Comparative Study: Content Based Image Retrieval using Low Level Features

Ankita Sharma[1], Gurpreet Singh[2]
Department of Electronics and Communication, Lovely Professional University[1,2]

Abstract

Image retrieval plays an important role in many areas like architectural and engineering design, fashion, journalism, advertising, entertainment, etc. How to search and to retrieve the images that we are interested in is a fatal problem: it brings a necessity for image retrieval systems. As we know, visual features of the images provide a description of their content. Content-based image retrieval (CBIR), emerged as a promising mean for retrieving images and browsing large images databases. It is the process of retrieving images from a collection based on automatically extracted features by generally using low level features. To improve CBIR, human perception i.e. high level feature extraction can be included for better efficiency.
In this paper, a comparative analysis is performed so as to achieve efficient results.

Keywords:CBIR, Low Level Features, High Level Features, Feature Vector

I. INTRODUCTION

Recent advances of technology indigital imaging and digital storage devices make it possible to easily generate, transmit, manipulate and store large number of digital images and documents. As a result of advances in the Internet and new digital image sensor technologies, the volume of digital images produced by scientific, educational, medical, industrial, and other applications available to users increased dramatically. Various methods and algorithms have been proposed for imageaddressing. Such studies revealed the indexing and retrieval concepts which have further evolved to Content Based Image Retrieval (CBIR).Concept of CBIR evolved in early 1990's.

Content Based Image Retrieval (CBIR) is a technology that helps to organize and retrieve digital image database by their visualcontent. It is the process of retrieving images from a collection based on automatically extracted features by generally using low level features. CBIR evolved as an advancement of Text Based Image Retrieval (TBIR) which used annotations to describe images. In it images were indexed using keywords, headings, codes, which were then used as retrieval keys. It suffered from many problems like; (1) it was non-standardized, because different users used different keywords for annotation, (2) It remained sometimes incomplete. (3) Humans were required to personally describe every image, therefore, it was cumbersome and labour intensive.Due to the above mentioned disadvantages, CBIR proved to be more promising and efficient than TBIR.

CBIR basically uses the visual contents of an image such as colour, shape, texture, and spatial layout to represent and index the image. In typical content-based image retrieval systems, the visual contents of the images in the database are extracted and are then described by feature vectors. The feature vectors of the images in the database form a database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities between the feature vectors of the query example and those of the images in the database are then measured and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database. Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process in order to generate semantically more meaningful retrieval results [2].

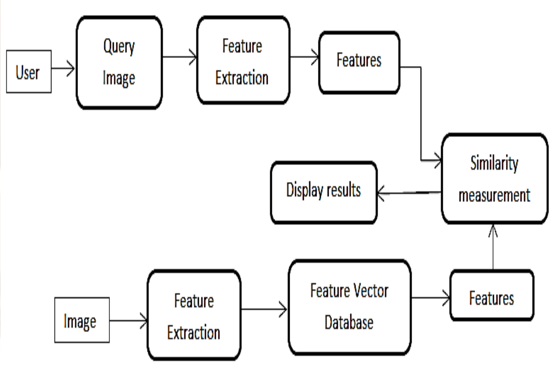

Figure 1: Content Based Image Retrieval System

The figure 1 shows how CBIR system works. Feature vector of both the query image and the database image is firstly calculated and then the similarity measurement is done between the two. The database image which matches the most with the query image is ranked as first and so on. A point to focus upon is that the retrieval result is not a single image; rather it is number of ranked images.

II. IMAGE CONTENT DESCRIPTORS

Feature

It can be defined as the representation of any distinguishable characteristic of an image. It can be classified into 3 levels:

1. Low Level: It includes colour, texture, shape, spatial information and motion.
2. Middle Level: It includes arrangement of particular types of objects, roles and scenes.
3. High Level: It includes impressions, emotions and meaning associated with the combination of perceptual features.

Feature Vectors

Extracting the features of an image and calculating its feature vector is one of the most important tasks in CBIR. Generally, low level features are extracted and calculated. These include:

1. Color: Color reflects the chromatic attributes of the image. A range of geometric color models (e.g., HSV, RGB, Luv) for discriminating between colors are available. Color histograms are amongst the most traditional technique for describing the low-level properties of an image. Color Coherent Vectors (CCV) and Color Moments can also be used to calculate the color feature vector.

2. Texture: It is another important property of images. Various texture representations have been investigated in pattern recognition and computer vision. Basically, texture representation methods can be classified into two categories: statistical and transform based methods. Statistical methods include calculating the Gray level Co-occurrence Matrix (GLCM) which further calculates the Energy, Entropy, Contrast and Inverse Difference moment. Transform Based Methods include calculating the Gabor transform or wavelet transform etc.

3. Shape: Shape features of objects or regions have been used in many content-based image retrieval systems. Compared with colour and texture features, shape features are usually described after images have been segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to special applications where objects or regions are readily available. The state-of-art methods for shape description can be categorized into either boundary-based or region-based methods. A good shape representation feature for an object should be invariant to translation, rotation and scaling.

4. Spatial: Regions or objects with similar colour and texture properties can be easily distinguished by imposing spatial constraints. For instance, regions of blue sky and ocean may have similar colour histograms, but their spatial locations in images are different. Therefore, the spatial location of regions or the spatial relationship between multiple regions in an image is very useful for searching images.

III. SIMILARITY MEASUREMENTS

For comparing the similarities between the query image and the database image Euclidean distance formula can be used.

$$\text{Similarity}(Q, DB) = \sqrt{\sum_{n=1}^{16}(Q - DB)^2} \quad \cdots(1)$$

Where Q is the query image and DB is the data base image.

Other formulas for distance measurement which can be used are:

Minkowski-Form Distance: If each dimension of image feature vector is independent of each other and is of equal importance, the Minkowski-form distance is appropriate for calculating the distance between two images [12].

Quadratic Form Distance: The Minkowski distance treats all bins of the feature histogram entirely independently and does not account for the fact that certain pairs of bins correspond to features which are perceptually more similar than other pairs. To solve this problem, quadratic form distance was introduced.

IV. PERFORMANCE EVALUATION

For evaluating the performance of the CBIR system, the value of Precision and Recall are considered. However, these two parameters don't give information about ranking order of the image. Therefore NMRR (Normalized Modified Retrieval Rank) is taken into account.

Precision: It is defined as the ratio of the number of relevant images retrieved to the total number of images retrieved. It is a measure of accuracy of retrieval and is given as:

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of images retrieved}} \quad \ldots(2)$$

Recall: It is defined as the number of relevant images retrieved to relevant images available in the database.

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{No. of relevant images available in the database}} \quad \ldots(3)$$

NMRR: It stands for Normalized Modified Retrieval Rank. It is basically the normalized MRR score. It gives information about the retrieval rank of the images. Its value lies in between 0 and 1. The smaller the values are, better are the results.

$$\text{NMRR} = \frac{\text{MRR}(q)}{K + 0.5 - 0.5 * NG(q)} \quad \ldots(4)$$

V. ANALYSIS OF SOME LOW LEVEL FEATURE BASED TECHNIQUES

1. CBIR using advanced Color and Texture Features [1]

This presents an efficient Content Based Image Retrieval (CBIR) system using color and texture. Here, two different feature extraction techniques are employed. A universal content based image retrieval system uses color, texture and shape based feature extraction techniques for better matched images from the database. In proposed CBIR system, color and texture features are used. The texture features were extracted from the query image by applying block wise Discrete Cosine Transforms (DCT) on the entire image and from the retrieved images the color features were extracted by using moments of colors (Mean, Deviation and Skewness) theory.

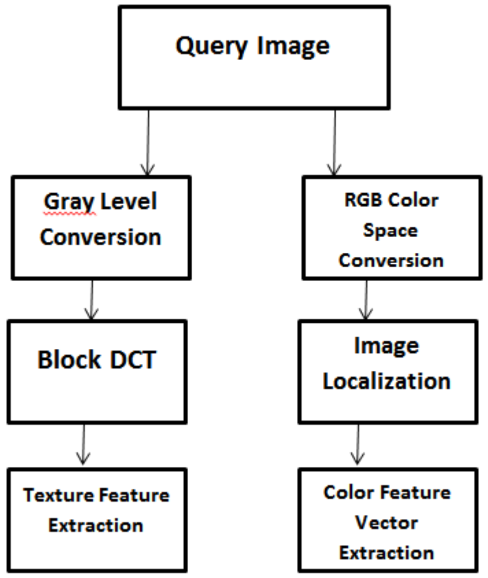

Figure 2: Block Diagram of Feature Extraction [1]

Color Feature Extraction: Color is one of the most widely used visual features in CBIR because it is relatively robust. As shown in the block diagram, the RGB components of the image are extracted. They are then divided into block sizes of 8x8 each and the three moments are retrieved using the equations given below[1]:

$$E_i = \frac{1}{mn}\sum_{j=1}^{mn} P_{ij} \ldots (5)$$

$$\sigma_i = \frac{1}{mn}\left[\left[\sum_{j=1}^{mn}(P_{ij} - E)^2\right]^{\frac{1}{2}}\right] \ldots (6)$$

$$\alpha_i = \frac{1}{mn}\left[\sum_{j=1}^{mn}(P_{ij} - E)^3\right]^{1/3} \ldots (7)$$

Where E_i is an average of each color channel (i= R, G, B), σ_i is a standard deviation, α_i is a skewness. P_{ij} is a value of each color channel at jth image pixel. m.n are the total number of pixels per image.

Texture Feature Extraction: For efficient texture feature calculation, we can use DCT coefficients [5].

Figure 3: DCT-9 Feature Vectors [1].

For each sub block containing one DC coefficient and other AC coefficients, we extract a feature set of 9 vector components by averaging specific regions of coefficients as shown by figure 4, considering following characteristics:

1) The DC coefficient of each sub block represents the average energy of an image (vector component 1).

2) All remaining coefficients within a sub block contain frequency information which represents a different pattern related to image variation (vector components 2 to 4).

3) The coefficients of some regions within a sub block also represent some directional information (vector components 5 to 9).

For DCT transform first we convert RGB image into gray level image. We then used block based DCT transformation. The equation used for the DCT calculation for each pixel is given as follows:

$$F(u,v) = \frac{c_u c_v}{4}\sum_{i=0}^{7}\sum_{i=0}^{7} f(i,j)\cos\frac{(2i+1)u\Pi}{16}\cos\frac{(2i+1)v\Pi}{16} \ldots (8)$$

Where f (i, j) is the pixel value at the (x, y) coordinate position in the image, F (u, v) is DCT domain representation of f (i, j), where u and v represent vertical and horizontal frequencies, respectively.[1]

RESULTS

There are three different techniques for retrieval of images. They are as follows:

1) Using Color (HSV and YCbCr) And Texture
2) Using only Color (HSV and YCbCr)
3) Using only Texture

Amongst all three techniques the first one 'Using Color (YCbCr) And Texture' is giving very efficient results than any other technique. It retrieved 60% of total available images in the database which are similar to given query image.

The results based on only color and texture were approximately 40% [1].

2. CBIR using Pulse Coupled Neural Networks (PCNN) [2]

In this technique a method for CBIR based on PCNN is shown. The PCNN is originally proposed as a model of biological neural network in the cat visual cortex, which can reproduce synchronous activities in the brain [6][7][8][9].

The method of the image matching using the PCNN had been proposed in [10]. In this conventional study, the method utilizes the PCNN-Icon to achieve image matching. Where the PCNN-Icon is obtained from PCNN firing activity when the image data inputs to the PCNN. The PCNN-Icon is a time series of the number of firing neurons, and it is obtained by the observation of neuron firing in the PCNN.The most important characteristics of the PCNN-Icon is that the correlation between PCNN-Icons of same or similar images has large value. One of the remarkable features of this method is that the method can recognize the rotated, or scale modified image.

This image matching method using PCNN is achieved by means of the PCNN-Icon. The PCNN-Icon is defined as the time series of the number of the firing neurons in the PCNN. To obtain the PCNN-Icons,a number of the firing neurons are observed in every time step from $t = 0$ to $t = t_{max}$, where t_{max} is defined arbitrary. Namely, the PCNN-Icon is also defined as a t_{max}-dimensional natural number vector. Where we assumed that the t_{max} is 100 in this study. It is empirically known that the PCNN-Icon is unique to the input image.

To retrieve the images, the system calculates the correlation among the PCNN-Icons of the images in the database and the query image. The way of presentation of the results is simply assumed that the system presents the image data in descending order of the correlation between the PCNN-Icon of the image in the database and that of the query image.

RESULTS

In the results of the image retrieval using our CBIR system, it was shown that the images in relevant category were listed on top of the ranking. The characteristics of retrieved images have also been evaluated, and it is shown that the image recognition method using PCNN is valid for the CBIR system [2].

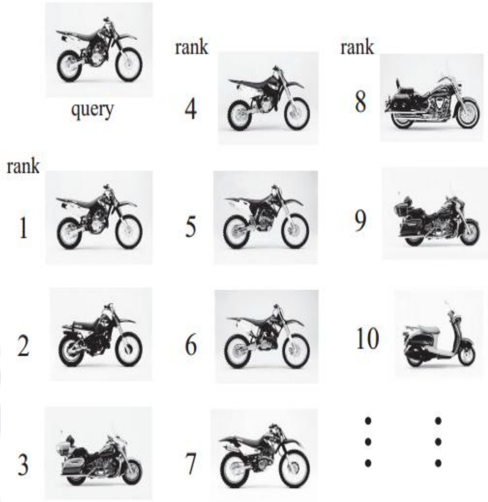

Figure 4: Result of CBIR using PCNN

3. CBIR using Shape Feature [3]

The focus of this technique is on shape based image retrieval. Like any other features based on human perception, the major problem in the use of shape is how to represent shape information. The proposed technique is a contour based approach. Canny operator is used to detect the edge points of the image. The contour of the image is traced by scanning the edge image and re-sampling is done to avoid the discontinuities in the contour representation. The resulting image is swept line by line and the neighbour of every pixel is explored to detect the number of surrounding points and to derive the shape features [3]

Contour-based shape description technique is implemented in this work. The contour detection is one of the most used operators in image processing. In this method, lengths of the shape's radii from centroid to boundary are used to represent the shape. If a shape is used as feature, edge detection is the first step to extract that feature. Here, the canny edge detector is used to determine the edge of the object in the scene. After the edge has been detected the important step is tracing the contour of the object in the scene. For this the edge image is scanned from four directions (right to left, left to right, top to bottom, bottom to top) and the first layer of the edge occurred is detected as image contour. To avoid discontinuities in object boundary the contour image is then re-sampled. Next, the central point of the object is located. The centroid is computed using the given equations:

$$x_c = \sum_{i=1}^{n} \frac{x_i}{n} \ldots(9)$$

$$y_c = \sum_{i=1}^{n} \frac{y_i}{n} \ldots(10)$$

Where n is the number of points of an object. The contour is characterized using a sequence of contour points described in polar form [11]. Here the pole at the centroid (x_c, y_c) is taken then the contour graph is obtained using the equation d=f(θ) and each point (x,y) has the polar description where x, y, d, θ are related using the given equations:

$$d = \sqrt{(x - x_c)^2 + (y - y_c)^2} \ldots(11)$$

$$\theta = \tan^{-1}(\frac{y - y_c}{x - x_c}) \ldots(12)$$

RESULT

The approach described herein has proven to be simple and effective for shape based image Retrieval and it performs better than salient points method. Average retrieval efficiency is 71.17% (for top 20 images), 70.07% (for top 25 images) and 80.625% (for top 30 images).

4. Image Retrieval using Local Color Features [4]

In this technique, an image indexing and retrieval approach using local color features and a modified weighted color distortion measure is proposed. In it, each image is segmented into several regions by a watershed segmentation algorithm, and then the mutual relationships between connected color regions are extracted as local color features. That is, an image can be represented as set of connected (adjacent) color regions and the mutual relationships between connected color regions. In the image retrieval stage, the similarity between a query image and a target image will contain not only direct region correspondence but also the mutual relationships between connected color regions. Watershed algorithm is used to partition an image into segmented regions. Because the watershed algorithm is a gradient-based region-growing method [12], each input color image 1 is converted into a gray-scale image L. To match the characteristics of the human visual system, the conversion is a weighted average conversion [13]

RESULT

To evaluate the performance of the proposed approach, three existing approaches for comparison are implemented. They are (1) the conventional color histogram (CCH) approach, (2) the fuzzy color histogram (FCH) approach [3], and (3) the unified feature matching (UFM) approach.
The performance of the proposed approach is found to be better than those of the three existing approaches. The average precision value is found to be 0.281 whereas for CCH it is 0.242, for FCH it is 0.260, for UFM 0.202.

VI. CONCLUSION

Research in content-based image retrieval (CBIR) in the past has been focused on image processing, low-level feature extraction, etc. In this review paper, the main focus is on Low Level Feature extraction. Features like color, shape, texture has been extracted by applying suitable techniques. Use of Neural networks has also given good results. These techniques give good results when the image database is not very large.

But, in recent trends, image retrieval needs more accurate results when the database increases. Extensive experiments on CBIR systems demonstrate that low-level image features cannot always give proper results. Therefore, high-level semantic concepts can be introduced as well, so as to decrease the semantic gap between the low level and the high level features. It is believed that CBIR systems should provide maximum support in bridging the 'semantic gap' between low-level visual features and the richness of human semantics.
Therefore, for better retrieval results the high level features can also be introduced along with the low level features for better retrieval results.

REFERENCES

[1] SagarSoman, MitaliGhorpade, VrushaliSonone, SatishChavan, (2011), "Content Based Image Retrieval Using Advanced Colour and Texture Features", International Conference in Computational Intelligence (ICCIA)

[2] Masato Yonekawa and Hiroaki Kurokawa,School of Computer Science, Tokyo University of Technology,2012, "The Content-Based Image Retrieval using the Pulse Coupled Neural Network", WCCI 2012 IEEE World Congress on Computational Intelligence.

[3] S.Arivazhagan, L.Ganesan, S.Selvanidhyananthan, 2007,"Image Retrieval using Shape Feature", International Journal of Imaging Science and Engineering (IJISE).

[4] Tzu-Chien Wu, Jin-Jang Leou,@ Member, IEEE, and Li-Wei Kang, 2005,"Image Retrieval Using Local Color Features", 0-7803-8838-0/05/$20.00 02005 IEEE

[5] Dr.H.B.Kekre, Sudeep D. Thepade and AkshayMaloo, "Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform," International Journal of Engineering Science and Technology. Vol. 2, No. 4, pp. 362-371, 2010

[6] R. Echorn, H. J. Reitboeck, M. Arndt, P. Dicke, "Feature linking via synchronization among distributed assemblies: Simulations of resultsfrom cat visual cortex," Neural Computation, vol.2, pp.293–307, 1990

[7] A. K. Engel, A. K. Kreiter, P. K ¨ onig, and W. Singer, "Synchronizationof oscillatory neuronal responses between striate and extrastriate visual cortical areas

of cat," Proc. Natl. Acad. Sci. USA, vol. 88, pp.6048–6052,1991

[8] W. Kruse and R. Eckhorn, "Inhibition of sustained gamma oscillations (35-80 Hz) by fast transient responses in cat visual cortex," Proc. Natl. Acad. Sci. USA, vol. 93, pp.6112–6117,1996

[9] R. Echorn, "Neural Mechanisms of Scene Segmentation: Recording fromthe Visual Cortex Suggest Basic Circuits for Liking Field Model," IEEETrans. Neural Network, vol.10, no.3, pp.464–479, 1999

[10] J. L. Johnson, M. L. Padgett, "PCNN Models and Applications," IEEE Trans. Neural Network, vol.10, no.3, pp.480–498, 1999

[11] Hwei-Jen Lin, Yang-Ta Kao, Shwu-Huey Yen, and Chia-Jen Wang. "A Study of shape-Based Image Retrieval", Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04)0-7695-2087-1, 2004.

[12] Y. Chen and J. 2. Wang, "A region-based fimy feature matching approach to content-based image retrieval," IEEE Trans on Putrern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1252-1267, September 2002.

[13] D. A. Adjeroh and M. C. Lee, "On ratio-based color indexing," IEEE Trum. onlmage Processing, vol. 10, no. 1, pp. 3648, January 2001.

Ankita Sharma,received her B.tech Degree from MaharanaPratap College of Technology, Gwalior in the year 2010. Presently, she is pursuing her M.tech (ECE) from Lovely Professional University, Punjab (2013). Her area of interest is image processing and in future her focus is to pursue Ph.D in the same domain.

Gurpreet Singh, received his B.tech degree from Rayat Institute of Engineering and Technology in the year 2010. Presently, he is pursuing M.tech (ECE) from Lovely Professional University, Punjab (2013). His area of interest is Image Processing and Neural Networks. In future, his aim is to pursue Ph.D in the same area.

.