International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 1, January 2014)

CBIR By Integration of Color and Texture Features

Vaishali Khandave[1], Nitin Mishra[2]
[1]M.Tech (I.T) Student, [2]Assistant Professor (I.T), NRI Institute of Information Science and Technology, Bhopal

***Abstract*— Content Based Image Retrieval (CBIR) is the retrieval of images based on visual features such as color, texture and shape. Image retrieval using single feature cannot provide good solution for accuracy and efficiency. High level feature describes the concept of human brain will reduce the query efficiency and low level features such as color, texture and shape will reduce query accuracy, so it is better way to use multi features for image retrieval. The most important visual features are Color and texture. In this paper technique used for retrieving the images based on their content namely color, texture and combination of both color and texture. The technique verifies the superiority of image retrieval using multifeature than the single feature.**

***Keywords*— Combined feature, Content Based Image retrieval (CBIR), Co-occurrence matrix, Euclidian distance**

I. INTRODUCTION

With the development of the Internet, and the availability of image capturing devices such as digital cameras, image scanners, the size of digital image collection is increasing rapidly. It is very important to efficiently store and retrieve images for different application such as fashion design, crime prevention, medicine, architecture, etc. For this purpose, many general purpose image retrieval systems have been developed. They are text-based and content-based [1].

The idea of text-based approach was originated at 1970s.In this images are stored in the database associating it with a keywords, number, texts. Then we search the images based on associated metadata such as keywords, texts and other elements that defined the image. This method is not only time consuming but it is dependent to the people who categorized it.

To describe an image such as in Figure 1, perception of each of the person who looks at the image will differs. One person will see the image as sunset scene and one person will interpret this image as river or lake.

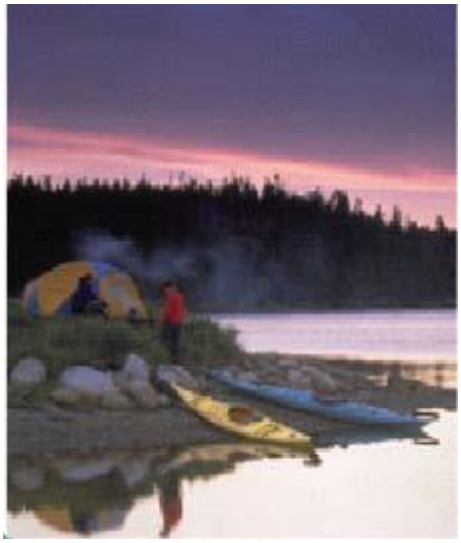

Figure 1: Example of scenery image

In this situation, "A picture is worth a thousand words" is being applied [2]. To overcome the above disadvantages in text-based retrieval system, content based image retrieval (CBIR) was introduced in the early 1980s. In CBIR, images are indexed by their visual content, such as color, texture, shapes. The CBIR mainly consists of two steps. One is the feature extraction and another one is the similarity matching. Means in CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image.

Here we are using the Color and texture visual features. First the work of color feature extraction is obtained by using HSV color space. Similarly, the work of texture feature is obtained by using gray-level co-occurrence matrix (GLCM) and color co-occurrence matrix (CCM). Then we combine color features and GLCM as well as color features and CCM separately [3]. System using combination of color and texture would provide accurate results than using the single feature.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 1, January 2014)

II. EXISTING METHOD

A. *Color Based Image Retrieval*

In exiting method they retrieved the images only based on Color. So even though if one image is same domain but in different color mean it will not retrieved the image [10].

B. *Color Histogram*

Color is the most widely used "feature" owing to its intuitiveness compared with other features and most importantly, it is easy to extract from the image. The color histogram depicts color distribution using a set of bins. However, a CBIR system based on color features is often found to yield distorted results, because it uses global color feature which cannot capture color distributions or textures within the image in some cases. To improve the preferment of the color extraction we divide the color histogram feature into global and local color extraction.

Using Global Color Histogram (GCH), 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.

Local color histogram (LCH) can give some sort of spatial information, however the con associated with it is that it uses very large feature vectors. LCH includes information concerning the color distribution of regions. The first step is to segment the image into blocks and then to obtain a color histogram for each block. An image will then be represented by these histograms. When comparing two images, we calculate the distance, using their histograms, between a region in one image and a region in same location in the other image. The distance between the two images will be determined by the sum of all these distances.

However, it does not include information concerning the color distribution of the regions, so the distance between images sometimes cannot show the real difference between images. Moreover, in the case of a GCH, it is possible for two different images to have a very short distance between their color histograms. And the accuracy of this system is less then (HSV model based feature extraction), this is the main disadvantage [9].

III. COLOR FEATURE EXTRACTIION

In image retrieval color feature is one of the most widely used featu re. Colors are defined on a selected color space. There are many color spaces like RGB, LAB, LUV, HSV (HSL), YCrCb and the huemin-max-difference (HMMD).

Color-covariance matrix, color histogram, color moments and color coherence vector. storing, filtering and retrieving audiovisual data are common color features or descriptors in CBIR systems. This thesis includes the HSV color space to extract the color feature.

A. *Color Feature Extraction Using HSV Color Space*

HSV stands for Hue, Saturation and Value. Here hue is used to distinguish colors, saturation is the percentage of white light added to a pure color and value refers to the perceived light intensity. Hue ranges from 0 to 360^0 .Saturation ranges from 0 to 100% and Value ranges from 0 to 100%. Because of a large range of each component, if directly calculate the characteristics for retrieval, then computation will be very difficult to ensure rapid retrieval. It is essential to quantify HSV space component to reduce computation and improve efficiency. At the same time, because the human eye to distinguish colors is limited, do not need to calculate all segments. Unequal interval quantization according the human color perception has been applied on H, S, and V components. Based on the color model of substantial analysis, we divide color into eight parts. Saturation and intensity is divided into three parts separately in accordance with the human eyes to distinguish [3].

B. *Algorithm for Color feature extraction using HSV color space*

Step 1: Convert the image's to HSV (hue (H), saturation(S) and value (V)) format.

*Step 2***:** Multiply 360 in H panel of the image.

Step 3: Apply this condition for extracting the color feature

$$H=\begin{cases}0 \text{ if } h\in[316,20]\\1 \text{ if } h\in[21,40]\\2 \text{ if } h\in[41,75]\\3 \text{ if } h\in[76,155]\\4 \text{ if } h\in[156,190]\\5 \text{ if } h\in[191,270]\\6 \text{ if } h\in[271,295]\\7 \text{ if } h\in[296,315]\end{cases}$$

$$S=\begin{cases}0 \text{ if } s\in[0,0.2]\\1 \text{ if } s\in[0.2,0.7]\\2 \text{ if } s\in[0.7,1]\end{cases}$$

$$V=\begin{cases}0 \text{ if } v\in[0,0.2]\\1 \text{ if } v\in[0.2,0.7]\\2 \text{ if } v\in[0.7,1]\end{cases}$$

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 1, January 2014)

For e.g. in H panel find if the pixel value is greater than 316 means convert that pixel as '0', also find if the pixel value is less than 20 means convert that pixel as '0'. Same as we need to implement for H, S and V.

Step 4: After extracting the feature from individual panel apply this eqn.

G = 9*H + 3*S + V

Step 5: Then convert 2D to 1D then find the sum, (this is called color feature).

Step 6: Apply Euclidean Distance for finding Similarity.

IV. Texture Feature Extractiion

The texture feature is another type of important and useful visual information for image retrieval. There exist different approaches to extract and represent textures. They can be classified into space-based, frequency-based models, and texture signatures. Some popular techniques i.e. wavelet transform, co-occurrence matrix, and Gabor filters are applied to express texture features for image. Here we used the Gray Level Co-occurrence matrix (GLCM) and Color Co-occurrence matrix (CCM) to extract the texture feature.

A. Algorithm for Texture feature extraction using GLCM

Step 1: Convert the image's to Gray format.

Step 2: Find the Energy of that gray image, for calculating the Energy follow this formula.

$$\text{Energy } E=\sum\sum p(i, j)^2$$

Step 3: Find the Contrast of that gray image, for calculating the Contrast follow this formula.

$$\text{Contrast } I=\sum\sum (i-j)^2\, p(i, j)$$

Step 4: Find the Entropy of that gray image, for calculating the Entropy follow this formula.

$$\text{Entropy } S=-\sum\sum p(i, j) \log p(i, j)$$

Step 5: Find the Inverse Difference of that gray image, for calculating the Inverse Difference follow this formula.

$$\text{Inverse difference } H=\sum\sum [\ 1/\ 1+(i-j)^2\ ] p(i, j)$$

Step 6: Apply Euclidean Distance for finding Similarity.

B. Algorithm for Texture feature extraction using CCM

Step 1: Convert the image to Gray format as well as HSV format.

Step 2: In CCM we are going to extract the feature from R panel, G panel, H panel and V panel.

Step 3: Find the Energy of the R panel, G panel, H panel and V panel, for calculating the Energy follow this formula.

$$\text{Energy } E=\sum\sum p(i, j)^2$$

Step 4: Find the Contrast of the R panel, G panel, H panel and V panel. for calculating the Contrast follow this formula.

$$\text{Contrast } I=\sum\sum (i-j)^2\, p(i, j)$$

Step 5: Find the Entropy of the R panel, G panel, H panel and V panel, for calculating the Entropy follow this formula.

$$\text{Entropy } S=-\sum\sum p(i, j) \log p(i, j)$$

Step 6: Find the Inverse Difference of the R panel, G panel, H panel and V panel, for calculating the Inverse Difference follow this formula.

$$\text{Inverse difference } H=\sum\sum [\ 1/\ 1+(i-j)^2\ ] p(i, j)$$

Step 7: Find the sum of all feature's in individual panel.

Step 8: Apply Euclidean Distance for finding Similarity.

V. Feature Extractiion By Combination Of Color And Texture

A. Algorithm for Combined feature extraction (HSV color space + GLCM)

Step 1: Get the Color features

Step 2: Get the GLCM features

Step 3: Apply the formula for combining those two features

$$D(A, B) = w1\frac{\sqrt{2}-D(F_{CA}-F_{CB})}{\sqrt{2}} + w2\frac{\sqrt{2}-D(F_{TA}-F_{TB})}{\sqrt{2}}$$

Formula in derivation form:

F(Color+Glcm) = (0.5.*((sqrt(2) - FColor)./sqrt(2)))+ (0.5.*((sqrt(2) - FGlcm)./sqrt(2)))

B. Algorithm for Combined feature extraction (HSV color space +CCM)

Step 1: Get the Color features

Step 2: Get the CCM features

Step 3: Apply the formula for combining those two features

$$D(A, B) = w1\frac{\sqrt{2}-D(F_{CA}-F_{CB})}{\sqrt{2}} + w2\frac{\sqrt{2}-D(F_{TA}-F_{TB})}{\sqrt{2}}$$

Formula in derivation form:

F(Color+Ccm) = (0.5.*((sqrt(2) - FColor)./sqrt(2)))+ (0.5.*((sqrt(2) - FCcm)./sqrt(2)))

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 1, January 2014)

VI. Relevance Feedback

Relevance Feedback was introduced in content based image retrieval to improve the performance by human intervention. To retrieve the image from the database, we first extract feature vectors from images (the features can be shape, color, texture etc), then store feature vectors into another database for future use. When given query image, we similarly extract its feature vectors, and match those features with database image features. If the distance between two images feature vectors is small enough; we consider the corresponding image in the database similar to the query. The search is usually based on similarity rather than on exact match, and the retrieval results are given to the user. Then user gives the feedback in the form of 'relevance judgments' expressed over the retrieval results. The relevance judgments evaluate the results based on a three value assessment. These three values are relevant, non-relevant and don't care. Relevant means the image relevant to the user, non- relevant means the image is definitely not relevant, and don't cares mean the user does not say anything about the image. If the user feedback is relevant, then feedback loop stops otherwise it continues until user get satisfied with results [4].

VII. Experimental Result And Evaluation

An experiment is conducted to explore the performance of system on image Set downloaded from http://wang.ist.psu.edu/docs/related/. Image Set consists of 1000 images. These images are grouped into 10 clusters with each containing 100 images. The images in the same cluster are considered as similar images.10 clusters are shown in Table I.

The precision and recall measurements are often used to describe the performance of an image retrieval system. The precision (P) and recall (R) are defined as follows:

P=n/L and R= n/N. Where L is the number of retrieved images; n is the number of relevant images in the retrieved images and N is the number of all relevant images in the database [8].

Table I
Ten Classes Of Image Set [8]

Classes	Semantic Name
1	African people and village
2	Beach
3	Building
4	Buses
5	Dinosaurs
6	Elephants
7	Flowers
8	Horses
9	Mountains and glaciers
10	Food

This experiment used each image in each class as a query image. This experiment was carried out with number L of retrieved images to compute the precision P of each query image and finally obtain the average precision P/100 (100 images of a class). The experimental results from proposed method and the other methods are shown in Table II. It is obvious that this proposed method has achieved a better average precision of various images than the other two methods [8].

Table II
The average precision of these methods on image

Semantic name	Proposed Method (Color+ Ccm)	Proposed Mehtod (Color+ GLCM)	CTDC IRS[8]	Jhanwar et.al.[09]	Hung and Dai's [10]
African people and village	0.61	0.60	0.562	0.4525	0.424
Beach	0.58	0.57	0.536	0.3975	0.4455
Building	0.70	0.69	0.61	0.3735	0.4105
Buses	0.932	0.92	0.893	0.741	0.8515
Dinosaurs	0.99	0.98	0.984	0.9145	0.5865
Elephants	0.612	0.60	0.578	0.304	0.4255
Flowers	0.91	0.90	0.899	0.8515	0.8975
Horses	0.82	0.81	0.78	0.568	0.589
Mountains and glaciers	0.578	0.562	0.512	0.2925	0.268
Food	0.732	0.721	0.694	0.3695	0.4265
Average	0.7464	0.7353	0.7048	0.52645	0.53245

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 1, January 2014)

The image retrieval results (dinosaurs) using different techniques is shown in following figures.

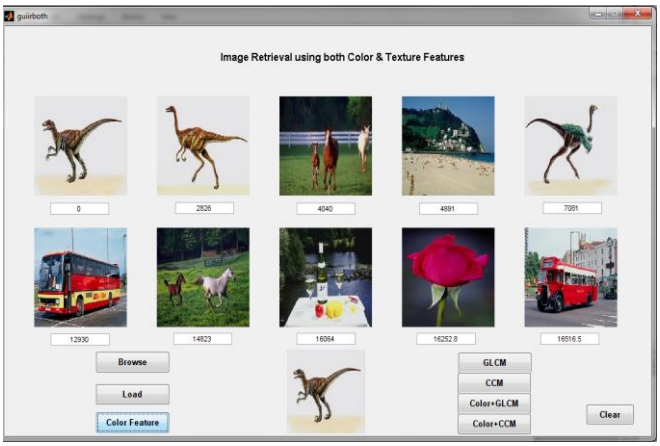

Figure 2: Retrieval Result Based on HSV color space

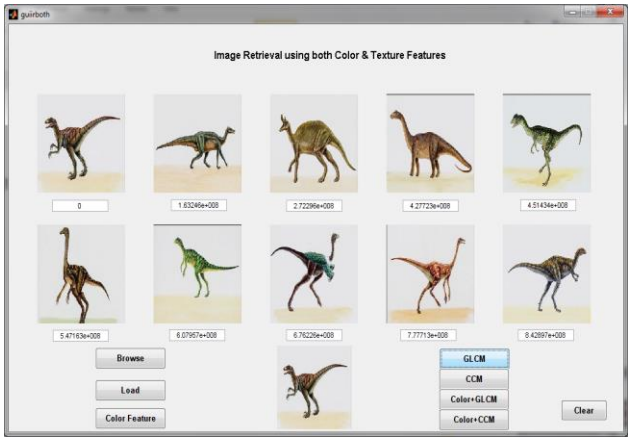

Figure 3: Retrieval Result Based on GLCM

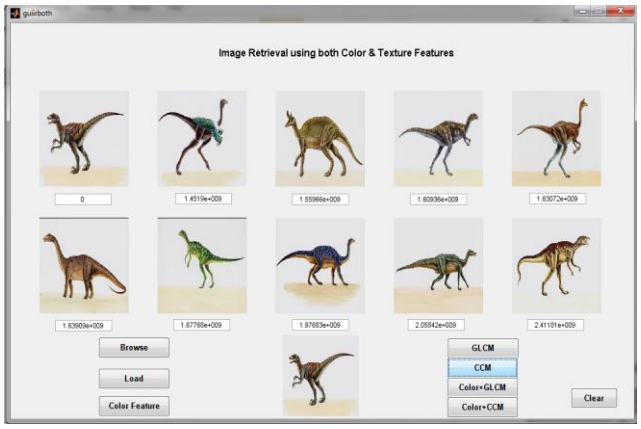

Figure.4: Retrieval Result Based on CCM

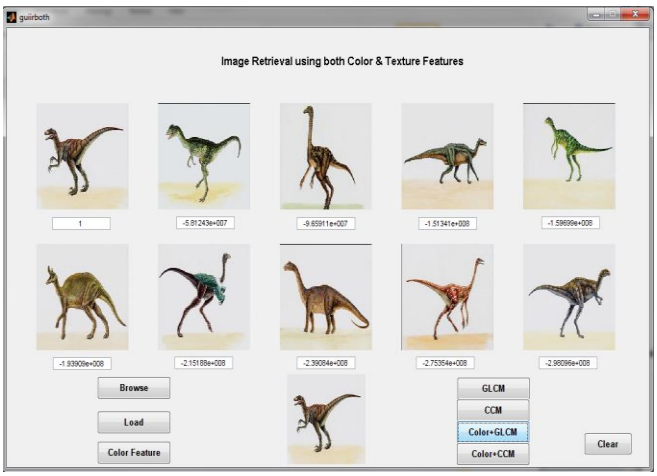

Figure 5: Retrieval Result Based on Color+GLCM

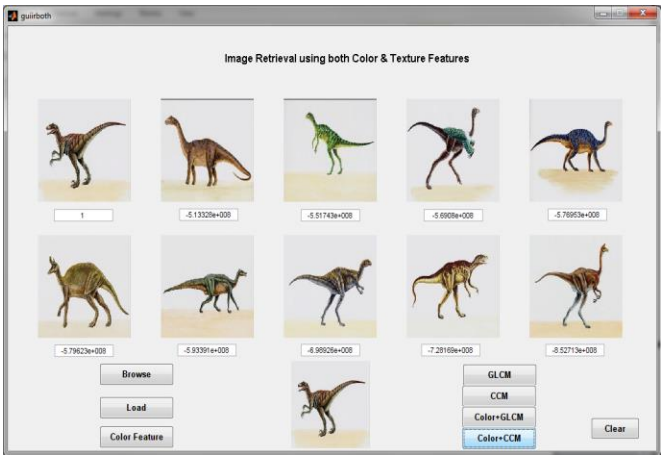

Figure 6: Retrieval Result Based on Color+CCM

VIII. CONCLUSION

For color feature extraction, HSV color space method is used. For texture, features such as energy, contrast, entropy, inverse difference etc. are used. These features together increase the retrieval efficiency of the texture feature. Both gray-level co-occurrence matrix (GLCM) and color co-occurrence matrix (CCM) algorithms are used for texture feature extraction. For combined feature extraction, we extract the color and texture features altogether i.e. we combine color features and GLCM as well as CCM separately. From the results it shows that image retrieval using integrated features gives more better and accurate results than image retrieval using single feature. The result of retrieval varies upon the content type of image.

REFERENCES

[1] Swapnalini Pattanaik, Prof.D.G.Bhalke ,"Beginners to Content Based Image Retrieval"International Journal of Scientific Research Engineering &Technology (IJSRET) Volume 1 Issue2 pp 040-044 May 2012.

[2] Ruziana Mohamad Rasli, T Zalizam T Muda, Yuhanis Yusof, Juhaida Abu Bakar, "Comparative Analysis of Content Based Image Retrieval Technique using ColorHistogram. A Case Study of GLCM and K-Means Clustering",Third International Conference on Intelligent Systems Modelling and Simulation,2012

[3] Fan-Hui Kong, "Image Retrieval Using Both Color And Texture Features", Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009.

[4] Pushpa B. Patil, Manesh B. Kokare," Relevance Feedback in Content Based Image Retrieval: A Review", Journal of Applied Computer Science & Mathematics, no. 10 (5) /2011, Suceava

[5] Rafael C.Gonzalez, Richard E.Woods "Digital Image Processing",Third Edition

[6] P.W. Huang, S.K. Dai, "Image retrieval by texture similarity", Pattern Recognition 36 (3) (2003) 665–679.

[7] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique," Content based image retrieval using motif co-occurrence matrix", Image and Vision Computing 22 (2004) 1211–1220

[8] M.Babu Rao, Dr. B.Prabhakara Rao, Dr. A.Govardhan, "CTDCIRS: Content based Image Retrieval System based on Dominant Color and Texture Features", International Journal of Computer Applications (0975 – 8887) Volume 18– No.6, March 2011

[9] Aman Chadha, "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval",Department of Electrical and Computer Engineering, University of Wisconsin-Madison, International Journal of Computer Applications (0975 – 8887) Volume 52– No.20, August 2012

[10] Rishav Chakravarti, Xiannong Meng, "A Study of Color Histogram Based Image Retrieval",Sixth International Conference on Information Technology,2009