A Review on Content Based Image Retrieval with Different Techniques

Ashutosh Gupta[1], Abhishek Verma[2]
[1, 2] Department of Information Technology
[1, 2] Galgotias College of Engineering and Technology

Abstract- *Content-based image retrieval is a very important area of research nowadays. Content Based mage Retrieval (CBIR) is a technique which uses visual features of image such as color, shape, texture, etc. CBIR technologies provide a method to find images in large databases by using unique descriptors from a trained image. A lots of research works had been completed in the past decade to design efficient image retrieval techniques from the image or multimedia databases.*

Keywords- CBIR, Color, Texture, Shape. Dwt, Svm.

I. INTRODUCTION

All Advancement in internet and multimedia technologies has imposed to have a system that organizes the large digital images for easy categorization and retrieval. Many researchers in information-technology field and leading academic institutions try to develop content based image retrieval system for very large image database. However, we cannot access or make use of this information unless it is organized to allow efficient browsing, searching, and retrieval. One of the main problems is the difficulty of locating a desired image in a large and varied collection. While it is perfectly feasible to identify a desired image from a small collection simply by browsing, more effective techniques are needed with collections containing thousands of items. Image retrieval attracts interest among researchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, database applications, and other related areas [1].

Content Based Image Retrieval (CBIR) is the method of retrieving images from the large image databases as per the user demand. It is also known as Query By Image Content (QBIC) and Content Visual Information Retrieval (CBVIR). In CBIR, content based means the searching of image is proceed on the actual content of image rather than its metadata. The Content Based Image Retrieval System is used to extract the features, indexing those features using appropriate structures and efficiently provide answers to the user's query. To provide the satisfactory answer to the user query, CBIR provides some flow of work. Firstly CBIR system takes the RGB image as an input, performs feature extraction, performs some similarity computations with the images stored in database and retrieves the output image on the basis of similarity computation[1].

II. CBIR FRAMEWORK

Basic idea behind CBIR is that, when building an image database, feature vectors from images (the features can be color, texture, shape, region or spatial features, etc.) are to be extracted and then store the vectors in another database for future use [2]. General image retrieval system is shown by Figure 1, is consists of three main modules, input module, query module, retrieval module.

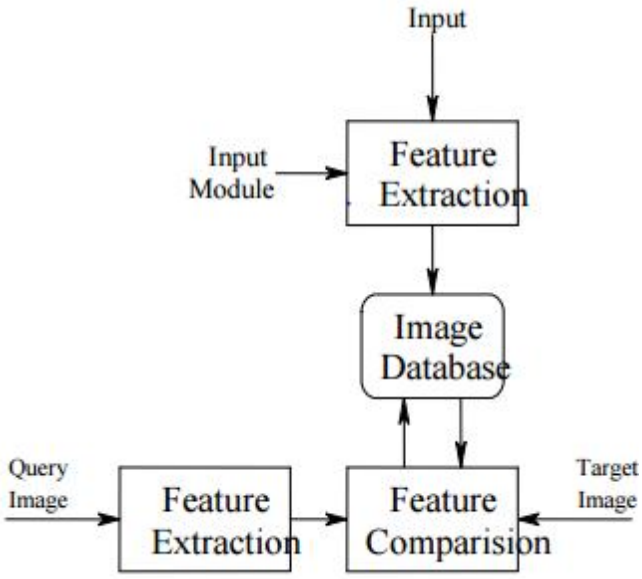

Fig 1: Framework of CBIR

In input module, the feature vector is extracted from each input image and stored into the image database with its input image. When query image is entered into the query module, the feature vector of the query image is extracted. In retrieval module, the extracted feature vector of query image is compared with the images stored in the database. Similar images are retrieved according to their similarity with the query image. Finally the target image will be obtained from the retrieved images [2].

III. APPLICATION OF CBIR

- CBIR is popular for police force for picture recognition in crime prevention.
- Medicine Diagnosis.
- Architectural and Engineering Design.
- Fashion and Publishing.
- Geographical information and remote sensing.
- Home Entertainment.

IV. FEATURE EXTRACTION

Features are divided into two categories respectively text based and visual based. Textual features are keywords, tags, annotations etc. Visual features are color, space and texture etc. Visual features are the important features of an image for pattern recognition.

A. Color:

color describes one of the important visual features in content based image retrieval. There are number of examples, where color features in retrieving image are used like histograms, moments, block-based. Color histogram is used for computing distance measures based on color similarity for each image. A color histogram is used to describe the global color distribution in an image and is more frequently used method because of its advantages like high efficiency. Other feature representation like color moments and color sets are also used than color histogram. The color is a widely used important feature for image representation. This is very important as it is invariant with respect to scaling, translation and rotation of an image [3]. Color space, color quantification and similarity measurement are the key components of color feature extraction. Color feature is not dependent upon size of image. The color models can be classified as User & Hardware based models; such as RGB and HSV. Many color spaces are there which offers different applications.

a) Color Histogram:

The color histogram depicts color feature which cannot capture color distributions or textures within the image. In this method color histogram feature is divided into global and local color extraction. Using 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 histograms [4]. Local color histogram gives spatial information. Local color histogram also gives the information related to the color distribution of regions. The first step is to divide the image into segment and then to obtain a color histogram for each block then image will 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 [4].

b) Geometric Moments:

In image processing, an image moment is ascertain particular weighted average (moment) of the image pixels' intensities, or a function of such moments, usually chosen to have some attractive property or interpretation [4]. Image moments are generally used to describe objects after segmentation. This feature use only one value for the feature vector, which mean that when the image size becomes relatively large, computation of the feature vector takes a large amount of time and the pros of using this feature combine with other features such cooccurrence, which can provide a better result to user [4].

c) Color Moments:

Color moments are the statistical• moments of the probability distributions of colors and have been successfully used in many retrieval systems, especially when the image contains just the object, it means color moment will work best when image has only object. Three parameters are calculated in this method: Mean, Variance and Skewness. Color moments have been proved to be efficient and effective in representing color distributions of images and it suffer from the problem that they fail to encode any of the spatial information surrounding the color within the image [5].

B. Texture:

Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. It is an inherent property of virtually all surfaces including clouds, trees, bricks, hair, and fabric. Texture provides useful information of the surfaces about their structures and the relationship with the surrounding. texture analysis can be studied at three levels i.e. on statistical level, a set of statistics extracted from the image is called texture. On the structural level, the primitives of the image and their placement rules are known as its texture. On the spectral level, the texture is defines as a set of coefficients in the transform domain. With the help of these levels the textures can be identified but the textures may not agree with human way of evaluating the textures.[6] These reasons are semantic gap and human perception subjectivity. Texture feature describes spectral features which are taken using wavelet transform, statistical features, tamura texture features etc. Tamura explored the texture representation from a different viewpoint. [7]Texture and color queries can be formulated in similar way, by selecting desired textures or by supplying an query image.

a) Discrete Wavelet Transform:

The Wavelet Transform is created by repeatedly filtering the image coefficients on a row by row and column by column basis. A two-dimensional DWT decomposition of image contains various band information such as low-low frequency approximation band, high-low frequency vertical detail band, low-high frequency horizontal detail band and high-high frequency diagonal detail band. Discreate wavelet transform is used to obtain good image retrieval base on the low computational cost. Discrete Wavelet Transform (DWT), which transforms a discrete time signal to a discrete wavelet representation[8].

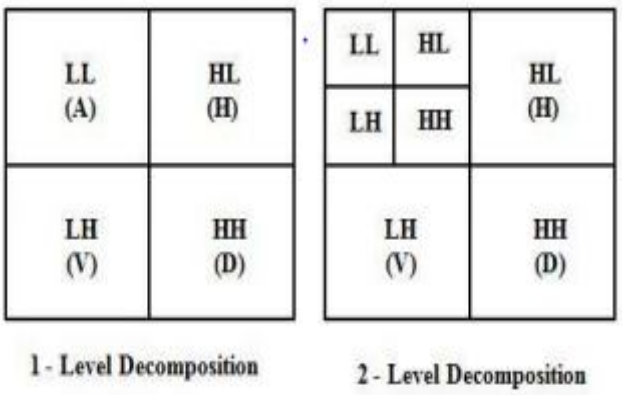

Fig 2: Image Decomposition

Wavelet Feature Extraction

Wavelet ψ (t) is a function that satisfy the following condition

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad . \qquad (2)$$

Discrete Wavelet Transform (DWT) is a spectral estimation technique, which decompose a function or a signal into different frequency sub-bands and leads to a set of wavelet coefficients. The low level physical features of the original image are buried in these wavelet coefficients; therefore the signal can be characterized by these wavelet coefficient[8].

b) Gaussian pyramid:

It is used to decompose images into information at multiple scales, to extract features and to remove noise from the image. It consists of low-pass filtered, down-sampled version of the previous level of the pyramid, where the base level is defined as the original image [9].

c) Fourier transform:

In this method, a signal is• decomposed into a number of sinusoids of different frequency. The Fast Fourier Transform (FFT) refers to a class of algorithms for efficiently computing the Discrete Fourier Transform (DFT) hence FFT is not an approximation of the DFT, it is the DFT with a reduced number of computations. DFT is most useful in digital signal processing, Convolution and digital filtering [9].

C. Shape:

Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out. In image retrieval, depending on the applications, some require the shape representation to be invariant to translation, rotation, and scaling, while others do not. Shape features of objects or regions have been used in many content-based image retrieval systems. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. Shape features are divided into two categories boundary based and region based. Boundary based shape features uses only boundary of the shape whereas region-based shape features uses entire shape region[5]. The term shape refers to the information that can be deduced directly from the image. Shape is represented through perceptually grouped geometric cues such as edges, contours, joints, and polygonal regions extracted from an image. Such a grouping can serve as a spatial layout or as a rough sketch by additional post processing. Shape features are known as geometric features. shape feature are commonly used – global features such as aspect ratio, circularity and moment invariants and local features such as sets of consecutive boundary.[3]

a) Fourier Descriptors:

They are obtained by applying Fourier transform on shape boundary, the Fourier transformed coefficients are called the Fourier descriptors (FD) of the shape. For good shape description, an appropriate shape signature is essentially required to obtaining FD. The nice properties of FDs are its robustness and easy to derive. With Fourier descriptors, coarse shape features or global shape features are captured by lower order coefficients and the finer shape features are captured by higher order coefficients and Noise is not a problem with Fourier descriptors. With fast Fourier transform (FFT), the computation is efficient [10].

b) CSS Descriptors:

They are essentially the descriptors of key local shape features. Not only the locations of, but also the degree of concavities of shape boundaries are detected by dealing shape in scale space. These features are very important to human perception in judging the similarity between shapes. The dimension of CSS descriptors is very low. It only captures

the local shape features, the global features which are also important to shape representation are missed out from the representation. There may be no CSS descriptors for smooth convex shapes such as polygon composed of straight lines [10].

c) Zernike moments:

This method allows independent moment invariants to be constructed to an arbitrarily high order. Zernike moments descriptor does not need to know boundary information, making it suitable for more complex shape representation [10]. Zernike moments descriptors can be constructed to arbitrary order like Fourier descriptors, this overcomes the drawback of geometric moments in which higher order moments are difficult to construct[10].

Support Vector Machine

Vapnik proposed Support Vector Machines (SVM) in 1979 (Vapnik, 1995)[11], but they have only recently been gaining popularity in the learning community. The main idea of an SVM is to construct a hyper plane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. The notion that is central to the construction of the support vector learning algorithm is the inner product kernel between a support vector Xi and a vector X drawn from the input space. The support vectors constitute a small subset of the training data extracted by the support vector learning algorithm. The separation between the hyperplane and closet data point is called the margin of separation, denoted by ρ.

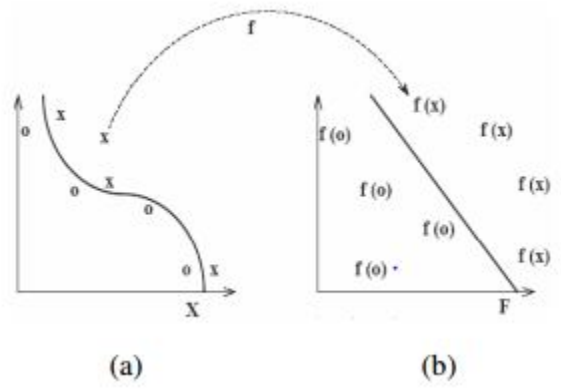

Figure 3. The function 'f' embeds the data in the original space

The goal of a support vector machine is to find a particular hyper plane for which the margin of separation ρ is maximized. Under this condition, the decision surface is referred to as the optimal hyper plane. The support vectors play a prominent role in the operation of this class of learning machines. In conceptual terms, the support vectors are those data points that lie closest to the decision surface, and therefore the most difficult to classify. They have a direct bearing on the optimum location of the decision surface. The idea of an SVM is based on the following two mathematical operations:

1. Nonlinear mapping of an input pattern vector on to a higher dimensional feature space that is hidden from both the input and output.
2. Construction of an optimal hyper plane for separating the patterns in the higher dimensional space obtained from operation1.

V. LITERATURE REVIEW

Amreen Posharkar et al (March 2015) [12] they proposed a system that capable to serve a hybrid approach to combined CBIR and e-commerce website to retrieval quality products for the user based on internal This system also allowed the user to jot down his/her thoughts since the system enables the user to draw the sketch of the desired product. So ,the result of the search returned will be of good quality demanded by the user. For every search, relevant results will be returned to the user. User's feedback will be taken into consideration so that more relevant search can be provided. Hence the proposed system provided shopping in a more fun loving manner and at the same time will be more efficient to the user.

Supreet Sidhu et al (May 2015) [13] in this paper they focused on. to the development of content based image retrieval system i.e. CBIR. CBIR would retrieve images on the basis of visual content of the image like colour, texture, shape etc. rather than textual annotation. The desirable features of CBIR system include reducing the retrieval time and increasing the efficiency of the system. Thus a system needs to be developed that can automatically extract relevant objects from a huge dataset.

Gurmeet Kaur et al.(July 2015) [14] In this paper author presents content based image retrieval from color digital images using enhanced SVM techniques .They proposed The dramatic rise in the sizes of images databases has stirred the development of effective and efficient retrieval systems. The development of these systems started with retrieving images using textual connotations but later introduced image retrieval based on content. This came to be known as Content Based Image Retrieval or CBIR. Systems using CBIR retrieve images based on visual features such as texture, colour and shape, as opposed to depending on image descriptions or textual indexing. The main objective of this paper is to retrieve the images from database in a fast and an

efficient manner using modified Support vector method(SVM).

Satish Tunga et al.(December, 2015)[15] presented a comparative study of CBIR techniques This paper presents a brief survey on work related to the exciting fields of content-based image retrieval and provides a detailed review of the works carried out in this field. This paper also discussed the various methodologies used for extracting the salient low level features and various distance measures to find the similarity between images in reducing the semantic gap between the low level features and the high level semantic concepts. A discussion of various approaches of CBIR and comparison of various techniques with respect to data are also made.

Pushpendra Singh et al.(September, 2015)[16] described how shape descriptor are classified into two methods boundary-based and region based and shape descriptor are further divided into classified structural and global descriptor for image retrieval. They also focused on two main shape descriptors a) simple shape descriptors b) Fourier descriptor.

Savita Gandhavni et al (October, 2015)[17] they compared compared different color, texture and shape feature extraction methods that are most popularly used in image understanding studies. Our comparison shows that there is considerable performance variability between the various feature extraction methods. One of the features of this study is the use of a publicly available benchmark that further studies can use.

Roshni S. Tadse et al(December, 2015)[18] The proposed method infer that mobile document images are sending to the server and retrieved the document of that particular image. An efficient method of retrieving document images from content based information retrieval and searching methods that search the information with respect to content of images. Also getting the optimized output at the user end in an efficient manner. JPEG compression is there, to low complexity is introduced to reduce the query delivery latency while maintaining comparable search accuracy.

VI. COMPARISON OF VARIOUS TECHNIQUES

TABLE I. Comparison of The Various Techniques

Feat-ures	Tech-niques	Accur acy	Dime nsion	Adv-antage	Dis-Advantage
Color Feature	Color Moment	Low	Low	Lower computatio nal complexity	Precision is low
	HSV histogram	High	Mediu m	Simple ,Fast computatio n	No spatial information
Texture Feature	Gabor Filter	High	High	Achieves highest retrieval results	Computationall y intensive
	Gabor Moment	Low	Low	Lower Dimension ality	Low retrieval result compare to Gabor filter
Shape Feature	Moment Invariant	High	Low	Invariable to translation, rotation and scale	Limited recognition power
	Zernike moments	High	Low	Invariable to translation, rotation and scale	Computational Complexity is High

VII. PERFORMANCE MEASUREMENT

a) Euclidean Distance:

In the process of a query image compute the distance between the transformed feature vector of each image of the query image and that of 1000 images in the database this distance is measured as:

$$dist(q,d) = \frac{1}{10}\sqrt{\sum_{k=1}^{10}\left[dist(q_k,d_k\right]^2} (13)$$

Where q is query image and d is distance.

b) City block distance:

In this we are getting the value which is greater than or equal to 0.If 0 comes then identical points are considered, and when it is greater than 0 there it is low similarity.

$$dist(q,d) = \frac{1}{10}\sum_{k=1}^{10}\left[dist(q_k,d_k\right] (14)$$

c) Precision

Precision is defined as the ratio of the number of retrieved relevant images to the total number of retrieved images. We denote the precision by P [3]

$$P = \frac{\text{Number of relevant images are retrived}}{\text{Total number of images retrived}}$$

d) Recall

Recall is defined as the ratio of the number of retrieved relevant images to the total number of relevant images in the database. We denote to the recall by R [3].

$$R = \frac{\text{Number of relevant images are retrived}}{\text{Total number of relevant images in database}}$$

VIII. CONCLUSION

There are many CBIR systems developed across the world. They all use features like color, shape, texture and point of interest individually or in combination. We have focused on some important feature aspects related to CBIR. If these features are used individually they give results accordingly, which means if only color feature is used, images having same color features will be retrieved. In this paper literature of different content retrieval method is discussed like SVM based retrieval, SVM with relevance feedback method, DWT based method etc in which some of the methods are efficient to shorten the semantic gap between the image while some are less so in future work need to develop such technique which much efficiently and effectively reduces the semantic gap and increases the information gain.

REFERENCES

[1] Priyanka Malode, Prof. S. V. Gumaste "A Review Paper on Content Based Image Retrieval" International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 09 | Dec-2015.

[2] Sagar Soman, Mitali Ghorpade, Vrushali Sonone and Satish Chavan. Article: Content Based Image Retrieval using Advanced Color and Texture Features. IJCA Proceedings on International Conference in Computational Intelligence (ICCIA2012) iccia(9):-, March 2012. Published by Foundation of Computer Science, New York, USA.

[3] Henning Müller, Nicolas Michoux, David Bandon, Antoine Geissbuhler (2004), ―A review of content-based image retrieval systems in medical applications–―clinical benefits and future directions‖, International Journal of Medical Informatics 73, 1—23.

[4] Aman Chadha, Sushmit Mallik, and Ravdeep Johar, "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval," International Journal of Computer Applications (0975 – 8887) Volume 52– No.20, August 2012, pp.35-42. [7] P.S.SUHASINI, Dr. K.SRI RA

[5] Amandeep Khokher, and Rajneesh Talwar, "Content-based Image Retrieval: Feature Extraction Techniques and Applications," Proceedings published in International Journal of Computer Applications, pp.9-14

[6] Chiu, C.Y., Yang, S. and Lin, H.C.(2002), ―Learning human perceptual concepts in a fuzzy content based image retrieval system‖, 5th International Conference on Computational intelligence and Multimedia Applications, pp. 330-335.

[7] Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma.(2007),‖ A survey of content-based image retrieval with high-level semantics‖ The Journal of the pattern recognition society. Pattern Recognition 40 (2007) 262 – 282

[8] Jeyanthi Prabhu and Jawahar Senthil Kumar "Wavelet Based Content Based Image Retrieval Using Color And Texture Feature Extraction By Gray Level Coocurence Matrix And Color Coocurence Matrix" Journal of Computer Science 10 (1): 15-22, 2014.

[9] Swapna Borde and Udhav Bhosle, "Feature Vectors based CBIR in Spatial and Transform Domain," International Journal of Computer Applications (0975 – 8887), Volume 60– No.19, December 2012, pp.34-42.

[10] Dengsheng Zhang and Guojun Lu, "Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study".

[11] Kommineni Jenni, Satria Mandala, Mohd Shahrizal Sunar "Content Based Image Retrieval Using Colour Strings Comparison" Content Based Image Retrieval Using Colour Strings Comparison

[12] Amreen Posharkar, Shabina Sayed, Sneha Jha and Amruta Jaitpal, Content Based Image Retrieval in E-Commerce for Quality Products, International Journal of Advanced Research in Computer Science and Software Engineering 5(3), March - 2015, pp. 706-708

[13] Supreet Sidhu and Jyoti Saxena, Content Based Image Retrieval A Review, International Journal Of Research In Computer Applications And Robotics, Vol.3 Issue.5, Pg.: 84-88 ,May 2015

[14] Gurmeet Kaur and Er. Arshdeep Singh Content Based Image Retrieval from Colored Digital Images using

Enhanced SVM Technique, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 7, July 2015.

[15] Satish Tunga, D. Jayadevappa and C. Gururaj, A Comparative Study of Content Based Image Retrieval Trends and Approaches, International Journal of Image Processing (IJIP), Volume (9) : Issue (3) : 2015 Pg.: 127-155.

[16] Review on Shape based Descriptors for Image Retrieval, International Journal of Computer Applications (0975 – 8887) Volume 125 – No.10, September 2015

[17] Savita Gandhani and Nandini Singhal ,Content Based Image Retrieval: Survey and Comparison of CBIR System Based on Combined Features, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.8, No.10 (2015), pp.155-162

[18] Roshni S. Tadse et al, International Journal of Computer Science and Mobile Computing, Vol.4 Issue.12, December- 2015, pg. 211-214.