International Journal of Computer Applications Technology and Research
Volume 5– Issue 6, 358 - 363, 2016, ISSN:- 2319–8656

Features Analysis in CBIR Systems

Poonam Bhatia Anand
MMICT&BM
Maharishi Markandeshwar University
Mullana, Ambala
India

Neetu Bhatia
MMICT&BM
Maharishi Markandeshwar University
Mullana, Ambala
India

Abstract: Image retrieval is the major innovations in the development of images. Mining of images is used to mine latest information from the general collection of images. CBIR is the latest method in which our target images is to be extracted on the basis of specific features of the specified image. The image can be retrieved in fast if it is clustered in an accurate and structured manner. In this paper, we have the combined the theories of CBIR and analysis of features of CBIR systems.

Keywords: Image retrieval, Content based image retrieval, CBIR techniques, Systems and features.

1. INTRODUCTION

In the present era, images play a big role in every part of life. As images are increasing day by day on the web, Retrieval of efficient and accurate images from the web is very crucial task in the field of image retrieval. Presently there are two methods that are used for retrieving the images. One method is Text Based Image Retrieval in which retrieval of images is done using textual features of the images. This method does not use the visual features of the images. Due to this drawback some problems exist to meet accurate images in a right way [19] when we search the image database. On the Other Hand, the second method used for retrieval of image is CBIR. Its goal is to search the image from a huge image database according to the query given by the user. In CBIR, the color, texture and shape are visual features extracted from the image. Therefore it is also called visual image Retrieval. In the next section the model of CBIR is discussed.

2. CONTENT BASED IMAGE RETRIEVAL

CBIR uses the exclusive part of image to signify and true to use. In common CBIR system is divided into following three phases which are shown in figure 1:

- Extraction of features
- Feature Matching
- Semantic image retrieval

In this system, first the query to the system provide by the user. The user's query can be either in the form of wording or in the form of an picture. Using feature extraction techniques, from the image, the features like color, texture and shape are extracted. After that, these features are matched with the images that are stored in the database of images. Then various similarity measures are used by a variety of systems to allocate the index value of the images or objects.

CBIR is considered as one of the efficient ways for retrieving images. It points at budding new procedure that supports well-organized probing and browsing of huge libraries of digital image depend on derived features of the images. First, the image is divided into chunks of equal size. This system uses the image content directly that will be searched in an image database. The key plan of this is to analyze information of the image by using features of color, texture, shape, face detection and layout of an image.

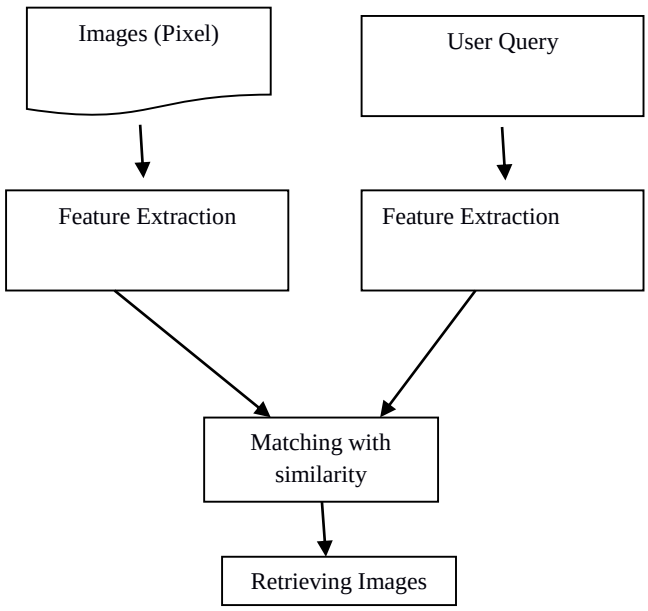

Fig 1: CBIR System [21]

The various features used in CBIR Systems are explained in the next section.

3. FEATURES USED IN CBIR

Features are the properties of the image. It describes the image in various ways. There are two types of features used in CBIR Systems i.e. Low Level Features and High Level Features [3,12]. These Features are described as follows:

3.1 Low level features

The primary features used in CBIR Systems are low level features. Following are the basic low level features:

3.1.1 Color

This feature refers to the color of different parts of the images. This feature can be extracted by methods such as Histogram method Arithmetical method Color / Shade model.

3.1.2 Texture

Texture refers to visual patterns in images. Texture is used to represent texture in images. Texture is also one of the important kind of an image. This feature refer to the visual patterns including the surface of clouds, trees, bricks, hair, and fabric. A variety of algorithms have been proposed for texture testing: Gray Level Co occurrence, The Tamura Texture Feature, The Model of Markov Random Field, Gabor Filtering, Binary Local Patterns.

3.1.3 Shape

Shape refers to shape of particular region in the image. Shapes can be determined by applying segmentation of the image. Shape Descriptors are used to translate, rotate and scale the images. Boundary based and Region based Shape Representation is used to represent the image.

3.2 High level Features

The high level features are discussed as below:

3.2.1 Metadata [19]

Meta data is data about data that are related to the process of the image creation. The Meta attributes include image acquisition date, image identification number and name, image modality device, image magnification, etc. Due to manual annotation in images retrieval following limitations exists:

1. Excess time consumption
2. Expensive task for large image database
3. Retrieval of non subjective, context sensitive and incomplete data.

Therefore CBIR is used for image retrieval to overcome the disadvantages of keyword annotation method.

3.2.2 Semantics

Semantic image retrieval [7] begins by request made by a person, e.g. "find pictures of Mahatma Gandhi". The above said task is very difficult to perform for computers as you will not find Mahatma Gandhi always posing in the same pose or in front of the camera. To evade this problem, CBIR systems uses feature of low level such as color, shape and texture. The combination of these features is used with databases which are already well equipped and qualified to match these features, i.e. fingerprints, faces or shapes, etc. Although in general, high-level identification concepts of image retrieval require human feedback.

In the next section CBIR Systems are explained.

4. CBIR SYSTEMS

There are lots of areas such as academia, commerce, government and hospitals in which large group of digital images are being formed. Some of those images are the manufactured goods of digitizing accessible group of diagrams such as drawings, paintings, and prints in these image collections [8]. Normally, keyword indexing is the only method of searching this collection. In this section we explain some technical phase of current CBIR systems. There are number of image database systems, retrieval of images, or Systems of multimedia information has been published. The major target of this clarification is to provide a review of the short-term CBIR systems.

Some of the commonly used CBIR Systems are described as follows:

4.1 Alexandria Digital Library(ADL)

In this CBIR system [17], the only texture feature is extracted from the images using with texture features images can explore. With the help of browser map the user can interrelate magnifies a two dimensional world map to find its significance area, and choose a question area that has to involve with the database images. Right now, a list of images, query parameters can be selected by the user for example, aerial of the photo, map, images of remote sensor etc. and then with the help of browser map the images that are overlapped with area are retrieved.

4.2 Advanced Multimedia Oriented Retrieval Engine(AMORE)

In this CBIR system [18], the images are divided into eight sections by region or all the same color and downsize into pixels of 24*24. At firs, the user selects a different group of images. The first image group is being chosen at random. Then selected images, generally

parallel descriptions images can be getting back. The query of images can also be specific by its Uniform Resource Locater. Then the comparative significance of color and shape is indicated by user.

4.3 Blobworld

In this CBIR system [5], the color, shape of regions which are called blobs, texture and location are features of the query image. In Lab space the color description is done by the coordinates of the histogram of 218 bins. First, categories have the limiting search is selected by the user. After selecting a blob, the user points to the importance of the region such as somewhat and very. After then the user points at blob's color, texture, location, and shape. For querying, more than one region can be used by the user.

4.4 Candid

In this CBIR system [13], each image is represented by signature containing Gaussian functions. The clusters of spectral bands are used by color features. The meaning of every pixel and Skegness are used by texture. A query image is provided by the user.

4.5 Content based Visual Query

The system enables queries by texture is called as CBVQ. First, a 9-dimensional texture vector is calculated for each pixel in the image. After then by using the filtering of non linear techniques, the pixels of the image are merged into texture regions of homogeneous to reduce noise followed by algorithm of a sequential labelling then by overlapping regions and renovation of the images is done from the images of the binary sub band [23]. By using 9 dimensional binary vectors, each region is represented. Spatial information is also mined for each texture region detected; a global color histogram is also calculated for every image database. With queries by example the system allows outline in which one of the display random images can be selected by the user or address of any image gives the URL and by a color histogram user can give direct queries. The available search methods is chosen by the user selects in a query for example histogram of color, manipulation of texture To build a new query histogram, a histogram of the query image can be used.

4.6 Chabot

This system is made up of an energetic text record of the collection, the location of the picture[20]. A 20 bins of color histogram are worked out for each image. There is certeria of a search are presented by user such as keywords, location and colors. Although there are some partial options of the color criterion for the user. The user has the many options to organize thoughts, which is combination of criteria of search in which thought make happy. For instance, as a sunset is grouping of keyword and mainly red or mainly orange is the color criterion used for the concept of 'sunset' is described.

4.7 Flexible Image Database System(FIDS)

In this CBIR system [4], the features are also taken in the formation of the sub image grid such as rows and columns. The histogram values is obtained after applying the filter of Sobel edge. The inquiry image as an image is chosen by the user. The feature distance measures can also be chosen by the user, and merge them. As a next inquiry image the result image is displayed.

4.8 Fast Object Color Based Query System(FOCUS)

In this CBIR system [6], every image is distributed in a group of cells having dimensions of 100×100 pixels. In HSV space each cell is grouped as a color histogram. SPG (Spatial proximity graph) is built in many stages to obtain the target of characterizing spatial relationships between color regions. The base of connecting two nodes is

a) In the same cell resultant peaks situated.
b) Resultant peaks situated in different groups, but have the same color.

All linked nodes of the same color are then unified by merging them to a single node. Adjacency matrix representation is then used for storing linked nodes. A global color histogram is calculated and determined the relationship of color region for the inquiry image at the pixel level.

4.9 Jacob

The color and texture is used in this CBIR system and a histogram in RGB space is used for Color characterizing [15]. In this, query may be straight. By inserting a few standard queries which made straight by indicating the histogram and/or features of the texture. The user must give an image, for an example of a query.

4.10 Metaseek

In this CBIR system [3], to select target searching engines in a presentation, list MetaSeek is used. After then QBIC, VIRAGE, WebSeek and the VisualSeek are the CBIR systems followed for the real matching. The client can choose a class and offer a keyword, then give a URL of an image. With the help of Metaseek, the sender validates the list presentation performance upon acceptance of a query. These databases contain scores of history query successes.

4.11 Multimodel Information Retrieval System

This system merges various techniques to get images of descriptions semantically. With the help of natural language processing (NLP) techniques, picture's information is extracted in analyzing image captions [24]. This information consists the people with their names, location and time of picture taken with optical description such as gender and hair color. The pictures are detected by NLP in which peoples are presented which are identified by face to verify the guess. When the area of face's is cropped out then for remaining images, a histogram is calculated. For the scenery, a histogram of color correlation is computed in which no face is identfied by the NLP analysis.

By consisting of text string, an image and different topics about sports, politics and entertainment, query is constructed. At this, the relative significance of text vercus content as well as foreground versus background can also be painted by the user.

4.12 Netra

In this CBIR system [16], By using color, the database of Images is created by division of images. To characterize the shape of regions, three feature vectors are used. The first one is the curvature function of the curve. The second one is the region of the centroid function and the third one is the function of complex coordinate. In 25 categories, there are 2500 Corel photo images which are gathered and organized with 100 images in each category. As the enquiry image, user can decide on any one of them. All database images are segmented in all the same regions. Color, spatial location, texture and shape are main characteristics of the image which the user can choose. Color and spatial location can be directly used for image example. The spatial location is used to illustrate the area of interest of two bound box, for confine the favorite region point to the inner box and for confine the region of objects point to the outer box.

4.13 Photobook

In this CBIR system [22] , three types of looms are applied to create an image demonstration for purpose query, each for a specific type of image contents: faces, 2D shapes and images of texture. The first two demonstrations are same because they propose an explanation by using covariance of the eigenvectors. To perform a query for present an explanation filter is used from network of images the user chooses from a little explanation of images. From that display, another query image might be selected by the client and repeats the search.

4.14 Pictoseek

The queries based on content color and shape are classified as features in this CBIR System [10]. Different color, features, the effect of various imaging conditions is examined below the statement of a model of dichromatic reflection. In this system reflection is used for designing that how a cluster of color is constructed. By the system's Web crawlers, with help of giving a URL address the query images from the database is chosen to compose. The desired variance is chosen by the user before giving the query, which points to the system.

4.15 Query by Image Content (QBIC)

In this CBIR system [26] the features of texture are used with customized versions of the roughness, dissimilarity. We put queriea by using image content. The shape area is used by shape features circularity, eccentricity, major axis orientation. These shape features are planned partially involuntarily and mined for all objects. On the base of image example, color, patterns and constructed sketch by user, queries are allowed by QBIC system. By using the slider, the percentage of color is adjusted in the image.

4.16 Virage

This system is an extensible framework structure for constructing CBIR systems [1]. In this CBIR system, the basic concept is the type of feature, computation and distant matching. There are five abstract data types: a) global value, b) color histograms, c) local value, d) histograms and e) graphs. The VIRAGE system gives a set of general primitives. The VIRAGE system gives a GUI tool set essential for the growth of a GUI edge. These contain services for placing of images, image queries and maintain many popular formats of image file. Query by sketch is also used where picture can be sketch by user with drawing tools and color palette.

4.17 Visualseek

In this CBIR system [23], using the back-projection technique, the color region extraction takes place. In this system of the population, every image is decayed robotically into region of foremost colors. Features and Spatial properties are used in every region. For start a query, a number of regional sketches by the user and selects a color for each region. After then it points to location boundary, size and relationship between regions.

5. ANALYSIS OF FEATURES IN CBIR SYSTEMS

Low level features, High level features and Keywords which are used by CBIR Systems are shown in table 1. Some systems uses high level features and some systems used low level features and some systems uses keywords. High level features are face detection and layouts. Color, shapes and texture are low level features. Color features can be described as Color moments, Global histogram, Dominant colors and Correlation histogram. Shape features are described as

Table1: Feature Analysis in CBIR Systems

CBIR SYSTEMS	Features														
	Low Level Features												High Level Features		
	SHAPE				TEXTURE				COLOR						
	Fourier Descriptor	Elementary Descriptor	Bounding Box	Template Matching	Atomic Texture	Wavelet, Gabor, Fourier	Edge Statistics	Random Fields	Dominant Colours	Colour Moments	Global Histogram	Correlation Histogram	Face Detection	Layout	KEY-WORDS
ADL						Yes									Yes
AMORE				Yes					Yes						Yes
BLOBWORLD		Yes			Yes									Yes	
CANDID					Yes				Yes						
CBVQ		Yes				Yes					Yes				
CHABOT											Yes				Yes
FIDS						Yes	Yes				Yes			Yes	
FOCUS		Yes												Yes	
JACOB					Yes						Yes	Yes			
METASEEK					Yes						Yes				Yes
MIR					Yes	Yes					Yes		Yes	Yes	Yes
NETRA	Yes					Yes								Yes	
PHOTOBOOK								Yes					Yes		
PIC TO SEEK	Yes									Yes	Yes				
QBIC		Yes		Yes	Yes					Yes	Yes			Yes	Yes
VIRAGE															Yes
VISUALSEEK		Yes	Yes						Yes					Yes	

Fourier descriptor, Elementary descriptor, Bounding box and Template matching. Texture features are described as Atomic texture features, wavelet, gabor, fourier edge statistics and random fields etc.The eigen imae is an example of dominant color, because it's feature is copied from the value of global image color. The features of atomic texture are represents gap of an image, regularity, directionality and smoothness of an image. The edge statistics represents histogram of image and orientation. The centroid, quarter, orientation axix, strangeness are the kinds of elementary shape descriptors. The layout feature represents the total or partial locality of the color, shape and texture.

The simple features can be easily mined, and it is easy to apply in the CBIR systems, and the easiest to use in the accurate method At last, color feature is established effective, for the reason that good quality color feature is not very complicated to design and implement. As retrieval of images is very challenging task so .

Researchers have proposed and used various CBIR Systems according to various characteristics of systems. These CBIR Systems perform the task of retrieving the images based on the various features available in it. Some systems retrieve the images by using only color, texture or shape and some uses only keywords or metadata to retrieve the images there are very less number of systems which uses both low level and high level features to retrieve the images in effective manner.

From table 1 it can be shown that QBIC is the most effective system that is widely used by the researchers and if we talk about the features in this system then it used all the features like color, texture,shape, layout as well as keyword based retrieval of images. According to the features used by various CBIR Systems, QBIC uses both high level features, low level features as well as keywords for retrieving images according to the requirements of user.

6. CONCLUSION

The techniques of CBIR are still under beneath investigation. To increase the performance of image retrieval capability of systems of CBIR are used. In this paper we have discussed many features used in CBIR systems. It can be drawn from the analysis that the wide varieties of features are used by the QBIC CBIR systems. Still this area needs more research for finding a best system for image retrieval process so that images could be retrieved in effective manner. Most systems are under research is the main product of these systems.

7. REFERENCES

[1] Bach J.R., Fuller C., Gupta A., Hampapur A., Horowitz B., Humphrey R., Jain R.C. and Shu C.F., 1996. Virage image search engine: an open framework for image management. In Electronic Imaging: Science & Technology (pp. 76-87). International Society for Optics and Photonics.

[2] Balan S., and T. Devi. 2012. Design and Development of an Algorithm for Image Clustering In Textile Image Retrieval Using Color Descriptors." International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol. 2(3).

[3] Benitez Ana B., Mandis Beigi, and Shih-Fu Chang. 1998. Using relevance feedback in content-based image metasearch." Internet Computing, IEEE vol. 2(4) ,pp.59-69.

[4] Berman Andrew P. and Linda G. Shapiro.1998. A flexible image database system for content-based retrieval. Pattern Recognition,. Proceedings. Fourteenth International Conference on. Vol. 1. IEEE.

[5] Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M., & Malik, J. 1999. Blobworld: A system for region-based image indexing and retrieval. Visual Information and Information Systems. Springer Berlin Heidelberg.

[6] Cerra D. & Datcu, M. 2012. A fast compression-based similarity measure with applications to content-based image retrieval. Journal of Visual Communication and Image Representation, vol. 23(2), pp. 293-302.

[7] Chadha A., Mallik S. and Johar, R. 2012. Comparative study and optimization of feature-extraction techniques for content based image retrieval, arXiv preprint arXiv:1208.633.

[8] Chadha D. and Singh N. 2012. Query by Image for efficient Information Retrieval- A Necessity, International Journal of Computer Applications.

[9] Chary R., Lakshmi, D. R. & Sunitha, K. V. N. ,2012. Feature extraction methods for color image similarity. arXiv preprint arXiv:1204.2336.

[10] Gevers T., & Smeulders A. W., 2000. PicToSeek: combining color and shape invariant features for image retrieval. Image Processing, IEEE Transactions on, vol. 9(1), pp.102-119.

[11] Huang T. Rui, Y., & Chang, S. F. 1997. Image retrieval: Past, present, and future ,In International Symposium on Multimedia Information Processing (Vol. 108).

[12] Kalel, M. D. S., Pisal, M. P. M., & Bagawade, M. R. P.: 2016. Color, Shape and Texture feature extraction for Content Based Image Retrieval System: A Study, International Journal of Advanced Research in Computer and Communication Engineering, vol. 5(4).

[13] Kelly P. M., Cannon, T. M., & Hush, D. R. 1999. Query by image example: the comparison algorithm for navigating digital image databases (CANDID) approach. In IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology (pp. 238-248). International Society for Optics and Photonics, March 1

[14] Khokher A. & Talwar R. 2011. Content Based Image Retrieval: State-of-the-Art and Challenges. International Journal of Advanced Engineering Sciences and Technologies, vol. 9(2), pp. 207-211.

[15] La Cascia M. and Ardizzone, E. 1996. JACOB: Just a content-based query system for video databases" In Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings. IEEE International Conference on IEEE., Vol. 2, pp. 1216-1219.

[16] Ma W. Y. and Manjunath B. S. 1999.Netra: A toolbox for navigating large image databases. Multimedia Systems,vol. 7(3),pp. 184–198.

[17] Manjunath B. S. 1995. Image Browsing in the Alexandria Digital Library (ADL) Project D-Lib Magazine, vol. 1(2).

[18] Mukherjea S., Hirata, K., & Hara, Y. 1999. Amore: A world wide web image retrieval engine. World Wide Web, vol. 2(3), pp. 115-132.

[19] Nalini M. P. & Malleswari B. L. 2016. Review on Content Based Image Retrieval: From Its Origin to the New Age., International Journal of Research Studies in Science, Engineering and Technology Vol. 3(2), pp. 18-41.

[20] Ogle V. E., & Stonebraker, M. 1995.Chabot: Retrieval from a relational database of images. Computer, vol. 28(9),pp. 40-48.

[21] Pattanaik, S. Prof. DG Bhalke. Beginners to Content Based Image Retrieval. International Journal of Scientific Research Engineering & Technology, 1, 040-044.

[22] Pentland A., Picard, R. W., & Sclaroff, S. 1996. Photobook: Content-based manipulation of image databases. International journal of computer vision, vol. 18(3), pp. 233-254.

[23] Smith J. R. & Chang, S. F. 1997. Querying by color regions using the VisualSEEk content-based visual query system. Intelligent multimedia information retrieval, vol. 7(3), pp. 23-41.

[24] Srihari R. K., Zhang, Z., & Rao, A. 2000. Intelligent indexing and semantic retrieval of multimodal documents". Information Retrieval, vol. 2(2-3), pp. 245-275.

[25] Su, J. H., Huang W. J., Yu P. S. & Tseng, V. S. 2011. Efficient relevance feedback for content-based image retrieval by mining user navigation patterns. Knowledge and Data Engineering, IEEE Transactions on, vol. 23(3), pp. 360-372.

[26] Wayne, N., Ron, B., Will, E., Flickner, M. D., Glasman, E. H., Petkovic, D.,& Taubin, G. 1993. The QBIC project: Querying images by content using color, texture, and shape. In Proc. SPIE ,Vol. 1908, No. 1, pp. 173.