Enhanced CBIR using Color Moments, HSV Histogram, Color Auto Correlogram, and Gabor Texture

Shiv Raj Singh, Dr. Shruti Kohli

Department of Computer Science
BIT, Mesra, Ranchi
NOIDA Campus, INDIA

Abstract: *Content-based image retrieval (CBIR) is an emerging field of research due to the exponential rise of multimedia files. CBIR is an application of computer vision techniques. CBIR retrieves the similar digital images from a large collection of images. The term "content or features" in this context might refer to color, shape, texture, or any other information that can be derived from the image itself. On the basis of these features, feature vector is created for each and every image and stored as feature vector database. In this approach, Support Vector Machine (SVM) is used as an image classifier which gives the better result in retrieval of similar images. Similarity matching is performed using various distance measuring techniques like Euclidean distance and city block distance etc.*
.

Keywords: *computer Vision; contents or features; feature vector; query image; similarity matching; Euclidean distance; city block distance; support vector machine.*

I. Introduction

Content-based image retrieval (CBIR), a technique which uses visual contents of digital image for searching similar images from a huge collection of digital images, has been an active research area since 1990s. Although, a significant progress has been observed in this field however, there remain many challenging research problems that continue to attract researchers from multiple disciplines.

In earlier approach, the images were first annotated with text and then searched using text-based approach from traditional database management systems. This approach is known as Text-based Image Retrieval (TBIR). Due to an advancement in technologies, volume of digital images produced by various fields like scientific, military, educational, medical, industrial, and other applications increased exponentially. The difficulties faced by earlier approach (TBIR) became more and more severe. This need formed the driving force behind the emergence of CBIR.[1]

Researchers from various fields like computer vision, database management, human-computer interface, and information retrieval were attracted to this field.

CBIR, uses the basic visual features of an image such as color, shape, texture and spatial layout to represent and index the images. In typical content-based image retrieval systems, the features of every image are extracted and grouped together to form the feature vector. All the feature vectors are stored in the feature database. Similarity matching between the feature vector of the query image and feature vector database is carried out to retrieve the similar images. The CBIR, can successfully be used in the applications of scientific databases, art museum, medical science database, collection of photographs, crime prevention, military applications, historical research, fashion and graphics design, remote sensing images, education and training, architectural and engineering design, automatic inspection system, oil-spill detections and many more.

II. Related work

In Enhanced Multistage Content-based Image Retrieval [7] explained the Search criteria based on three-layer feed-forward architecture. First layer consists of the comparison of color feature, second layer consists of the comparison of texture feature and the last layer consists of the comparison of shape feature. The output of first stage is presented to input of second stage, and the output of second stage is presented to input of third stage.

In the Web Based Image Retrieval System Using Color, Texture and Shape Analysis Comparative Analysis [8] research work shows the client-server functionality, searching images without any compatibility issue. For the retrieval of images, primitive features of images have been used. The working principle of various applications like QBIC, Netra, Image Management System, and Kingfisher has been explained.

In the Sketch4Match - Content-based Image Retrieval System using sketches,[9] explained that the human being can recall the visual information more quickly than the textual information since the text is the human abstraction of image. In this paper, a user interface is proposed with some drawing area in which the user draws a color sketch or a blob. Features of the color sketch are extracted and form a feature vector. The feature vector of sketch or blob, matches with the stored feature vector database and the similar images are retrieved.

In Content-Based Image Retrieval System in Medical Applications [10] shows, how CBIR can be helpful in the Medical Application due to ever increasing image database. Digital images are produced in a huge amount by the Radiology, Cardiology, and by the Pathology departments.

A need has been arisen to build a system which can handle such a huge number of digital images. According to this, CBIR can successfully be implemented in diagnosis of disease by comparing the stored images.

Sudhir Ramdas in his research work "A Efficient Content-based Image Retrieval System using GMM and Relevance Feedback" [11] explained that how a CBIR can be efficient and accurate. According to him, there is always a gap between low-level and high-level features of an image. To bridge the gap, Relevance Feedback were introduced to get more accurate result

A multi-feature model for the Content Based Image Retrieval System [12] is proposed by combining the color histogram, color moment, texture, and edge histogram descriptor features. In this approach, users can select desired feature extraction method for retrieving the images. They found good results for most of the query images. The result can be improved further adding relevance feedback.

In "Content-Based Image Rretrieval – Extraction by objects of user interest" [13] explained that the query Image may have many objects. User can select a particular object from the query image according to their interest. Features like color moments, wavelet texture, and region-based shape feature have been used. The search is performed in two stages. The first stage integrates global color and texture feature to narrow down the search space. The second stage combines the color, texture, and shape feature of the object of user's interest for image retrieval. The researchers found that this method retrieves images accurately specifically when many objects are present in the query image.

III. METHODOLOGY

The technique explained in this document is based on color and texture features of digital images. The algorithm for proposed system is as follows:

Input: digital color images (png, jpg, bmp).

Output: Retrieved images from a collection of images.

Step 1: Select directory of images.

Step 2: All images are resized to 384 x 256.

Step 3: Convert image from RGB to HSV color space, quantize image to 8x2x2 and compute HSV histogram. Normalize histogram to unit sum.

Step 4: Compute Color Auto Correlogram.

Step 5: Compute Color Moments (up to 3rd order) for Red, Green, and Blue planes of the image.

Step 6: Compute mean amplitude, Mean squared energy from Gabor Wavelet (number of scales = 4 and no. of orientations = 6) and compute wavelet moments.

Step 7: Construct dataset of feature vectors using features extracted from HSV histogram, Color Auto Correlogram, Color Moments, Mean amplitude, Mean Squared Energy, and wavelet moments.

Step 8: Load dataset.

Step 9: Select Query image.

Step 10: Extract features of query image using steps from 2 to 6 and construct feature vector for query image.

Step 11: Match the feature vector of the query image with the stored dataset of feature vectors of all the images using Euclidean distance or City block distance and retrieve similar images.

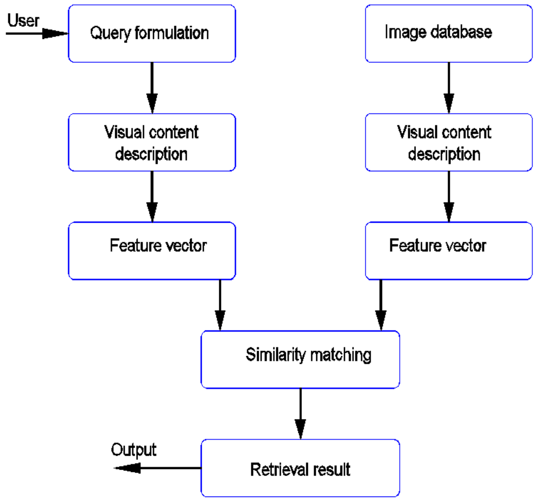


Fig-2 Proposed Block diagram of CBIR

A. *Color features*

Color is an important feature used in CBIR. Color is a property that depends on the reflection of light to the eye and

the processing of that information in the brain. Color features are relatively easy to compute and plays an important role for effective indexing and searching of color images in the image database.[3]

Before extracting the color feature, we have to determine the color space. A color space usually a multidimensional space in which each dimension represents its primary colors. RGB is a 3D color model which is used in CBIR. However, RGB color space is not perceptually uniform. The HSV (Hue, Saturation, Value) color space is approximately perceptually uniform and easily interchangeable to RGB color space. In this approach, both color spaces have been used for feature extraction.

The HSV values of a pixel can be transformed from its RGB representation according to the following formula:

$$H = \cos^{-1} \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}$$

$$S = 1 - \frac{3[\min(R,G,B)]}{R+G+B}$$

$$V = \left(\frac{R+G+B}{3}\right)$$

The *hue* (H) refers to an attribute associated with the dominant wavelength in a mixture of light waves. Alternatively, it represents the dominant color as perceived by an observer. When we call an object red, orange, or yellow, we are referring to its *hue*. *Saturation(S)* refers to the relative purity or the amount of white light added with a *hue*. The pure spectrum colors are fully saturated. The color red is pure saturated color and colors such as pint (red plus white) and lavender (violet plus white) are less saturated. The *value* (V) of a color, also known by its intensity level and describes how dark the color is. Intensity level 0 is black, and 1 is white for *value* (V).

B. Color Moments

Color moments are successfully used in many retrieval systems. Color distribution of an image can be interpreted as a probability distribution of colors and can be characterized by its moments. Color moments are scaling and rotation invariant. Most of the color distribution information's are found in the low-order moments. So, in this approach, only the first three color moments have been used as features extraction. The first order moment gives the mean value of color distribution, the second order gives the standard deviation, and the third order gives the skewness color moment.

i. Mean

Mean shows the average color value in the image and it can be calculated by using the following formula:

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij}$$

Where N represents the number of pixels in the image and p_{ij} is the value of the j^{th} pixel of the image at the i^{th} color component.

ii Standard Deviation

The second color moment is the standard deviation, which is obtained by taking the square root of the variance of the color distribution.

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^2\right)}$$

Where E_i is the mean value, or first color moment, for the i^{th} color component of the image.

iii Skewness

The third color moment represents the skewness. Skewness measures the asymmetric distribution of colors in an image and can be computed as:

$$s_i = \sqrt[3]{\left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^3\right)}$$

Output feature vector is computed using nine components i.e., Mean, standard deviation, and skewness for each of the color component (Red, Green, and Blue).

C. HSV Histogram

The color histogram is widely known feature to be used in CBIR. It characterizes the distribution of colors in the image. Moreover, it is robust to translation and rotation about the view axis. Any pixels in the color image is represented by three components i.e., red, green, and blue in RGB color space and hue, saturation, and value in HSV color space. A histogram, i.e., the distribution of the number of pixels for each quantized bin, can be defined for each component image. However, with increase in number of bins, discrimination power increases which may increase the computational cost.

In this approach, Histogram in HSV color space has been computed using quantization level 8 for H(hue), 2 for S(saturation), and 2 for V(Value) components. A feature vector of 1x32 is computed and normalized it to *unit sum.*

D. Color Autocorrelogram

A color correlogram expresses how the spatial correlation of pairs of color changes with distance. Any scheme that is based on only global properties may be sensitive to changes in appearance while the correlogram is more stable against these changes.[4]

Suppose *I* be an n x n image. Using the distance d ϵ {1,2,.....,n}, the correlogram of image I for i, j ϵ {1,2,.....,m} and k ϵ {1,2,.....d} is defined as

$$\gamma_{ci,cj}^{(k)}(I) = \Pr_{p_1 \in I_{ci}, p_2 \in I}\left[p_2 \in I_{cj} \mid \left[p_1 - p_2\right] = k\right]$$

The color correlogram is a tabled index by color pairs; Where the k^{th} entry for (i, j) determines the probability of finding a pixel with color j at a distance k from a pixel with color i in the image. This feature makes image retrieval systems robust against large changes in appearance of the same scene.[7]

Color Auto Correlogram is a subset of color correlogram. Color Auto Correlogram only captures the spatial correlation between identical colors. In comparison of color histogram and Color Coherence Vector (CCV), the auto-correlogram gives the better retrieval results. The Auto Correlation Coefficient is defined as:

$$\alpha_c^{(k)} = \gamma_{c,c}^{(k)}(I)$$

In this approach, Color Auto Correlogram is computed over a quantized color space (RGB) with four distances: 1, 3, 5 and 7. For each color-distance pair(c, k), the probability of finding the same color at exactly distance k away was computed.

E. Texture features

Image can be described by its texture, has a discrimination power to identify an image. Although there is not a universal agreed upon definition of texture, image processing techniques usually associate the notion of texture with image properties such as smoothness, coarseness, regularity, homogeneity, directionality. Low-level texture features play a vital role in CBIR. Similar images will be having similar texture patterns, so texture features are important for content-based image retrieval [5].

Gabor wavelet is widely used to extract texture of an image. Basically Gabor filters are a group of wavelets, each wavelet captures energy at a specific frequency and specific orientation. Frequency and orientation are similar to those of the human visual system. The scale and orientation are adjustable property of Gabor filter which is useful for texture analysis. The design of Gabor filter is as follows: [3].

For a given image *I*(x,y), Discrete Gabor wavelet transform is given by a convolution:

$$G_{mn}(x,y)=\sum_{s}\sum_{t}I(x-s,y-t)\Psi^{*}_{mn}(s,t)$$

Where, s and t represent the filter mask size variables, Ψ^*_{mn} is a complex conjugate of Ψ_{mn} which is a class of self similar functions generated from dilation and rotation of the following mother wavelet:

$$\Psi_{mn}(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi Wx)$$

Where σ_x and σ_x are the standard deviations of the Gaussian envelopes along the x and y direction, W is the modulation frequency.

In this approach, Output vector computed using Mean amplitude, Mean squared energy and wavelet moments.

F. Computing similarity

Two feature vectors (feature vector of the query image and feature vector database) in the feature space can be compared using the distance between them or, conversely, establishing their degree of similarity. For two feature vectors $a = (a_1, a_2, \ldots, a_n)^T$ and $b = (b_1, b_2, \ldots, b_n)^T$, the distance calculation is as follows:

- Euclidean distance:

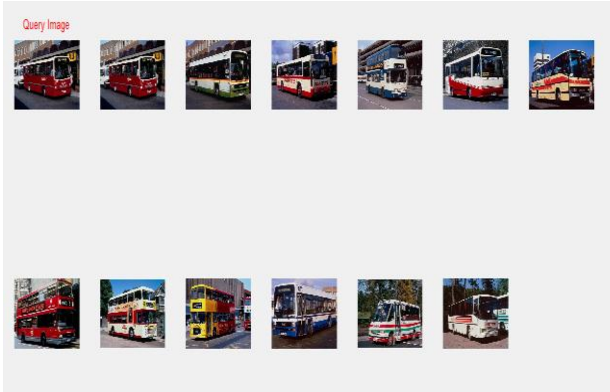
$$d_E=\sqrt{\sum_{i=1}^{n}(a_i-b_i)^2}$$

- City-block distance:

$$d_M=\sum_{i=1}^{n}|a_i-b_i|$$

Most similar images get retrieved using these distances using one to one matching of features of the query image and feature vectors in dataset.

The retrieved images are the result as shown in following figure. Images are retrieved using city-block distance. The performance is analyzed with the error matrix, generating using Support Vector Machine (SVM). SVM is a type of supervised learning technique which gives the better result for image classification. An error matrix called confusion matrix is generated using SVM which summarizes the relationship between the various classes.



G. Performance analysis:

In the proposed document, Support Vector Machine (SVM) has been used for analyzing the data. The result is analyzed with the help of the confusion matrix, also known as a contingency table or an error matrix. Confusion matrix is a specific table layout which shows the performance of an algorithm, specifically in case of supervised learning. Each column represents the instances in a predicted class while each row represents the instance in an actual class.

	W1	W2	W3	W4	W5
W1	90.00% (45)	4.00% (2)	6.00% (3)	0	0
W2	8.00% (4)	78.00% (39)	6.00% (3)	6.00% (3)	2.00% (1)
W3	10.00% (5)	8.00% (4)	76.00% (38)	6.00% (3)	0
W4	2.00% (1)	2.00% (1)	6.00% (3)	90.00% (45)	0
W5	0	0	0	0	100.00% (50)

The main diagonal of the matrix indicates the number of cases where the classifier was successful; a perfect classifier would show all off-diagonal elements equal to zero. In this case, class W5 were correctly classified all the time. Classification errors were highest (24%) for inputs labelled as class W3. The most common confusion incurred by the classifier was labelling an input of class W3 as W1 (10% of the time). Moreover, the classifier's performance for class W1 and W4 are worth commenting: although 10% of inputs labelled as class W1 were incorrectly classified (4% as class W2 and 6% as class W3). Similarly, 10% of input in class W4 are incorrectly classified (2% as W1, 2% as class W2, and 6% as class W3).

From the above matrix the error rate would be:

Misclassified data of class W1 = 10% (4%+6%)

Misclassified data of class W2 = 22% (8%+6%+6%+2%)

Misclassified data of class W3 = 24% (10%+8%+6%)

Misclassified data of class W4 = 10% (2%+2%+6%)

Misclassified data of class W5 = 0

Performance analysis is done based on error rate. The error rate measures the number (percentage) of error observed when testing of classifier against a test-set. From the above results the error rate would be:

(10+22+24+10+0)/(5 x 100) = 13.2%.

The result shows the performance of the system is quite good. Although, for analysis of performance other factors may also be important such as speed and computational complexity, but error rates are essential.

IV. CONCLUSION AND FUTURE WORK

Almost all CBIR systems use basic features like color, texture and shape, spatial information, etc. In this approach, most suitable features have been used which shows better result. HSV histogram is perceptually related with human vision system. The system suggested in this approach can also be extended for making applications for security check and crime prevention applications, medical diagnosis, oil-spill detection, automatic inspection system, etc. The researchers can use a different combination of features for further improvement in the system

REFERENCES

[1] Richa Jain1 Krishna2 “An Approach for Color Based Image Retrieval” International Journal of advanced electronics & communication systems Issue 2 Vol 2, Sept-Oct, 2012.

[2] Deepak S. Shete1, Dr. M.S. Chavan2 “Content Based Image Retrieval : Review” International Journal of Emerging Technology and Advanced Engineering” (ISSN 2250-2459, Volume 2, Issue 9, September 2012).

[3] S. Mangijao Singh, K. Hemachandran “Content-Based Image Retrieval using Color Moment and Gabor Texture Feature” IJCSI International Journal of Computer Science issue, Vol. 9, Issue 5, No 1, September 2012.

[4] Mohsen Fathian and Fardin Akhlaghian Tab “A Novel Content-Based Image Retrieval Approach Using Fuzzy Combination of Color and Texture” AICI'11 Proceedings of the Third international conference on Artificial intelligence and computational intelligence - Volume Part III Springer-Verlag Berlin, Heidelberg 2011.

[5] Ms. R. Janani1, Sebhakumar2 “An improved CBIR Method Using Color and Texture Properties with Relevance Feedback” International Journal of Innovative Research in Computer and Communication Engineering – Vol.2, Special Issue 1, March 2014.

[6] Ritendra Dutta, Jia Li, James Z. Wang “Content-Based Image Retrieval – Approaches and Trends of the New Age” ACM 2005.

[7] Nandish Chauhan1, Mahesh Goyani2 “Enhanced Multistage Content Based Image Retrieval” IJCS, Vol. 2, Issue. 5, May 2013,.

[8] Amol P Bhagat1, Mohammad Atique2 “Web Based Image Retrieval System using Color, Texture and Shape Analysis: Comparative Analysis” International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-3 Number-3 Issue-12 September-2013.

[9] Shadma Parveen1, Shweta Yadav2, Neelu Chauhan3 “Sketch4Match – Content-based Image Retrieval System Using Sketches” International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013.

[10] R.Senthil Kumar1, Dr.M.Senthilmurugan2 “Content-Based Image Retrieval System in Medical Applications” International Journal of Engineering Research & Technology (IJERT) ISSN:2278-0181 Vol-2 Issue 3, March-2013.

[11] Ramadass Sudhir, “A Efficient Content based Image Retrieval System using GMM and Relevance Feedback” International Journal of Computer Application (0975-8887) Volume72-No.22. June 2013.

[12] Rajshree S. Dubey, Rajnish Choubey, Joy Bhattacharjee, “Multi Feature Content Based Image Retrieval” (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 06, 2010, 2145-2149.

[13] D.Kishore Kumar1, E.Usha sree2, K.Suneera3, P.V.Chaitanya Kumar4 “Content Based Image Retrieval – Extraction by objects of user interest” International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No.3 Mar 2011.