A

Seminar-II Report

on

CONTENT BASED IMAGE RETRIEVAL

Submitted in Partial Fulfillment of
the Requirements for the Degree
of

Bachelor of Engineering

in

Computer Engineering

to

North Maharashtra University, Jalgaon

Submitted by

Madhuri M. Sathe

Under the Guidance of

Mr. Akash D. Waghmare

DEPARTMENT OF COMPUTER ENGINEERING
SSBT's COLLEGE OF ENGINEERING AND TECHNOLOGY,
BAMBHORI, JALGAON - 425 001 (MS)
2016 - 2017

SSBT's COLLEGE OF ENGINEERING AND TECHNOLOGY, BAMBHORI, JALGAON - 425 001 (MS)

DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the seminar-II entitled *CONTENT BASED IMAGE RETRIEVAL*, submitted by

Madhuri M. Sathe

in partial fulfillment of the degree of *Bachelor of Engineering* in *Computer Engineering* has been satisfactorily carried out under my guidance as per the requirement of North Maharashtra University, Jalgaon.

Date: October 7,2016
Place: Jalgaon

Mr. Akash D. Waghmare
Guide

Prof. Dr. Girish K. Patnaik
Head

Prof. Dr. K. S. Wani
Principal

Acknowledgement

I would like to thank all those who have contributed to the completion of the seminar and helped me with valuable suggestions for improvement. Our sincere thanks to principal Dr. Prof. K. S. Wani sir SSBT COET for providing me with best facilities and atmosphere for the creative work guidance and encouragement. Our deep gratitude goes to Dr. G. K. Patnaik sir ,head of the department,for granting us opportunity to conduct this Seminar work. I would like to thank my guide Mr. Akash D. Waghmare sir for all help and support extend to me. I thank all Staff members of my college and friends for extending their cooperation during my seminar. Above all I would like to thank my parents without whose blessings, I would not have been able to accomplish my goal.

Madhuri M. Sathe

Contents

List of Figures

Abstract

Content based image retrieval technique is used to retrieve images from large amount of image collection by using primitive features of images like color, shape, texture etc. It avoids the annotation of name with each image as in image retrieval using textual metadata. The main requirements of CBIR systems are techniques for feature extraction, query specification, user interactions etc.Techniques for feature extraction are described with currently existing techniques like color retrieval, shape retrieval and texture retrieval. The importance of user interaction and different methods for it is also described. The performance evaluation methods and practical applications of CBIR are also discussed in this topic.

Chapter 1

Introduction

Content-based image retrieval is a part of image processing and is also comes under artificial intelligence. Interest in digital images is growing day by day; Users in many professional fields are exploiting the opportunities offered by the ability to access and manipulate remotely-stored images in all kinds of new and exciting ways. The problems in image retrieval is becoming widely recognized, and the search for solutions an increasingly active area for research and development. So here is the technique Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases.

In Section 1.1 Content based image retrieval is presented. Architecture of content based image retrieval is presented in Section 1.2. In Section 1.3 requirement of content based image retrieval is presented Finally summary is presented in Section 1.4.

1.1 Content Based Image Retrieval

Advances in data storage and image acquisition technologies have enabled the creation of large image datasets. In order to deal with these data, it is necessary to develop appropriate information systems to efficiently manage these collections. Image searching is one of the most important services that need to be supported by such systems. In general, two different approaches have been applied to allow searching on image collections: one based on image textual metadata and another based on image content information. The first retrieval approach is based on attaching textual metadata to each image and uses traditional database query techniques to retrieve them by keywords. However, these systems require a previous annotation of the database images, which is a very laborious and time-consuming task. In fact, different users tend to use different words to describe a same image characteristic. The lack of systematization in the annotation process decreases the performance of the keyword-based image search. These shortcomings have been addressed by the so-called Content-Based

Image Retrieval (CBIR) systems. In these systems, image processing algorithms (usually automatic) are used to extract feature vectors that represent image properties such as color, texture, and shape. In this approach, it is possible to retrieve images similar to one chosen by the user (query-by-example) as shown in Figure 1.1. One of the main advantages of this approach is the possibility of an automatic retrieval process, contrasting to the effort needed to annotate images.

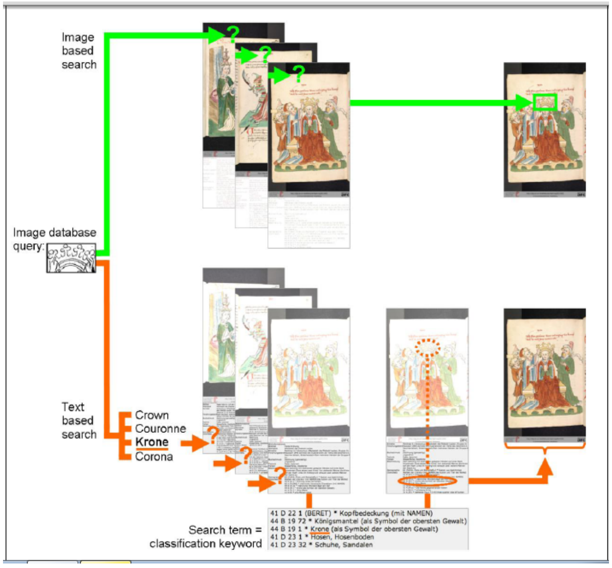


Figure 1.1: Image based search

1.2 Architecture of CBIR

Following Figure 1.2 shows a typical architecture of a content-based image retrieval system. Two main functionalities are supported: data insertion and query processing. The data insertion subsystem is responsible for extracting appropriate features from images and storing them into the image database (see dashed modules and arrows). This process is usually performed off-line [1]. The query processing, in turn, is organized as follows: the interface allows a user to specify a query by means of a query pattern and to visualize the retrieved similar images. The query-processing module extracts a feature vector from a query pattern and applies a metric (such as the Euclidean distance) to evaluate the similarity between the query image and the database images. Next, it ranks the database images in a decreasing order of similarity to the query image and forwards the most similar images to the interface module.

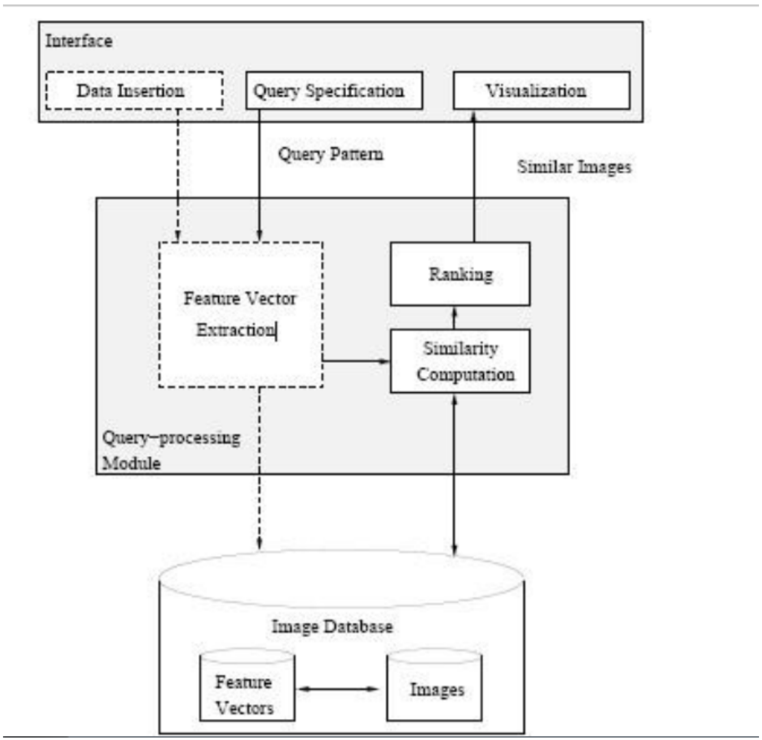


Figure 1.2: Architecture of CBIR

1.3 Requirements for CBIR systems

1.3.1 Techniques to extract metadata containing primitive features of images.

The metadata depict images with primitive features. Each primitive feature has its own representation format, such as the color histogram, which was first introduced in and has been widely used to represent the color feature. Another example is the shape feature, which can be represented by sets of consecutive boundary segments. With the appropriate metadata, current CBIR systems can achieve retrieval by color, by texture, by shape, by spatial relationship, and by a combination of the above.

1.3.2 Interface to obtain user's query demand

As in any retrieval system, the retrieval process begins with a query requirement. Hence, it is crucial to capture the users query demands accurately and easily. For current CBIR systems, the image retrieval process is usually carried out through an example image provided by the user, called query-by-example. However, users cannot always submit an example image to the retrieval system. Typically, current CBIR systems solve this problem by offering an interface to specify or choose some primitive features for providing an example image. For instance, when using IBMs QBIC system.

1.3.3 Methods to compare the similarities (or differences) between images

CBIR systems require methods that are based on primitive features to compare the similarities or differences between an example image and all the images in the image collection. However, the similarities or differences between images cannot be quantized in an ideal manner. The extent to which the images are similar will change when query requirements are varied. For instance, in the case of two pictures, one of a blue sea with a sunrise and the other of a green mountain with a sunrise, when the sunrise is considered, the similarity between these two images should be high, but if the object of interest is the blue sea, the similarity between these two images should be low. We believe that it would be very difficult to find a method to measure the similarities or differences between images accurately for all kinds of query demands. In other words, every retrieval method will have its own limits. For example, it will be hard for a color-based image retrieval technology to differentiate an image of blue sky from an image of blue sea. Therefore, when evaluating a CBIR technology, one should remember that the retrieval effectiveness of that technology depends on the types of query requirements that users make. As far as we know, there are no Department of Computer Science Content Based Image Retrieval. Benchmarks that contain a set of images on which CBIR technologies can be tested, nor there representative queries and their expected answers set to evaluate the retrieval effectiveness of CBIR technologies.

1.3.4 Techniques to efficiently index and store metadata.

For a large image collection, the storage space required for the metadata is usually non-trivial. A CBIR system must possess efficient techniques to compress the metadata. Also, retrieving images and other visual resources, such as video, demands that standards be set to describe the metadata. When a query is processed against a large image database, it is often unacceptable to compare the similarity between the query image and all the images, one by one. Because users only need images having high similarity to the example image, index structures, which can help to prevent sequential searches and improve query efficiency, should be used in CBIR systems. Further, for frequently varied image databases, dynamic index structures are necessary.

1.4 Summary

In this chapter the brief introduction about image retrieval systems is presented. In the next chapter literature survey is presented.

Chapter 2

Literature Survey

A literature survey or literature review means that anyone read and report on what the literature in the field has to say about any topic or subject. There may be a lot of literature on the topic or there may be a little. Either way,the goal is to show that he has read and understand the positions of other academics who have studied the problem/issue that he is studying and include that in ones seminar. It allows reader to establish his theoretical framework and methodological focus. Even if literature is proposing a new theory or a new method,literature is doing so in relation to what has be done.

In Section 2.1 background is presented. Related work is presented in Section 2.2.

2.1 Background

The use of digital images and pictures in communication is new what olden days cave-dwelling ancestors did was they use to paint pictures on the walls of their caves, and the use of maps and building plans to convey information almost certainly dates back to pre-Roman times. But the twentieth century has witnessed unparalleled growth in the number, availability and importance of images in daily life. They now play an important role in fields as diverse as medicine, journalism, advertising, design, education and entertainment. The term CBIR seems to have originated in 1992, when it was used by T. Kato to describe experiments into automatic retrieval of images from a database, based on the colors and shapes present. Since then, the term has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features. The techniques, tools and algorithms that are used originate from fields such as statistics, pattern recognition, signal processing, and computer vision.

2.2 Related Work

Related work is described as follows.

2.2.1 Color-based techniques

There are two traditional techniques for color-based image retrieval: Global Color Histograms that represent images with single histograms; and Local Color Histograms that divide images into fixed blocks and, for each block, obtain its color histogram. Global Color Histograms do not capture the content of images adequately, whereas Local Color Histograms contain more information and also enable the color distances among regions between images to be compared.

- The Global Color Histogram.

Using the Global Color Histogram, an image will be encoded with its color histogram, and the distance between two images will be determined by the distance between their color as shown in Figure 2.1.

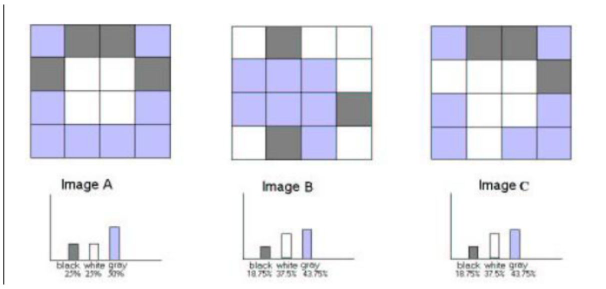


Figure 2.1: Histograms

In the sample color histograms there are three bins: black, white, and gray. We note the color histogram of image A : 25percent, 25percent, 50percent, the color histogram of image B :18.75percent,37.5percent,43.75percent and image C has same color histogram as B. If we use Euclidean metric to calculate histogram distance between A and B for GCH is

$$d_{GCH}(A, \bar{B}) = \sqrt{(0.25 - 0.1875)^2 + (0.25 - 0.375)^2 + (0.5 - 0.4375)^2} = 0.153;$$

The GCH is traditional method for color based image retrieval. However, it does not contain information about color distribution of regions, so the distance between images sometimes cannot show the real difference between images.

- The Local Color Histogram

This approach includes information concerning the color distribution of regions. The first step is to segment the image into blocks and then to obtain a color histogram for each block. An image will then be represented by these histograms. When comparing two images, we calculate the distance, using their histograms, between a region in one image and a region in same location in the other image. The distance between the two images will be determined by the sum of all these distances as shown in Figure 2.2.

$$d_{GCH}(A, B) = \sqrt{(0.25 - 0.1875)^2 + (0.25 - 0.375)^2 + (0.5 - 0.4375)^2} = 0.153;$$

Figure 2.2: The Local color Histograms

Using LCH to compute the distance between images A and B

$$\begin{aligned} d_{LCH}(A, B) = \; & \sqrt{(0.5 - 0.25)^2 + (0.25 - 0.25)^2 + (0.25 - 0.5)^2} \\ & \sqrt{(0.5 - 0)^2 + (0.25 - 0.75)^2 + (0.25 - 0.25)^2} \\ & \sqrt{(0 - 0.25)^2 + (0.25 - 0.25)^2 + (0.75 - 0.5)^2} \\ & \sqrt{(0 - 0.25)^2 + (0.25 - 0.25)^2 + (0.75 - 0.5)^2} \quad = \quad 1.768 \end{aligned}$$

In some scenarios, using LCHs can obtain better retrieval effectiveness than using GCHs. The above examples show that the LCH overcomes the main disadvantage of the GCH, and the new distances between images may be more reasonable than those obtained using the GCH. However, since the LCH only compares regions in the same location, when the image is translated or rotated, it does not work well.

2.2.2 Shape retrieval

There are generally two types of shape representations, i.e., contour - based and region-based. Contour-based methods need extraction of boundary information which in some cases may not available. Region-based methods, however, do not necessarily rely on shape boundary information, but they do not reflect local features of a shape. Therefore, for generic purposes, both types of shape representations are necessary.

- Fourier descriptors

Fourier descriptors are obtained by applying Fourier transform on shape boundary (usually represented by a shape signature), the Fourier transformed coefficients are called the Fourier descriptors of the shape. Shape boundary is a set of coordinates (xi , yi) , i = 1, 2, , L, which are extracted out in the preprocessing stage by contour tracing technique. Centroid distance function,

$$r_i = ([x_i - x_c]^2 + [y_i - y_c]^2)^{1/2}, \quad i = 1, 2, \ldots, L$$

where x c , y c are averages of x coordinates and y coordinates respectively. Due to the subtraction of centroid (which represents the position of the shape) from boundary coordinates, the centroid distance representation is invariant to shape translation. Fourier transform of ri , i = 0, 1, , N-1 is then given by

$$u_n = \frac{1}{N} \sum_{i=0}^{N-1} r_i \exp\left(\frac{-j2\pi n i}{N}\right), \; n = 0, 1, \ldots, N-1$$

The coefficients un , n = 0, 1, , N-1, are usually called Fourier descriptors (FD) of the shape, denoted as FD n , n = 0, 1, , N-1. The similarity measure of the query shape and a target shape in the database is simply the Euclidean distance between the query and the target shape feature vectors.

Advantages of fourier descriptor(FD):- The nice properties of Fourier descriptors are its robustness, being able to capture some perceptual characteristics of the shape and easy to derive. Noise is not a problem with Fourier descriptors, for noise only appears in very high frequencies which are truncated out.

2.2.3 Texture retrieval

Texture extraction process could be used to retrieve color images. Using a texture extraction process alone to retrieve color images, would result in badly ranked recalls (e.g. best match for the query in terms of color composition may be placed as the lowest ranked image). The main purpose of image retrieval is to retrieve a set of images matching the query and the best match always should be ranked highest otherwise it defeats the whole purpose of image retrieval.

Color image consists of three layers i.e. Red, Green and Blue (RGB). The color composition could be represented in a RGB space as shown in Figure 2.3. The diagonal running through the space is the grayscale area. From the RGB space we can see that any color is a combination of red, green and blue elements.

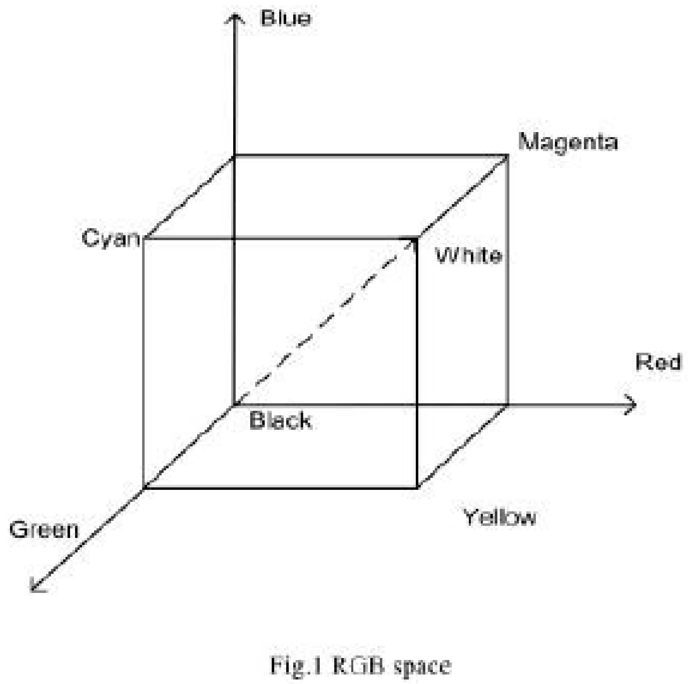


Figure 2.3: RGB Space

An area with yellow would have a combination of red and green color elements only. Similarly an area with magenta would have combination of red and blue elements only. Thus if you consider an image of a object with yellow color, then the object is best described in red and green layers. Harnessed on this theory to enhance the capabilities of a texture extraction process to retrieve color. Certain texture/shape features which appear in the red layer does not appear with the same intensity in the green or blue layers and vice versa is also true. Hence, if the feature vectors of each layer of the query have matched with feature vectors of each layer of a particular image in the database; conclude that the query image color composition has matched with a color Composition of an image in database as shown in Figure 2.4.

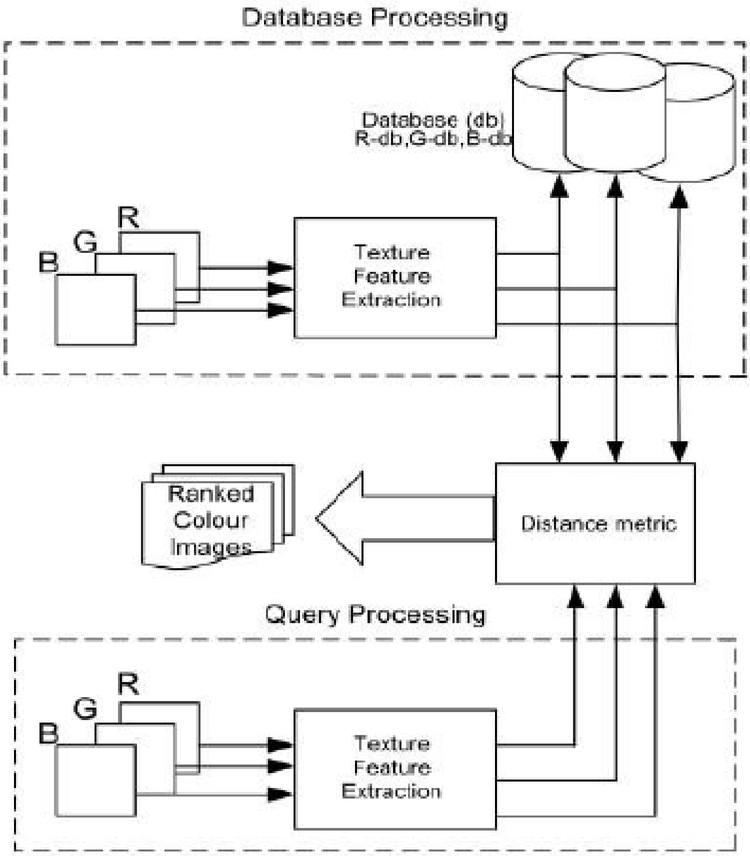


Figure 2.4: Database and query processing

2.2.4 User interaction

CBIR systems have access to feedback from their users that can be exploited to simplify the task of finding the desired images. There are, therefore, two fundamental problems to be addressed. First, the design of the image representation itself and, second, the design of learning mechanisms to facilitate the interaction. In response to a query, the CBIR system suggests a few images and the user rates those images according to how well they satisfy the goals of the search. Because each image usually contains several different objects or visual concepts, this rating is both difficult and inefficient.

2.3 Summary

In this chapter literature survey is presented. Methodology is presented in next chapter.

Chapter 3

Methodology

A system of methods used in a particular area of study or activity. A methodology for investing the concept of focal points. Methodology is the systematic,theoretical analysis of the methods applied to a field of study. A methodology does not set out to provide solutions-it is ,therefore,not the same thing as a method. Instead,it offers the theoretical underpinning for understanding which method,set of methods or so called best practices can be applied to specific case.

In Section 3.1 query techniques is presented. Performance evaluation method is presented in Section 3.2.In Section 3.3 summary is presented.

3.1 The Query Techniques

Different implementation of CBIR makes use of different types of user queries.

3.1.1 Query Example

Query by example is a query technique that involves providing the CBIR system with an example image that it will then base its search upon. The underlying search algorithms may vary depending on the application, but result images should all share common elements with the provided example. Options for providing example images to the system include:

- A preexisting image may be supplied by the user or chosen from a random set.
- The user draws a rough approximation of the image they are looking for, for example with blobs of color or general shapes.
- This query technique removes the difficulties that can arise when trying to describe images with words.

3.1.2 Semantic Retrieval

The ideal CBIR system from a user perspective would involve what is referred to as semantic retrieval, where the user makes a request like "find pictures of dogs" or even "find pictures of Abraham Lincoln". This type of open-ended task is very difficult for computers to perform -pictures of Chihuahuas and Great Danes look very different, and Lincoln may not always be facing the camera or in the same pose. Current CBIR systems therefore generally make use of lower-level features like texture, color, and shape, although some systems take advantage of very common higher-level features like faces.

3.1.3 Other query methods

Other methods include specifying the proportions of colors desired (e.g. "80percent red, 20percent blue") and searching for images that contain an object given in a query image. CBIR systems can also make use of relevance feedback, where the user progressively refines the search results by marking images in the results as "relevant", "not relevant", or "neutral" to the search query, then repeating the search with the new information.

3.2 Performance evaluation methods

Performance evaluation methods are described in detail. They are as follows.

3.2.1 user comparisons

User comparison is an interactive method. Users judge the success of a query directly after the query. It is hard to get a large number of such user comparisons as they are time-consuming.

3.2.2 Before-after comparison

This is the easiest test method. Users are given two or more different results and allowed to choose the one which is preferred or found to be most relevant to the query. This method needs a base system or another system for comparison.

3.2.3 Rank of the best match

Whether the most relevant image is in either the first 50 or first 500 images retrieved. 50 represents the number of images returned on screen and 500 is an estimate of the maximum number of images a user might look at when browsing.

3.2.4 Precision and recall

The most common evaluation measures used in IR are precision and recall usually presented as a precision vs. recall graph (PR graph).

$$precision = \frac{\text{No. relevant documents retrieved}}{\text{Total No. documents retrieved}},$$
$$recall = \frac{\text{No. relevant documents retrieved}}{\text{Total No. relevant documents in the collection}}.$$

3.2.5 Target testing

Users are given a target image and the number of images which the user needs to examine before finding the target image is measured. Starting with random images, the user marks images as either relevant or non-relevant.

3.2.6 Error rate

Error rate is calculated for retrieval process.

$$\text{Error rate} = \frac{\text{No. non-relevant images retrieved}}{\text{Total No. images retrieved}}.$$

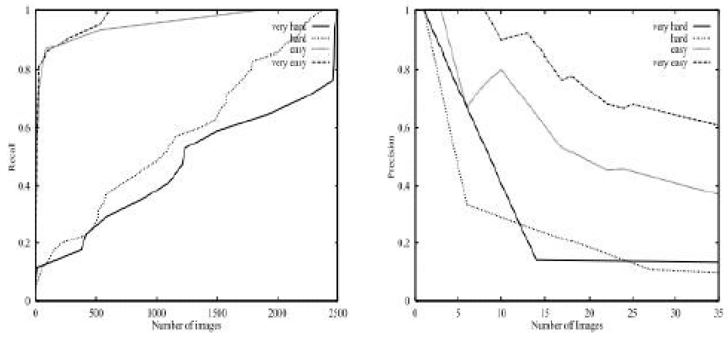
3.2.7 Retrieval efficiency

If the number of images retrieved is lower than or equal to the number of relevant images, this value is the precision; otherwise it is the recall of a query.

$$\text{Retrieval efficiency} = \begin{cases} \frac{\text{No. relevant images retrieved}}{\text{Total No. images retrieved}} & \text{if No. retrieved} \\ & \text{No. relevant} \\ \frac{\text{No. relevant images retrieved}}{\text{Total No. relevant images}} & \text{otherwise.} \end{cases},$$

3.2.8 Correct and incorrect detection

The numbers of correct and incorrect classifications are counted. When divided by the number of retrieved images, these measures are equivalent to error rate and precision.



(a) recall graph (b) precision graph

The recall graph can distinguish well between the hard and easy queries, but not too well between the easy and very easy one. A complete precision graph does not contain much information in this case, that is the reason for printing a partial one. The recall graph looks more positive than a PR graph, especially when a few relevant images are

3.3 summary

In this chapter the query techniques and performance evaluation method is presented. In next chapter advantages ,disadvantages and application are presented.

Chapter 4

Discussion

After discussing the methodology of content based image retrieval in the previous chapter, now concentrate on the discussion of content based image retrieval.

In Section 4.1 advantages is presented. Disadvantages is presented in Section 4.2. In section 4.3 application are presented. Finally summary is presented in section 4.4.

4.1 Advantages

Advantages are as follows:

- CBIR has become a very active area research for two major research communities. Database Management and Computer Vision
- Feature Extraction methods are easy, effective and less expensive.
- Time require is less to find all those related images.
- More than one related outcomes occur by only one search(If more one equally likely image present in the database).

4.2 Disadvantages

Disadvantages are as follows.

- We not yet have a universally acceptable means of characterizing human vision,more specifically in the context of image understanding.
- Hence it is not surprising to see continuing efforts towards it,either building up on prior work or exploring novel directions.
- Accuracy is not up to the mark.
- Continuous filtering is needed for achieving accurate output.

4.3 Practical application of CBIR

A wide range of possible applications for CBIR technology has been identified. Potentially fruitful areas include:

4.3.1 Crime prevention

Law enforcement agencies typically maintain large archives of visual evidence, including past suspects facial photographs (generally known as mug shots), fingerprints, tyre treads and shoe prints. Whenever a serious crime is committed, they can compare evidence from the scene of the crime for its similarity to records in their archives. Strictly speaking, this is an example of identity rather than similarity matching, though since all such images vary naturally over time, the distinction is of little practical significance.

E.g: AFIX used by FBI

Face recognition is also a reasonably mature technology. Most current systems use either a version of the eigen face method initially developed for the Photobook system at MIT.

4.3.2 Intellectual Property

Trademark image registration, where a new candidate mark is compared with existing marks to ensure that there is no risk of confusion, has long been recognized as a prime application area for CBIR. Copyright protection is also a potentially important application area. Eg: QBIC and Virage.

4.3.3 Architectural and engineering design

Architectural and engineering design share a number of common features the use of stylized 2- and 3-D models to represent design objects, the need to visualize designs for the benefit of non-technical clients, and the need to work within externallyimposed constraints, often financial. Such constraints mean that the designer needs to be aware of previous designs, particularly if these can be adapted to the problem at hand. Hence the ability to search design archives for previous examples which are in some way similar, or meet specified suitability criteria, can be valuable. Eg: SAFARI.

4.3.4 Medical diagnosis

The increasing reliance of modern medicine on diagnostic techniques such as radiology, histopathology, and computerized tomography has resulted in an explosion in the number and importance of medical images now stored by most hospitals. While the prime requirement for medical imaging systems is to be able to display images relating to a named patient, there is increasing interest in the use of CBIR techniques to aid diagnosis by identifying similar past cases. I2C system for retrieving 2-D radiological images from the University of Crete.

4.3.5 Cultural heritage

Museums and art galleries deal in inherently visual objects. The ability to identify objects sharing some aspect of visual similarity can be useful both to researchers trying to trace historical influences, and to art lovers looking for further examples of paintings or sculptures appealing to their taste. IBMs QBIC system has received extensive trials in managing art library databases.

4.3.6 Web searching

Text-based search engines have grown rapidly in usage as the Web has expanded; the well-publicized difficulty of locating images on the Web [Jain, 1995] indicates that there is a clear need for image search tools of similar power. Eg: WebSEEk, ImageRover, Yahoo! Image Surfer (1998).

4.4 Summary

In this chapter Advantages,Disadvantages,Applications is presented. Conclusion is presented in next chapter.

Chapter 5

Conclusion

5.1 Conclusion

The most important contribution of this research is the proposed hybrid method combining the advantages of low-level image characteristics extraction with textual description of image semantics. Most systems use color and texture features, few systems use shape feature, and still less use layout features. The retrieval on color usually yields images with similar colors. Retrieval on texture does not always yield images that have clearly the same texture, unless the database contains many images with a dominant texture. Searching on shape gives often surprising results. Apparently the shape features used for matching are not the most effective ones. It is difficult to evaluate how successful contentbased image retrieval systems are, in terms of effectiveness, efficiency, and flexibility. Of course there are the notions of precision (the ratio of relevant images to the total number of images retrieved) and recall (the percentage of relevant images among all possible relevant images). From these statements coming to a conclusion that CBIR systems currently existing are primitive and need some more advanced techniques.

Bibliography

[1] Yu Xiaohong,Xu Jinhua The Related Techniques of Contentbased Image Retrieval *International Symposium on Computer Science and Computational Technology*, 2008 .

[2] Ryszard S. Choras,Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems, *International Journal of biology and biomedical engineering*, Issue 1, Vol. 1, 2007

[3] Alaa H. Halawani, Alexandra Teynor, Lokesh Setia, Gerd Brunner,Hans Burkhardt, Fundamentals and Applicationsof Image Retrieval: An OverviewDatenbank- Spektrum 18/2006.

[4] Tobias Weyand and Thomas Deselaers, Combining Content based Image Retrieval with Textual Information Retrieval , Department of Computer Science, RWTH, Aachen, October 2005.

[5] Zijun Yang and Jay Kuo, Survey on content-based analysis, indexing and retrieval techniques and status report of MPEG- 7,*Tamkang journal of science and engineering*, vol.2, No.3 pp.101-118, 1999.