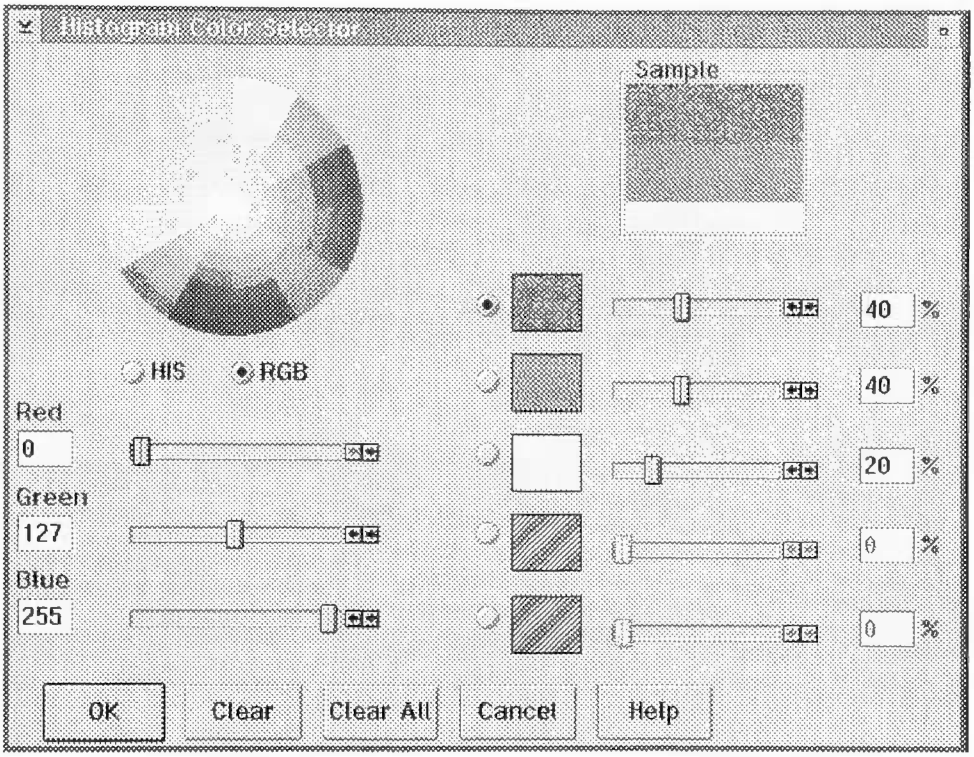

Figure 6 - The QBIC Histogram Color Selector Window[9]

A drawing area allows users to draw a raw sketch with dominant lines (see Figure 4, page 18.) In the Histogram Color Selector window, different colors can be selected along with a percentage to search for images with a corresponding overall color distribution. Selected shapes of a sample image can be used as a query criteria as well. Finally, the criteria can be combined and weighted.

In 1995, IBM presented an enhanced version of QBIC.[10] In addition to the features of the earlier version, it offers a color painting window where users can construct a color layout for a region-based color query.

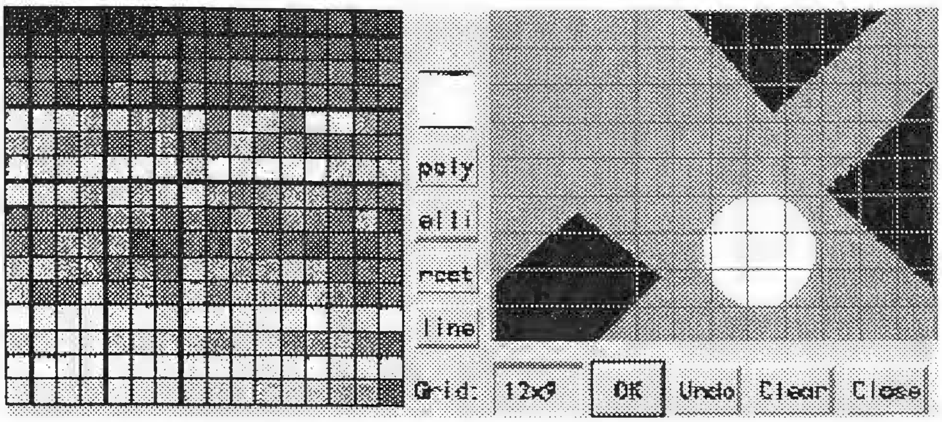

Figure 7 - The QBIC Color Painting Window[11]

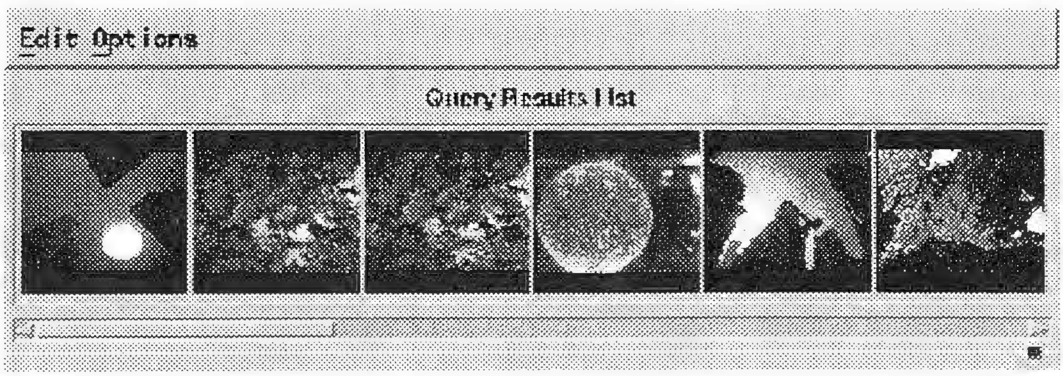

Figure 8 - QBIC Search Results[12]

These images were returned as results of the query displayed in Figure 7.

At the same congress, IBM's Ultimedia Manager — a commercial implementation of the QBIC project — was presented.[13] Ultimedia Manager consists of an Image Classifier and an Image Query component. The two components offer all the previously mentioned features of the QBIC application.

The Classifier component is used for the data entry and feature calculation while the Query component is used to perform a search. Ultimedia Manager also offers a function to merge catalogs which were individually built by several people. The software is available as an OS/2 client application in a LAN (Local Area Network) or as an OS/2 stand-alone application.

A different approach to content-based image retrieval was made by Storm Software. The company designed Apple PhotoFlash and EasyPhoto for Windows platforms as "shrink-wrapped end-user software . . . street priced from $50-$100."[14] Further important design decisions were that the software could run on personal computers with 80486 or 68040 processors and that users would typically work with image collections of one hundred to a thousand images. The catalog file created by the software contains information about the images, such as keywords and file name, and a thumbnail of eighty-by-eighty pixels. The images are divided into sixty-four (eight-by-eight) rectangular regions, or cells, of identical size and classified by calculating the average color of the pixels within each cell. The applications allow users to search for images using the thumbnail of a sample image or by drawing a simple sketch in a drawing area of eight-by-eight cells. The two products are clearly designed for small office or home use and cannot meet the needs of stock photography agencies with tens of thousands of images.

Researchers at the Perceptual Computing Section of the Media Laboratory at the Massachusetts Institute of Technology (MIT) developed the Photobook system for content-based image retrieval. One of its most remarkable features is its capability to find images of people in a face database even if the appearance of the persons is altered with faked beards or hair-style changes. Promising results have been achieved with more general image databases as well.[15]

Image Storage and Data Compression

For many years, storage space has been one of the major concerns when dealing with digital images. Images which are used for high quality reproduction quickly take up megabytes of storage space. Advances in image-data compression and storage-media development have helped to reduce this concern. Today, magnetic, optical, and magneto-optical disks with capacities of up to nine gigabytes per disk are the media of choice for on-line digital image databases. Optical media offer a greater lifetime of the stored data than magnetic media and therefore are preferred for archiving purposes.

Another important concern is the accessibility of the stored data. The data access in a LAN is affected by the bandwidth of the communication lines, disk access time, and data throughput. An image server for a LAN environment must be optimized for multiple simultaneous access to large files, called Binary Large Objects (BLOBs). A parallel file system in a Multi-Processor-Multi-Disk (MPMD) environment has been proposed to meet these needs.[16] In the MPMD architecture, a number of disk drives are connected to a corresponding number of interconnected disk-node managing processors. The disk-node managing processors are then connected to the server-interface processor. At saving time, the images are partitioned into a number of extents which are distributed sequentially and as uniformly as possible among the available disk nodes. This approach speeds up the file transfer from the file server to the client significantly.

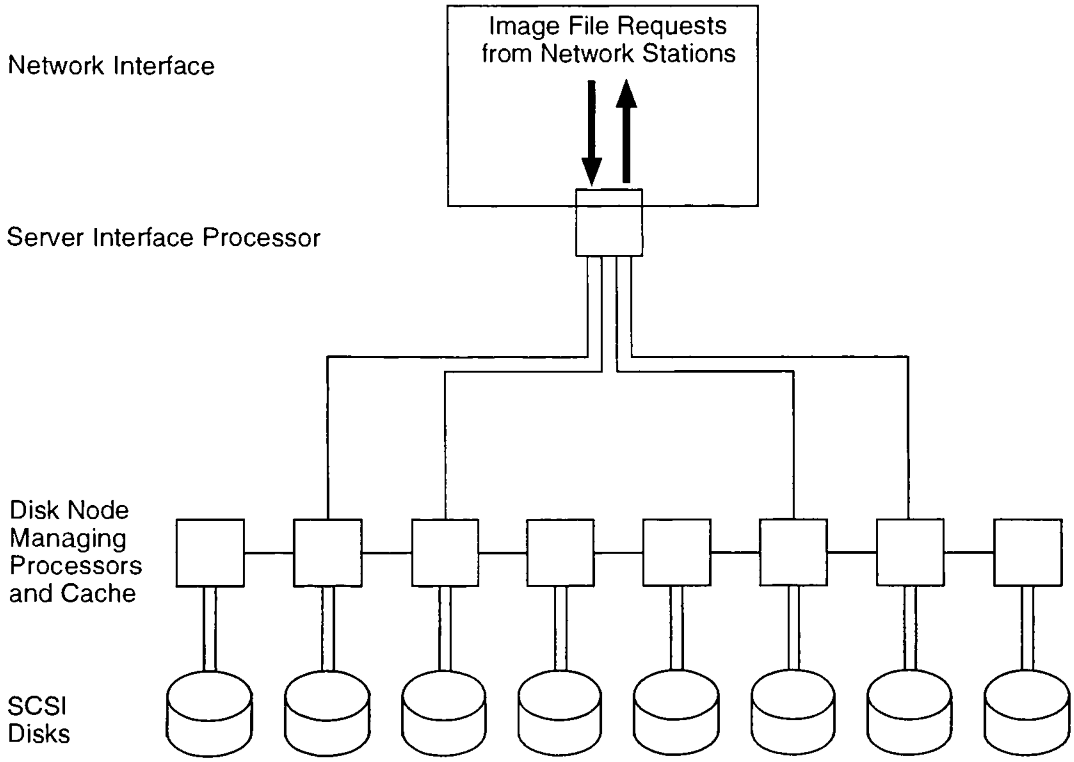

Figure 9 - Illustration of the MPMD Architecture[17]

Data retrieval requests in a client/server environment are usually handled on a "First Come First Serve" basis meaning that all other requests have to wait until the first request is completed. This approach may cause unnecessary delays if the data is stored in near-line devices, such as an optical jukebox. The system may use excessive time to switch disks, reducing the performance of the system. A cyclic request scheduling scheme that completes all requests for one disk before switching to the next disk shows an increased performance resulting in overall shorter waiting times for the individual users.[18] It has to be pointed out that this improvement applies to near-line systems where the time needed for switching disks is typically greater than data transfer times.

Data compression is a popular method to decrease storage space requirement and consequently increase the system's performance in terms of data access times. The Joint Photographic Expert Group's JPEG compression scheme has become one of the most popular compressed image file formats. However, JPEG is a lossy compression scheme which can cause visible artifacts,

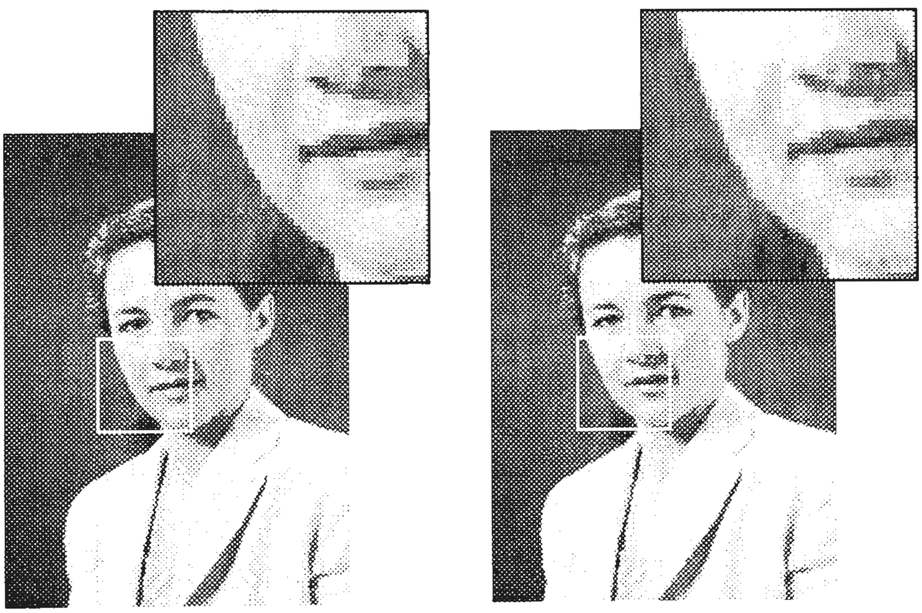

Figure 10 - Artifacts Caused by JPEG Compression

The image stored using JPEG compression (R) looks different than the uncompressed original TIFF file (L). The artifacts caused by JPEG compression are clearly visible in the enlarged detail (enlargement: 300%.)

such as banding, in the decompressed image (see Figure 10, page 25.) The results of discarding image information become especially obvious with high compression ratios and in low contrast image areas. The image deterioration caused by the compression scheme limits the reproduction scale for the final output size. Nevertheless, JPEG is a robust and reliable compression scheme which achieves reasonable image quality with relatively high compression ratios if used carefully.[19]

Lempel-Ziv-Welch (LZW) compression is a popular lossless algorithm which retains the original image information. The decompressed image consists of the exact same information as the original and no artifacts occur during the compression-decompression cycle. The drawback for preserving the original information is that the compression ratios achieved with LZW are lower than those with JPEG.

Advances in image processing and pattern recognition will lead to better compression schemes in the future. Wavelet Coding and fractal compression are two promising approaches to improved image compression which are currently being researched.[20] However, the introduction of a new standard may be problematic since it adds another file format to the already large variety of graphics file formats.

Image Database Management

Database applications were originally designed for highly structured alphanumeric and numeric information. With the advent of multimedia the applications had to be redesigned to be able to deal with BLOBs , such as image files. Today, most image database applications allow users to build a relational catalog file, containing thumbnails of the images along with additional information, such as keywords, date and other production information, and the location of the original image files. In some cases, the catalog file also contains information for image content-based search engines. Keyword searches are based on the Structured Query Language (SQL) of traditional text databases.

Due to the large storage space requirements for images, large image databases are usually distributed over a LAN. Frequently used images may be stored in on-line storage devices, typically magnetic drives, while less frequently used images may reside in near-line storage devices, such as optical jukeboxes. Newly added images may first reside on an image-input server and later moved to a different on-line server or near-line storage device. Since the catalog files contain pointers to the original image files, the Database Management System (DBMS) must be able to keep track of the images over the whole LAN. This problem has been addressed in the Mass Storage Reference Model of the IEEE Computer Society.[21] The model suggests a layered storage setup with on-line and near-line devices. The client applications communicate directly with the Bitfile Server. An application called Bitfile Mover moves files from near-line to on-line storage upon request. The same application also moves files from on-line to near-line devices if the on-line storage is full. The Bitfile Server keeps track of the files at any stage of a placement change.

Besides keeping track of the files, the DBMS must provide a number of other functions. It must allow individual users to make temporary copies of files which reside on the harddrive of a user workstation. This copy function helps to decrease access times for individual users.[22] The DBMS must keep track of different versions of an image and prevent unintended overwriting of original image data with manipulated data. Furthermore, the DBMS must provide a tight security system with different levels of permissions for individual users. Finally, it must provide backup and restore functions in case of system crashes or storage hardware failures.

Data Security Issues

Copyright has always been a big concern of photographers and stock photography agencies. Digital imaging — which allows users to easily copy, manipulate, and distribute images — has increased this concern significantly. This is one of the reasons why many agencies still shy away from getting involved in digital imaging. Currently, a number of different approaches to improved copyright protection are being explored. The most simple approach is to add a caption with a copyright note to the image data. The problem with this approach is that the text can easily be removed, therefore not offering good protection against intended copyright violations.

Data hiding is a term used for techniques which embed data, e.g., copyright information, into image data. "The primary purposes for data hiding in digital media are to provide solid proof of the copyright and assurance of content integrity. Therefore, the data should stay hidden in the host, even if the host signal is subject to manipulation, such as filtering, re-sampling, cropping or data compression."[23] Digital Watermarks for copyright proof can be visible alterations of the original image that display the copyright information. With this type of watermarks, the image can only be used for positioning purpose.

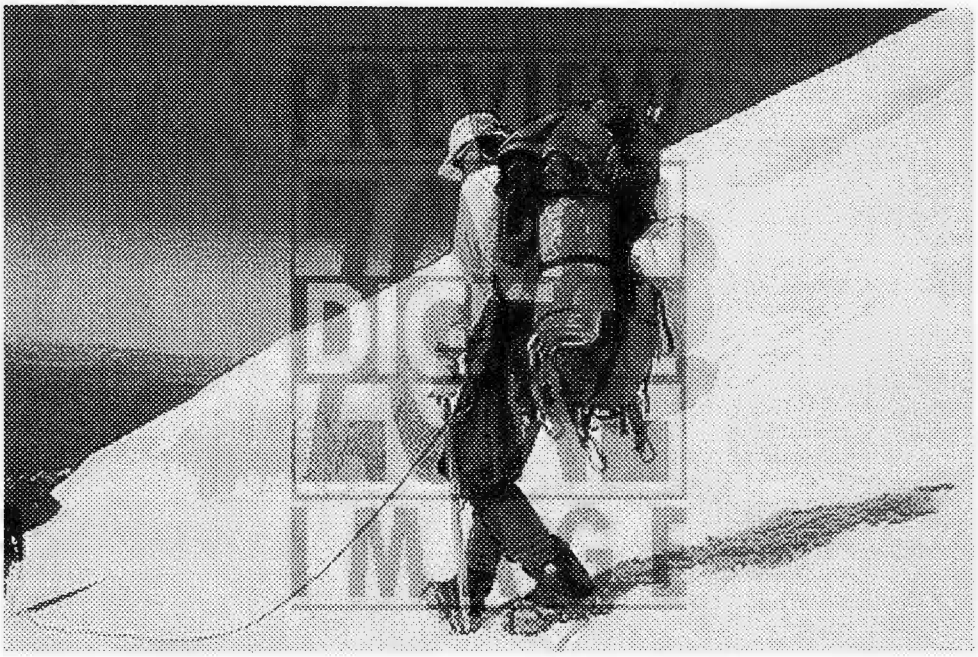

Figure 11 - Image Protected by a Visible Watermark [24]
Photography: Kevin Morris, Copyright: Digital Zone Inc.

If the original image has to be protected, the copyright information must be invisibly hidden in the image data. The copyright data can be added to the header of the image file or to the image data. Both methods have their disadvantages: information in the file header may easily be removed while information added to the image data may be affected or lost due to any type of image manipulation.

Tamper Proofing algorithms are used to detect whether or not an image has significantly been manipulated, e. g., an important item has been removed or added. However, the algorithms should not be triggered by small manipulations, such as gamma corrections or cropping.

Various procedures for data hiding have been suggested and tested in the last few years. Although successful in some areas, all techniques still have limitations, especially when image compression algorithms are applied after the data hiding process, and the goal of protecting hidden data against intended removal may be unachievable.[25] Furthermore, there is currently no information available regarding the impact of data hiding on content-based image retrieval.

Endnotes

[1] *Image Storage and Retrieval Systems: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 13 - 14 February, 1992*. SPIE Volume 1662, 1992

[2] *Storage and Retrieval for Image and Video Databases: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 2 - 3 February, 1993*. SPIE Volume 1908, 1993

[3] *Storage and Retrieval for Image and Video Databases II: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 7 - 8 February, 1994*. SPIE Volume 2185, 1994

[4] *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995*. SPIE Volume 2420, 1995

[5] Niblack, W., et al.,"The QBIC Project: Querying Images By Content Using Color, Texture, and Shape," in *Storage and Retrieval for Image and Video Databases : Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 2 - 3 February, 1993*. SPIE Volume 1908, 1993: 173 - 187

[6] Ibid.: 173

[7] *Query By Image Content (QBIC)*, [Online], Available http://www-i.almaden.ibm.com/cs/showtell/qbic/Initpage.html

[8] Ibid.

[9] *IBM Ultimedia Manager 1.1 and Client Search,,* [Online], Available http://www.software.ibm.com/data/umm/umm.html

[10] Ashley, J., et al., "Automatic and Semi-Automatic Methods for Image Annotation and Retrieval in QBIC," in *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995.* SPIE Volume 2420, 1995: 24 - 35

[11] *Query By Image Content (QBIC),* [Online], Available http://www-i.almaden.ibm.com/cs/showtell/qbic/Initpage.html

[12] Ibid.

[13] Treat, H., et al., "Searching Images Using Ultimedia Manager," in *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995.* SPIE Volume 2420, 1995: 204 - 213

[14] Equitz, Will, "Image Searching in a Shipping Product," in *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995.* SPIE Volume 2420, 1995: 187

[15] Pentland, A., Picard, R. W., Sclaroff, S., "Photobook: Tools for Content-Based Manipulation of Image Databases," in *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995.* SPIE Volume 2420, 1995: 34 - 47

[16] Hersch, Roger D., "Benefits of An Image Oriented Parallel File System" in *Storage and Retrieval for Image and Video Databases: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 2 - 3 February, 1993.* SPIE Volume 1908, 1993: 116 - 122

[17] Ibid.: 117

[18] Mantey, Patrick E., Levy, David E., "Electronic Libraries and Optical Jukebox Server Scheduling" in *Image Storage and Retrieval Systems:*

Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 13 - 14 February, 1992. SPIE Volume 1662, 1992: 146 - 153

[19] Rabbani, Majid, "Image Compression Fundamentals" in *The Compression Experience: Proceedings of The Rochester Section of The SMPTE (Society of Motion Picture and Television Engineers) Tutorial Held in Rochester, NY, October 28, 1995*: 7 - 24

[20] Ibid.: 25 - 26

[21] Franchi, Santina, Guarda, Roberto, Prampotini, Franco, "Mass Storage Management for Distributed Image/Video Archives" in *Storage and Retrieval for Image and Video Databases: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 2 - 3 February, 1993.* SPIE Volume 1908, 1993: 153

[22] Ibid.: 159

[23] Bender, Walter, Gruhl, Daniel, Morimoto, Norishige, "Techniques for Data Hiding" in *Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995.* SPIE Volume 2420, 1995: 164

[24] *Kevin Morris PhotoCD Sampler*, [Online], Available http://www.digitalzone.com/cdrom/km/km1.jpeg

[25] Bender, Walter, Gruhl, Daniel, Morimoto, Norishige 1995: 173

Chapter 4

Statement of Project Goals

The main purpose of this thesis project was to determine whether or not the stock photography market could benefit from the introduction of pattern-recognition technology to image Database Management Systems (DBMSs.) The basics of this technology and its implementation in search engines for image databases are described in Chapter 2 and 3. The research on providers and users of stock photography, as outlined in Chapter 5, was designed to prove the following statements.

> Currently, digital stock photography catalogs lack intuition in their search capability due to their language-based search engines. Pattern-recognition technology, which enables users to search for an image based on its visual content, is likely to overcome this deficit when combined with traditional search engines. However, other issues, such as data compatibility and copyright security, will need to be considered before this technology can be implemented in commercial image databases.

Chapter 5

Methodology

The analytical part of the thesis focuses on both providers and users of stock photography. Inquiries and interviews were used to collect the data needed to describe the current situation and future developments in the stock photography market.

Survey on Stock Photography Agencies

Questionnaires were sent to the current providers of on-line catalogs (Comstock, Kodak Picture Exchange, PressLink, and Seymour) as well as to other stock photography agencies which provide either CD-ROMs or printed catalogs. The questions mainly focused on the following topics (for a sample questionnaire see Appendix A.)

- What types of catalogs (printed, CD-ROM, on-line) are provided?
- What search functions (keyword, natural language, predominant color, others) are offered with digital catalogs?
- How are the images delivered (hardcopy/slide, SyQuest, on-line?)

- Which graphics-file formats are supported?
- How much time is necessary to perform an average on-line catalog search?
- How do providers perceive customer satisfaction concerning image retrieval?
- How much time needs to be spent on archiving the images and maintaining the archive?

Survey on Stock Photography Users

The inquiries among stock photography users was split into two segments: daily newspapers and magazines. Using the questionnaire provided in Appendix B, selected companies of each segment were interviewed on the following issues:

- How many stock photos are typically used in a publication?
- What types of catalogs (printed, CD-ROM, on-line) are used?
- What search functions (keyword, natural language, predominant color, others) are offered with the digital catalogs in use?
- How are the images delivered (hardcopy/slide, SyQuest, on-line?)
- Which graphics-file formats are supported by the providers?
- How much time is necessary to perform an average search for one particular image?
- Could pattern recognition improve the search engine?
- How do publishing companies maintain their own image archives?

Since the stock photography market as well as the publishing market are very diverse, a statistical analysis would have been beyond the scope of this thesis. The collected data was therefore discussed for stock photography agencies, newspapers, and magazines separately. This allowed the author to draw a general picture of the awareness of each group concerning the problems that are encountered when building, maintaining, and using large commercial image databases. The replies to this survey also indicated where it would be feasible to introduce content-based image retrieval in order to enhance keyword search engines.

Chapter 6

Results

For this survey, three different target groups had been chosen: stock photography agencies, newspapers, and magazines. The selected stock photography agencies received a different questionnaire than the newspapers and magazines (see Appendix A.) Due to the setup of this survey, i. e., three different target groups with small sample sizes, no statistical analysis was conducted. The findings of this survey are discussed for each group separately and compared where appropriate.

Stock Photography Agencies

According to Allen Russell, chairman of the Picture Agency Council of America, only a quarter of the council's ninety-seven members have embarked on digital technology so far. Nevertheless, few of them doubt that they will have to get involved sooner or later.[1] In order to achieve a more detailed picture of the current situation of stock photography, the following agencies were contacted.

Table 1 - List of the Participating Stock Photography Agencies

Agency	City, State
The Bettmann Archives / Bettmann Newsphotos	New York, NY
Comstock	New York, NY
The Image Bank	Dallas, TX
PressLink	Reston, VA
Ro-Ma Stock Photography	Burbank, CA
Seymour (Picture Network International)	Arlington, VA
Stock Footage Connection	Niles, IL
Tony Stone Images	New York, NY

The sizes of the catalogs offered by these agencies range from 3,000 to more than 300,000 images. Surprisingly, all the companies that were contacted already have some kind of digital catalog or are in the process of developing one and only two of them deliver images exclusively in analog form. The discrepancy between the previously mentioned findings of the Picture Agency Council of America and the results of this survey may be caused by the small sample size of the survey or a rapid change in the stock photography market.

Most of the participating agencies cover a large variety of different image categories, although some of them, e.g., PressLink, concentrate on one specific topic, such as news photography. The customer base of the companies varies according to the main focus of their archives. News-oriented agencies focus mainly on newspapers and news magazines, while others target book, magazine, and multimedia publishers and the advertising market.

Four of the participating companies offer only keyword search capabilities with their digital catalogs. The other three offer a natural language search engine where phrases or sentences can be entered, two of them allowing users to enter a predominant color as a search criterium as well. Also, some of the companies mentioned further criteria that are entered in separate fields, such as a picture number, the photographer's name, location, and other descriptive categories.

The companies were also asked to estimate the ratio between the time spent on cataloging the images and the time required for scanning. The replies ranged from fifty percent for both scanning and cataloging to as much as ninety percent for cataloging and ten percent for scanning. The resulting average estimate was about sixty-five percent for cataloging and thirty-five percent for scanning. Thus, the cataloging process accounts for a considerable amount of time and any possibility to reduce this time without affecting the performance of the search engine would be welcome.

The survey participants were asked to estimate how much time their customers spent on an average search for one particular image. These estimates ranged between thirty seconds to fifteen minutes with an average of six minutes and forty seconds. Most participants considered their customers being very satisfied with this performance and only two of them agreed that their customers would ask for a faster image retrieval. The participants were also asked about the accuracy of their search engines, i. e., how well the search results complied with their customers' needs and ideas. Here, less than half of the companies considered their customers to be satisfied while the others presumed their customers would welcome some improvements.

Finally, the participants were asked whether or not content-based image retrieval would, in their opinion, improve search engines significantly. Only one participant expected no improvement, three participants were uncertain and three expected significant improvements. However, only two participant expected that content-based would reduce the need for assigning keywords while all the others disagreed with this point.

In the short interviews held along with the questionnaire, there seemed to be overall positive expectations concerning the capabilities of content-based image retrieval. It was beyond the scope of this thesis project to determine under which circumstances a company would chose to change from an existing keyword-based system to a new system with content-based image retrieval or choose the later one in the first place.

Newspapers

For the survey on the needs of newspapers, the author selected several different newspapers. In order to have a good mix, newspapers published in various states and of various sizes in terms of their average circulation were chosen. The participating newspapers are the following:

Table 2 - List of the Participating Newspapers

Title of Publication	City, State	Average Circulation[2]
The Buffalo News	Buffalo, NY	302,490
Daily Herald	Chicago, IL	121,091
Democrat & Chronicle	Rochester, NY	211,588
The Kansas City Star	Kansas City, MO	328,472
Los Angeles Times	Los Angeles, CA	1,012,880
San Francisco Chronicle	San Francisco, CA	518,125
USA Today	Arlington, VA	1,904,844

The amount of stock photography used in these newspapers is very low, ranging from two to three percent. These numbers apply only to the editorial part of the newspapers and include neither advertising nor pictures from newswire services, such as the Associated Press Leafdesk system. The catalogs which are typically used to select pictures are either on CD-ROM or accessed on-line. Most of the catalogs in use offer just a keyword search function and, in some cases, reference numbers and other search fields.

Although most of the material is handled digitally, there are still some stock photography agencies which deliver slides or hardcopy pictures. As for the digitally delivered pictures, the common graphics-file formats are JPEG and TIFF. Internally, the newspapers handle the files as JPEG, TIFF, or EPS.

The time needed to find a specific picture in a digital catalog strongly varies by search engine. In cases where a reference number can be used, the result is returned within a few seconds, whereas keyword searches take ten minutes or longer.

The users' satisfaction with the speed of the retrieval and the quality of the result varies accordingly. Although improved search engines would be welcome, none of the survey participants felt that newspapers could profit from the capability to search for an image based on it's visual content. As a matter of fact, for a news photography editor the results returned by a content-based search could be rather disturbing as illustrated below.

The following pictures (Figures 13 through 15, page 44) coincidentally arrived at the *Los Angeles Time*'s graphics library the same time during the author's interview with the head of this operation. Although they display completely different events, the three images show astonishing similarities in their visual appearance. A simple sketch (Figure 12) used for a content-based image retrieval would probably return all three sample pictures.

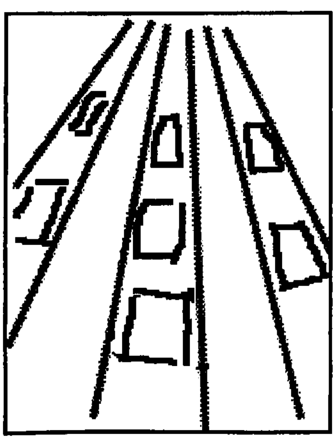

Figure 12 - Sample Sketch

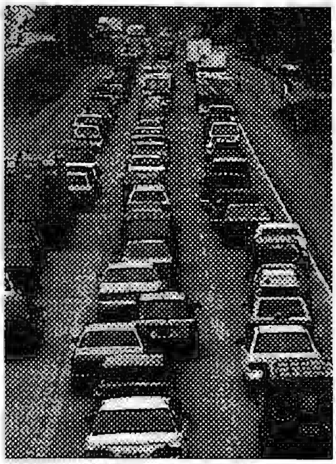

Figure 13 - Highway

Caption: "Northbound 101 at the 33 Junction was at a stand-still as a result of a water pipe break which caused mud and debris to flow over the freeway north of La Conchita. The freeway was closed both N/B and S/B for several hours." (Photography: *Los Angeles Times*)

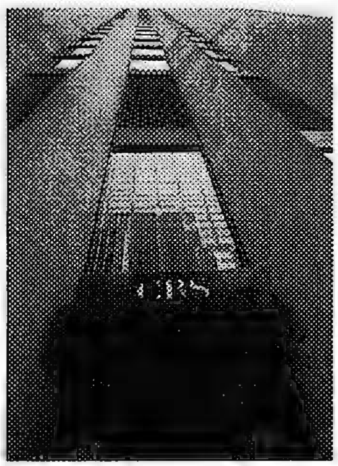

Figure 14 - CBS Building

Caption: "Television network CBS Inc. accepted a $5.4 billion, $81-a-share buyout offer from Westinghouse Electric Corp., Westinghouse said August 1. The headquarters of CBS, nick-named 'Black Rock,' in Manhattan are shown in this August 1 photo." (Photography: Mark Cardwell, REUTERS)

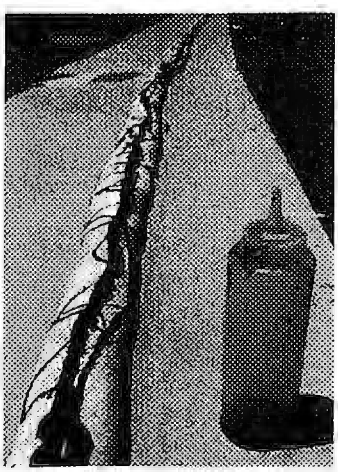

Figure 15 - Burrito

Caption: "Anaheim-The new world record burrito, weighing in at over 4,000 pounds and measuring 3,100 feet, was con-structed Monday at The Pond in Anaheim to benefit Olive Crest Homes and Services for Abused Children. Over two hundred people participated in the building of the burrito." (Photography: *Los Angeles Times*)

This example clearly demonstrates how little some news images have to do with the news that they are used with. The caption of an image often deals more with the circumstances under which an image was taken than with its visual content. Therefore, keywords are indispensable to news photography editors and content-based image retrieval may not be suitable for this group of stock photography users.

All of the contacted newspapers are maintaining their own image archive. One newspaper currently stores only hardcopies and slides, while all other newspapers maintain a digital archive as well. The author received detailed information on internal archiving from Mildred Simpson, Head of the Graphics Section in the Editorial Library of the *Los Angeles Times*. In her article:"The Electronic Picture Library At The Los Angeles Times,"[3] she describes the hardware and software in use, as well as the difficulties of setting up and maintaining an archive which currently has 65,000 images on-line.

> Even though our experience with electronic archiving has been generally successful, we often feel that we are running an experiment on a production basis. . . .
>
> We continue to have problems with both hard disk storage and floor space, performance and text handling. There are too many steps involved in achieving a fully indexed and edited record. We need to automate some of these processes.[4]

This emphasizes once more that the time required for cataloging images is one of the critical issues of digital image databases. However, none of the replies from newspapers indicated that shape-recognition technology could significantly speed up the archiving process in their environment by reducing the need for manually entered keywords.

Magazines

As with newspapers, the author tried to get answers from various points of view within the magazine publishing market. The list of contacted magazines was intended to represent the whole range from small to very large publications and from news magazines to special-interest magazines. However, some of the contacted magazines, such as Time Inc.'s *Life*, use only negligible amounts of stock photography. As it turned out, special-interest magazines are more likely to use stock photography than other magazines. Consequently, the variety of magazines contributing to the survey is not as great as originally intended. In addition, the magazine publishing market is heavily concentrated in New York City. Thus, it was impossible to determine whether any regional differences affected the survey results. The magazines which participated in the survey are the following:

Table 3 - List of the Participating Magazines

Title of Publication	Publisher	Circulation[5]
Backpacker	Rodale Press Inc. Emmaus, PA	186,113
Cosmopolitan	The Hearst Corp. New York, NY	2,714,639

Table 3 - continued

Title of Publication	Publisher	Circulation
Life	Time Inc. New York, NY	1,734,105
Travel & Leisure	American Express Publ. Corp. New York, NY	1,117,701
US News & World Report	US News. Washington, DC	N /A
Workbench	KC Publishing, Inc. Kansas, MO	860,000

The amount of stock photography used in these publications typically ranges from one to thirty percent. Using catalogs in any form is quite common although some of the photography editors send their request for specific images to professional image researchers. Fees for researching a single image may range from $75 flat rate to $450, or more.[6] The introduction of CD-ROM and on-line catalogs enables photography editors to do this research themselves which may be, although more time consuming, cheaper in the long run. Nevertheless, some magazines still work exclusively with printed catalogs.

Most commonly, the digital catalogs in use provide a keyword search and, in some cases, additional descriptive search fields. Some catalogs used by magazines are also able to search for images based on natural language entries, i.e., sentences or phrases in plain English. None of the survey participants mentioned catalogs where a predominant color or other criteria for content-based image retrieval was available. Obviously, these more advanced search engines are still very new to the stock photography market.

The estimates of how long an average search for one particular image would take ranged from ten to sixty minutes for printed catalogs. For digital catalogs, the estimates were generally lower, ranging from ten to forty minutes. The users' satisfaction in terms of speed and accuracy of the retrieval seemed to widely depend on the systems in use. Generally, the ratings of both speed and accuracy were lower than anticipated by the stock photography agencies. Thus, there seems to be a demand for improvements. Most of the contacted people could see advantages of content-based image retrieval for magazine publishing. Nevertheless, it was also mentioned that these retrieval methods could limit the performance of search engines in a similar way as keyword-based search engines do. Entering a particular color layout, for example, will return images which comply with this layout but may leave out interesting alternatives. The bottom-line is that the efficient use of these search engines requires as much training as keyword-based systems do.

Hardcopy and slides are still widely in use, although some magazines are acquiring images in digital form, mainly for positioning purposes. The most commonly used graphics-file formats are JPEG and TIFF, followed by PICT, GIF and Kodak CMS PhotoCD format, which agrees with the replies of the stock photography providers. Since all of these formats can be handled by common image manipulation software, compatibility does not seem to be an issue in this context.

Most of the contacted magazines maintain their own image archive in analog form, digitally, or both. None of the survey participants could see a significant improvement of the archiving process with content-based image retrieval since the need for manually entered keywords still remains.

Summary of the Survey Results

The replies from both users and providers of stock photography databases showed a demand for improvements of the archiving process and search engines. Magazines would profit more from the introduction of content-based image retrieval than newspapers would. However, the statement that content-based image retrieval would reduce the need for keywords clearly had to be rejected. Thus, from a provider's point-of-view, there are currently only a few reasons to change from a keyword-based system to a system with content-based image retrieval. Related issues, such as data compatibility, data transmission, and security do not seem to be as much of a problem as it had been stated in the Hypothesis of this thesis project.

Endnotes

[1] Kaplan, Karen, "The Cutting Edge: Computing / Technology / Innovation; A (Digital) Photo Opportunity; Agencies Begin The Move From Slides To CDs," *Los Angeles Times*, March 15, 1995, Part D: 4

[2] *Editor & Publisher Year Book*, Editor & Publisher Co., New York, NY, 1994

[3] Simpson, Mildred, "The Electronic Picture Library At The Los Angeles Times," *News Library News*, Vol. 16, No. 2, 1994: 12

[4] Ibid.

[5] *Benn's Media Directory*, 140th Ed., Vol.3, Benn Business Information Services, Tonbridge, Kent, UK, 1992

[6] Kaplan 1995

Chapter 7

Summary and Conclusion

The main purpose of this thesis project was to determine whether or not content-based image retrieval can improve digital stock photography catalogs. In order to obtain the necessary information, questionnaires were sent out to selected stock photography agencies, newspapers, and magazines. The replies were discussed for each group separately.

The replies obtained from stock photography agencies indicated interest in content-based image retrieval. However, corresponding technologies are not expected to replace manually assigned keywords describing the images' content and other significant information. Currently, the only software package offering encompassing content-based retrieval requires additional manual work at the database population stage to offer its users full profit from its capabilities. Since the amount of work involved with the database population is one of the major concerns of stock photography agencies, the chances of this software package succeeding are somewhat limited. From a provider's point of view, the question is whether or not the users' benefits from content-based image retrieval justify the required additional work.

The answer to this question mainly depends on the focus of a stock photography agency. For a news oriented agency, there may be no benefits at all. An agency which covers the whole range of stock photography may very well decide in favor of content-based image retrieval. Software packages which allow for content-based image retrieval will become more successful if the database population can be more automated. This may even convince an agency to make the transition from an existing keyword-based to a new content-based system.

Data compatibility did not appear to be an issue in the context of digital stock photography catalogs. The graphics-file formats in use are more or less standard and can be handled by common image manipulation and page layout applications. The problem of copyright security, on the other hand, has not been solved yet but does not affect the selection of a particular image database management system.

The survey on stock photography users showed mixed opinions about possible benefits from content-based image retrieval. As far as news photography is concerned, most survey participants denied any benefits. For images which are not news oriented, the participants agreed that content-based image retrieval could improve the performance of search engines. It would, however, not replace the need for keywords.

During his short interviews, the author realized that many participants had difficulties imagining what content-based image retrieval could possibly do.

For further research, the author therefore suggests to compare keyword-based and content-based systems with an existing image database. A corresponding project should compare the database population as well as the performance of the search engines. The author also suggests to select advertising agencies as a target audience where the images' relevance for news does not affect the results.

Bibliography

Bibliography

Benn's Media Directory, 140th Ed., Vol.3, Benn Business Information Services, Tonbridge, Kent, UK, 1992

Editor & Publisher Year Book, Editor & Publisher Co., New York, NY, 1994

Fisher, Yuval and Lawrence, Albert F. "Fractal Image Compression for Mass Storage Applications," in *Image Storage and Retrieval Systems: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA , 13 - 14 February, 1992*. SPIE Volume 1662, 1992

Gregory, Richard, "How do we interpret images?" in *Images And Understanding*, Barlow, Horace, Blakemore, Collin, Weston-Smith, Miranda ed., Cambridge University Press, Cambridge, UK, 1990: 310 -332

Hussain, Zahid, *Digital Image Processing: Practical Applications of Parallel Processing Techniques*. Ellis Horwood Limited, Chichester, UK,1991

IBM Ultimedia Manager 1.1 and Client Search, [Online], Available http://www.software.ibm.com/data/umm/umm.html

Image Storage and Retrieval Systems: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 13 - 14 February, 1992. SPIE Volume 1662, 1992

Kaplan, Karen, "The Cutting Edge: Computing / Technology / Innovation; A (Digital) Photo Opportunity; Agencies Begin The Move From Slides To CDs," *Los Angeles Times*, March 15, 1995, Part D: 4

Kevin Morris PhotoCD Sampler, [Online], Available http://www.digitalzone.com/cdrom/km/km1.jpeg

Query By Image Content (QBIC), [Online], Available http://www-i.almaden.ibm.com/cs/showtell/qbic/Initpage.html

Rabbani, Majid, "Image Compression Fundamentals" in *The Compression Experience: Proceedings of The Rochester Section of The SMPTE (Society of Motion Picture and Television Engineers) Tutorial Held in Rochester, NY, October 28, 1995*: 7 - 24

Rothman, Peter, Syntactic Pattern Recognition: Tutorial. *AI Expert*, Vol. 7, No. 10 (October 1992): 41 - 42

Seybold Publications. *The Seybold Report on Desktop Publishing*. Vol. 9, No. 1 (September 12, 1994)

Simpson, Mildred, "The Electronic Picture Library At The Los Angeles Times," *News Library News*, Vol. 16, No. 2, 1994: 12

Storage and Retrieval for Image and Video Databases: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 2 -3 February, 1993. SPIE Volume 1908, 1993

Storage and Retrieval for Image and Video Databases II: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 7 -8 February, 1994. SPIE Volume 2185, 1994

Storage and Retrieval for Image and Video Databases III: Proceedings of The SPIE (The Society of Photo-Optical Instrumentation Engineers) Congress Held in San Jose, CA, 9 - 10 February, 1995. SPIE Volume 2420, 1995

Appendix A

Questionnaire for Stock Photography Agencies

1. Which types of catalogs does your company provide

❑ Printed ❑ CD-ROM ❑ On-line

2. In which form do you deliver images to your customers

❑ Hardcopy/slide ❑ Electronic storage media (Disk, CD, etc.) ❑ On-line

3. How many images do you currently offer in your catalog

____________ images

4. On a scale from 1 to 5 (1 being negligible, 5 being very important): rate the importance of the following categories in your archive

__ Animals
__ Architecture
__ Art
__ Business
__ Celebrities
__ Environment
__ Education
__ Events & News
__ Home
__ Industry
__ Landscape
__ Nature

__ People
__ Politics
__ Recreation
__ Science
__ Sports
__ Still life
__ Travel
__ Transportation

Others — please specify:

____

____

5. How great is the overhead used for archiving the images and maintaining the archive as compared to the whole operation

❑ 0 - 10% ❑ 10 - 30% ❑ 30 - 50% ❑ 50 -70% ❑ More than 70%

6. On a scale from 1 to 5 (1 being negligible, 5 being very important): rate the importance of the following media for your business

__ Advertising __ Books __ Newspapers

__ News-Magazines __ Multimedia __ Special Interest Magazines

For companies providing *only* printed catalogs

7. Do you plan to digitize your catalog in the near future

❑ Definitely
❑ Probably
❑ Undecided
❑ Probably not
❑ Definitely not

7.1 If yes: what type of catalog do you plan to provide

❑ Electronic storage media ❑ On-line

8. Do you expect the overhead for scanning, categorizing and archiving to decrease or increase with a digital catalog

❑ Significantly decrease
❑ Decrease
❑ Stay the same
❑ Increase
❑ Significantly increase

For companies which provide digital catalogs (CD-ROM / On-line)

9. What search functions are offered with your digital catalogs

❑ Keyword ❑ Natural language ❑ Predominant color

❑ others — please specify:______________________________

10. In which file formats do you deliver your images to the customers (check all that apply)

❑ BMP ❑ EPS ❑ GIF ❑ JPEG

❑ Kodak CMS PhotoCD ❑ PCX ❑ PICT

❑ Targa ❑ TIFF

❑ others — please specify:______________________________

11. Considering the time spent on scanning and cataloging the images, how would you characterize the relationship between these two functions

Scanning:_____% Cataloging (incl. categorizing):_____%

12. How much time does the average customer need to search for a specific image in the catalog

___ Minutes

13. On a scale from 1 to 5 (1 being dissatisfied, 5 being very satisfied): how would you rate your customers' satisfaction concerning the *speed* of the image retrieval of digital image catalogs

❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5

14. On a scale from 1 to 5 (1 being dissatisfied, 5 being very satisfied): how would you rate your customers' satisfaction concerning the *results* that they get when they look for a specific image in a digital catalog, i.e., how well do the search results comply with their needs and ideas

❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5

Today's computer technology and storage media enable users to store large numbers of images digitally. Traditionally, keywords are assigned to the images during the cataloging process. These keywords have been the only means of retrieving images for later use. Recent advances in Artificial Intelligence (AI) and pattern recognition technology allow users to search for images based on their visual content using sample images, textures, predominant colors, and shapes. One of the goals of this research project is to determine the users' opinion about how this new technology could impact their work.

Please, consider the following two statements:

15.1 Software with shape recognition capabilities will enhance search engines of image databases significantly

❑ strongly agree
❑ agree
❑ may be, may be not
❑ disagree
❑ strongly disagree

15.2 Software with shape recognition capabilities will decrease the required overhead for image cataloging significantly since it reduces the need for manually entered keywords

- ❑ strongly agree
- ❑ agree
- ❑ may be, may be not
- ❑ disagree
- ❑ strongly disagree

Appendix B

Questionnaire for Stock Photography Users

1. How many stock photos are typically used in your publications as compared to other images

_____%

2. Which types of catalogs does your company use (check all that apply)

❑ Printed ❑ CD-ROM ❑ On-line

3. In which form are the images delivered to you (check all that apply)

❑ Hardcopy/slide ❑ Electronic storage media (Disk, CD, etc.) ❑ On-line

4. What search functions are offered with the *digital* catalogs that you use

❑ Keyword ❑ Natural language ❑ Predominant color

❑ others — please specify:______________________________

5. In which file formats do you receive digital images from stock photography agencies (check all that apply)

❑ BMP ❑ EPS ❑ GIF ❑ JPEG

❑ Kodak CMS PhotoCD ❑ PCX ❑ PICT

❑ Targa ❑ TIFF

❑ others — please specify:______________________________

6. Given the need for a specific picture, how long do you estimate an average search to take using a *printed* catalog

___ Minutes

7. Given the need for a specific picture, how long do you estimate an average search to take using a *digital* catalog (CD-ROM or on-line)

___ Minutes

8. On a scale from 1 to 5 (1 being dissatisfied, 5 being very satisfied): how would you rate your satisfaction concerning the *speed* of the image retrieval of digital catalogs

❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5

9. On a scale from 1 to 5 (1 being dissatisfied, 5 being very satisfied): how would you rate your satisfaction concerning the *results* that you get when you look for a specific image in a digital catalog, i.e., how well do the search results comply with your needs and ideas

❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5

10. Today's computer technology and storage media enable users to store large numbers of images digitally. Traditionally, keywords are assigned to the images during the cataloging process. These keywords have been the only means of retrieving images for later use. Recent advances in Artificial Intelligence (AI) and pattern recognition technology allow users to search for images based on their visual content using sample images, textures, predominant colors, and shapes. One of the goals of this research project is to deter-

mine the users' opinion about how this new technology could impact their work. Please, consider the following statement:

Software with shape recognition capabilities will enhance search engines of image databases significantly

- ❑ strongly agree
- ❑ agree
- ❑ may be, may be not
- ❑ disagree
- ❑ strongly disagree

11. Do you maintain your own image archive

❑ Yes ❑ No

11.1 If yes, which type of archive does your company maintain

❑ Hardcopy/slide ❑ Electronic storage media (Disk, CD, etc.) ❑ On-line

12. Consider the following statement:

Software with shape recognition capabilities will decrease the required overhead for image cataloging significantly since it reduces the need for manually entered keywords

- ❑ strongly agree
- ❑ agree
- ❑ may be, may be not
- ❑ disagree
- ❑ strongly disagree

Appendix C

Numerical Evaluation of the Replies from Agencies

1. Printed Catalogs: 3 CD-ROM Catalogs: 3 On-line Catalogs: 5

2. Hardcopy/slides: 5 Electr. storage media: 3 On-line: 4

3. Less than 10,000 images: 3
   10,000 to 100,000 images: 1
   More than 100,000 images: 3

4. Specialized in news: 2 Others: 5

5. 0 - 10%: 1 10 - 30%: 2 30 - 50%: 1 50 - 70%: 2

6.

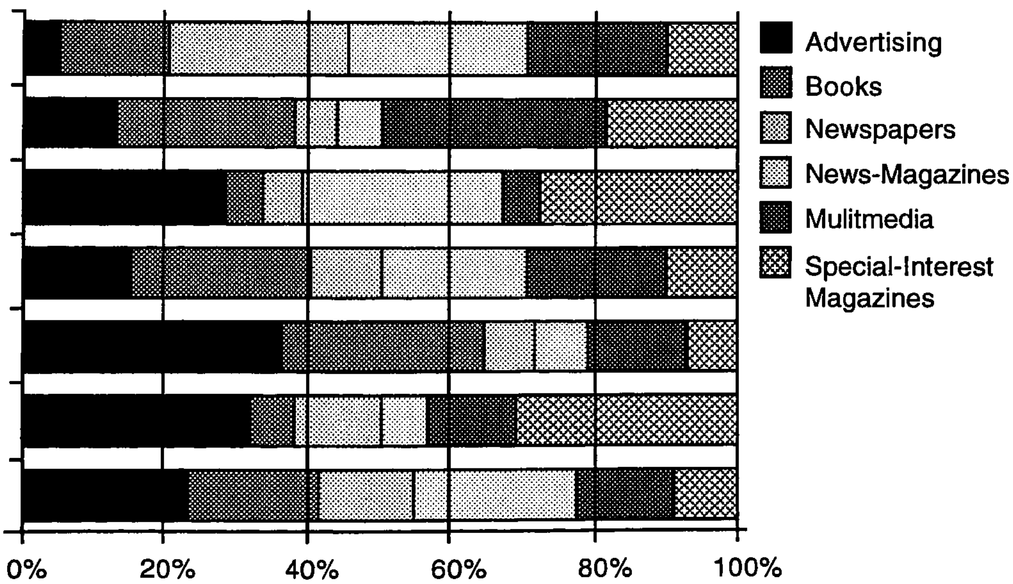

Figure 16 - Customer Base of the Participating Agencies

7. Probably: 1

7.1 Electronic storage media: 1

8. Significantly decrease: 1

9. Keyword: 5 Natural language: 3 Predominant color: 3
Others: Descriptive categories: 2 Picture number: 1

10. GIF: 1 JPEG: 4 Kodak CMS PhotoCD: 2 PICT: 1 TIFF: 4

11. Scanning 70% / Cataloging 30%: 1
Scanning 50% / Cataloging 50%: 2
Scanning 30% / Cataloging 70%: 3
Scanning 10% / Cataloging 90%: 1

12. 0.5 Min.: 1 1Min.: 1 1 - 2 Min.: 1 10 Min.: 1 12 Min.: 1
15 Min.: 1

13. Very satisfied: 3 Satisfied: 1 Undecided: 2 Dissatisfied: 1

14. Very satisfied: 2 Satisfied: 1 Undecided:3 Dissatisfied: 1

15.1 Strongly agree: 2 Agree: 1 May be, may be not: 3 Disagree: 1

15.2 Agree: 1 May be, may be not: 3 Disagree: 2 Strongly disagree: 1

Appendix D

Numerical Evaluation of the Replies from Users

Newspapers

1.	1% or less: 4	1 - 2 %: 2	3%: 1
2.	Printed Catalogs: 0	CD-ROM Catalogs: 3	On-line Catalogs: 7
3	Hardcopy/slide: 4	Electr. storage media: 5	On-line: 6
4.	Keyword: 6	Others (picture number): 2	
5.	EPS: 1 TIFF: 3	JPEG: 6	Kodak CMS PhotoCD: 1
6.	Does not apply		
7.	0,5 - 1 Minute: 1	10 - 15 Minutes: 3	Varies by search engine: 3
8.	Satisfied: 2	Undecided: 5	
9.	Undecided: 5	Dissatisfied: 2	
10.	Strongly agree: 1	May be, may be not: 2	Disagree: 4

11.	Yes: 7		
11.1	Hardcopy/slide: 7	Electr. storage media: 4	On-line: 2
12.	May be, may be not: 7		

Magazines

1.	1% or less: 1	20 %: 2	30%: 3
2.	Printed Catalogs: 3	CD-ROM Catalogs: 3	On-line Catalogs: 2
3	Hardcopy/slide: 5	Electr. storage media: 4	On-line: 2
4.	Keyword: 4 Reference number: 1	Naturale language: 2	Subject grouping: 1
5.	GIF: 1 TIFF: 3	JPEG: 3 PICT: 2	Kodak CMS PhotoCD: 1
6.	10 Minutes: 2	10 - 20 Minutes: 3	20 - 60 Minutes: 1
7.	10 Minute: 1 No reply: 2	10 - 20 Minutes: 2	40 Minutes: 1

8.	Undecided: 4	Dissatisfied: 1	No reply: 1
9.	Undecided: 4	Dissatisfied: 1	No reply: 1
10.	Agree: 1	May be, may be not: 4	Disagree: 1
11.	Yes: 5		
11.1	Hardcopy/slide: 5	Electr. storage media: 2	On-line: 1
12.	May be, may be not: 3	Disagree: 2	No reply: 1