Color Sketch Based Image Retrieval

N Raghu Ram Reddy[1], Gundreddy Suresh Reddy[2], Dr.M.Narayana[3]

B.Tech, Department of ECE, JPNCE, Andhra Pradesh, India [1]

B.Tech, Department of ECE, JPNCE, Andhra Pradesh, India [2]

H.O.D, Department of ECE, JPNCE, Andhra Pradesh, India[3]

ABSTRACT: This paper proposes a method of Content based image retrieval (CBIR) using color sketches, which is one of the most popular, rising research areas of the Digital image processing. The goal of CBIR is to extract visual content of an image automatically, like color, texture, or shape. The proposed method is based on free hand color sketch of image, which uses color and texture features. In this color features are extracted using HSV (Hue, Saturation and Value) color space. The HSV color space is quantified in non-equal intervals, a one dimensional feature vector is constructed and it is represented by cumulative histogram. Texture feature extraction is obtained by using gray-level co-occurrence matrix (GLCM). The combination of the color and texture features of an image provides a robust feature set for image retrieval. Euclidean distance of color and texture is used in retrieving the similar images. The image retrieval experiment indicates that the use of color and texture features in image retrieval using color sketches has obvious advantages. This paper will be very helpful in crime prevention. The proposed method has more retrieval rate than image retrieval using sketches.

KEYWORDS: CBIR, GLCM, HSV, texture shape, Euclidean distance.

I.INTRODUCTION

The rapid growth of multimedia applications followed by the emergence of large-scale image collections has brought about the need for efficient methods for storage, browsing, indexing and retrieval of images. In the early 1990s, the Content Based Image Retrieval (CBIR) was proposed to retrieve images based on visual features (contents) of images in substitute for the traditional annotation based retrieval, since the linguistic description of image by small number of key words is very limited in comparison to the richness of image contents. The visual contents called as features, are numeric descriptors capturing specific visual characteristics. Most CBIR systems have two-step approach to retrieve images from the databases. First step is indexing: for each image in a database, feature vector is computed and stored in feature space. Second step is searching: given a query image by a user, its feature vector is computed and the system retrieves images having feature vectors that best match the query feature vector. In CBIR, image feature selection is most critical problem because it largely affects the remaining aspects of the system design and eventual capabilities of the system.Our purpose is to develop a content based image retrieval system, which can retrieve using sketches infrequently used databases.

Using a sketch based system can be very Important and efficient in many areas of the life. In some cases we can recall our minds with the help of figures or drawing [7]. The SBIR systems [11] have a big significance in the criminal investigation. The identification of unsubstantial images, tattoos and graffities can be supported by these systems. Another possible application area of sketch based information retrieval is the searching of analog circuit graphs from a big database. The user has to make a sketch of the analog circuit, and the system can provide many similar circuits from the database.

In this paper color sketch based image retrieval system was developed by using color features and Gray-Level Co-occurrence Matrix (GLCM). Here we are using the color and Texture Feature for retrieving of images [8]. The proposed architecture is better than SBIR [10]. In SBIR color features are not used.

II. FEATURE EXTRACTION OF HSV COLOR

HSV defines a type of color space. The HSV color space has three components: hue, saturation and value. It is used to select various different colors needed for a particular picture. It gives the color according to human perception. In HSV, hue represents color. In this model, hue is an angle from 0 degrees to 360 degrees.

Table-1: Hue color ranges

Angle	Color
0-60	Red
60-120	Yellow
120-180	Green
180-240	Cyan
240-300	Blue
300-360	Magenta

Saturation indicates the range of grey in the color space. Its value is calculated from 0 to 1. When the value is '0,' the color is grey and when the value is '1,' the color is a primary color. A faded color is due to a lower saturation level, which means the color contains more grey. Value is the brightness of the color and varies with color saturation. Its value is calculated from 0 to 1. When the value is '0′ the color space will be totally black. With the increase in the value, the color space brightness up and shows various colors.

A. Non-interval quantization

Because of a large range of each component, if characteristics are directly calculated for retrieval, then computation will be very difficult to ensure rapid retrieval. It is essential to quantify HSV space component to reduce computations and improve its efficiency. At the same time, because the human eye can distinguish only limited colors, there is no need to calculate all segments. H,S,V components are quantized with unequal intervals according to the human color perception.

Based on the color model of substantial analysis, we divide color into eight parts. Saturation and intensity are divided into three parts separately in accordance with the human eyes to distinguish. In accordance with the different colors and subjective color perception quantification, quantified hue (H), saturation(S) and value (V) are showed as figure (1).In accordance with the quantization level above, the H,S,V three-dimensional feature vector for different values of with different weight to form one-dimensional feature vector named G[7].

$$G = Q_S Q_V\, H + Q_V\, S + V \tag{1}$$

Where QS is quantified series of S, QV is quantified series of V.
Here we set $Q_S = Q_V = 3$, then

$$G = 9H + 3S + V \tag{2}$$

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

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

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

In this way, three-component vector of HSV form one-dimensional vector, which quantize the whole color space for the 72 kinds of main colors. So we can handle 72 bins of one-dimensional histogram. This quantification can be effective in reducing the images by the effects of light intensity, but also reducing the computational time and complexity.

B. Characteristics of color cumulative histogram
Color histogram is derived by first quantize colors in the image into a number of bins in a specific color space, and counting the number of image pixels in each bin. One of the weaknesses of color histogram is that when the characteristics of images should not take over all the values, the statistical histogram will appear in a number of zero values. The emergence of these zero values would make similarity measure does not accurately reflect the color difference between images and statistical histogram method to quantify more sensitive parameters. Therefore, this paper represents the one-dimensional vector G by constructing a cumulative histogram of the color characteristics of image after using non-interval HSV quantization for G.

III. TEXTURE FEATURE EXTRACTION BASED ON GLCM

GLCM creates a matrix with the directions and distances between pixels and then extracts meaningful statistics from the matrix as texture features [12] . The main advantage of GLCM is its Rotation invariance characteristic [10]. It means the glcm property values of rotation form of an image are equal to its original image.GLCM texture features commonly used are shown in the following:

GLCM is composed of the probability value, it is defined by P (i, $j/d, \theta$) which expresses the probability of the couple pixels at θ direction and d interval. When θ and d is determined, P ($i, j/ d, \theta$) is showed by Pi, j . Distinctly GLCM is a symmetry matrix; its level is determined by the image gray-level. Elements in the matrix are computed by the equation showed as follow:

$$P (i, j/ d, \theta) = \frac{P (i,j| d,\theta)}{\Sigma\Sigma\, P (i,j| d,\theta)} \qquad (3)$$

GLCM expresses the texture feature according the correlation of the couple pixels gray-level at different positions [14]. It quantificationally describes the texture feature. In this paper, texture feature include four properties energy, contrast, entropy, homogeneity.

$$\text{Energy} \qquad E= \Sigma\, \Sigma\, p(x,y)^2 \qquad (4)$$

It is a gray-scale image texture measure of homogeneity changing, reflecting the distribution of image grayscale uniformity of weight and texture.

$$\text{Contrast I}= \Sigma\, \Sigma\, (x-y)^2 p(x,y) \qquad (5)$$

Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth.Contrast is large means texture is deeper.

$$\text{Entropy} \qquad S= - \Sigma\, \Sigma\, p(x,y) \log p(x,y) \qquad (6)$$

Entropy measures image texture randomness, when the space co-occurrence matrix for all values is equal, it achieved the minimum value; on the other hand, if the value of co-occurrence matrix is very uneven, its value is greater. Therefore, the maximum entropy implied by the image gray distribution is random.

$$\text{Homogeneity} \qquad H= \frac{\Sigma\Sigma\, 1\, p(x,y)}{1+(x-y)^2} \qquad (7)$$

It measures local changes in image texture number. Its value in large is illustrated that image texture between the different regions of the lack of change and partial very evenly. Here $p(x, y)$ is the gray-level value at the coordinate (x, y).

IV. EXPERIMENT AND ANALYSIS

In this paper, experimental data set contains 1000 Images from Wang database of images, divided into 10 Categories, each category has 100 images. Experimental images covers a wealthy of content, including Africa, beaches, moments, buses, food, dinosaurs, elephants, flowers, horses and mountains. We areusing rotated images of original images for the test purpose since GLCM is rotation invariant method which wasused for feature extraction. These images arestored in JPEG format.
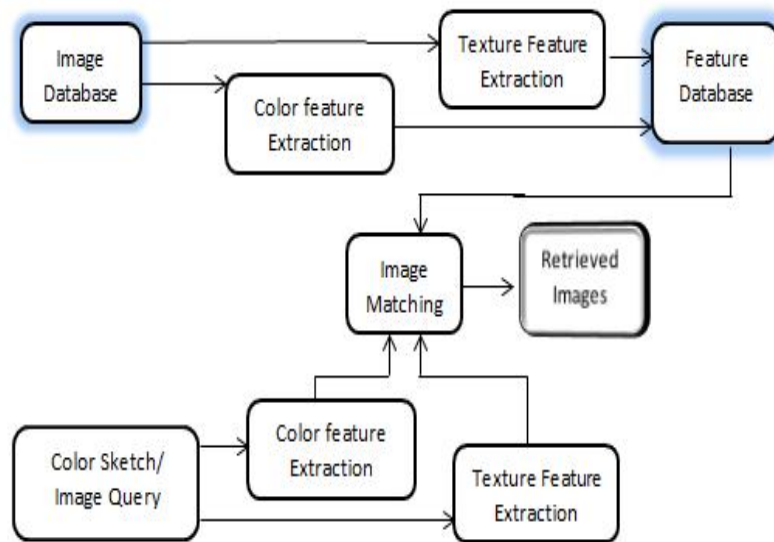

Fig 1: Proposed Block diagram of CSBIR

Here is the main objective of our program. As it is the color sketch based image retrieval system(CSBIR), we are using color sketch image as query. In this paper the query sketch was taken by converting original image into color sketch image. All the necessary sketch images were taken by converting their original images into color sketch form. For conversion photo to sketch converter software was used. These sketch images are given as query in the program while execution process.
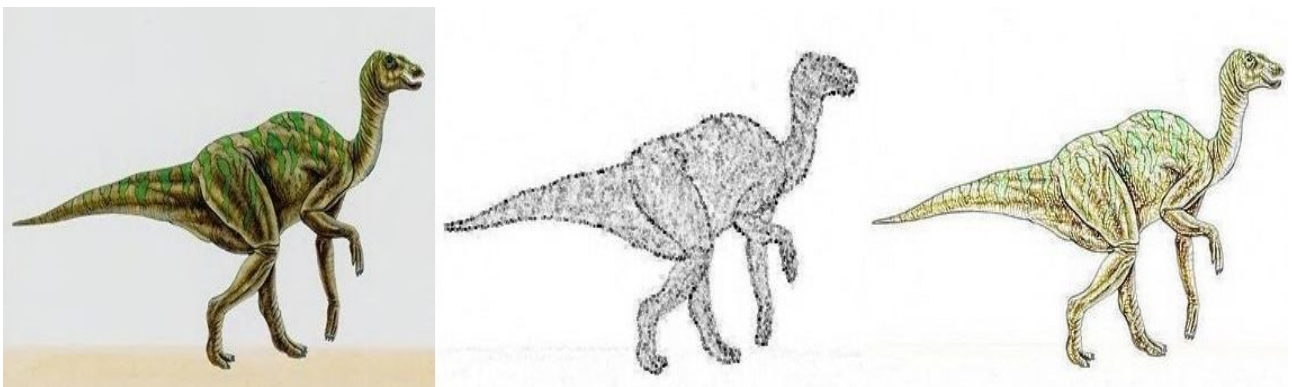

Fig:2 original image, sketch image and color sketch image

The texture feature and color features are extracted for the sketch image. The color sketch image feature values are compared with the feature database. The similarity between color, texture values of the color sketch image anddatabase

can be done by using distance metric. This comparison will be done at matching algorithm. The matching algorithm provides distance metric. Here we are using the Euclidean Distance metric for similarity measurement [4]. It is given by the following equation:

F. Computing the Similarity
Find the distances between feature vector of query image and the feature vectors of target images using normalized Euclidean distance. The similarity between query and target image is measured from two types of characteristic features which includes color and texture features to formulate the graph. So during the Euclidean similarity measure, when necessary the appropriate weights to combine them are also considered. Therefore, in carrying out Euclidean similarity measure we should consider necessary appropriate weights to combine them. We construct the Euclidean calculation model as follows:

$$D (A, B) = \omega 1 D (FCA, FCB) + \omega 2 D (FTA, FTB) \tag{9}$$

Here $\omega 1$ is the weight of color features, $\omega 2$ is the weight of texture features, *FCA* and *FCB* represents the normalized 72- dimensional color features for image A and B. For a method based on GLCM, *FTA* and *FTB* on behalf of 4- dimensional normalized texture features correspond to image A and B.Here, we combine color features and texture features. The value of ω through experiments shows that at the time $\omega 1 = \omega 2 = 0.5$ has better retrieval performance.

V. RESULTS AND DISCUSSIONS

Fig 3: Shows Top 20 retrieved images of SBIR and CSBIR based on rose image as query image

Fig 4: Shows Top 20 retrieved images of SBIR and CSBIR based on bus image as query image

Fig 5: Shows Top 20 retrieved images of SBIR and CSBIR based on building image as query image

The first test was conducted by using sketch image as query from database, the system was tested for top 20 retrieved images using Sketch Based Image Retrieval.In SBIR only GLCM features are used. The results of SBIR are shown in the figure 3, 5, 7. The retrieval rate is low for SBIR systems. The second test was conducted by using color sketch image as query from the database are shown in figure 4, 6, 8. In the second test both the color features and texture features are combined. Table 1 shows the image retrieval result. In this table, recall and precision are computed. The proposed architecture gives better performance than SBIR.

Table.2: Average Recall and precision rate

Retrieval mode	Recall (%)	Precision (%)
SBIR(sketch Based Image Retrieval)	24.1	37.2
CSBIR(Color Sketch Based Image Retrieval)	29.5	44.5

VI.CONCLUSION

In this paper, we have proposed system architecture for CSBIR based on HSV color space and texture characteristics of the image retrieval. In spite of the traditional text based image retrieval, color sketch based image retrieval is developed which gives best results. Through the quantification of HSV color space, we combine color features and Texture features using Euclidean distance classifier. Image retrievalexperiment, indicates that the use of color features and texture characteristics of the image retrieval method issuperior to sketch based image retrieval. In this work, it is observed that the proposed method achieved high retrieval rate using the color sketches of the images. This paper will be very useful in crime prevention. The proposed CSBIR has more retrieval rate than SBIR.

The future work will focus on improved retrieval performance of color and shape dominant images by exploring additional image features. Further, a research is in progress to improve the method aiming to increase the retrieval rate.

REFERENCES

1. Rui,Y.,Huang,T.S.,Mehrotra,S.[Sharad],"Retrieval with relevance feedback in MARS", In Proc of the IEEE Int'l Conf. on Image Processing, New York, pp.815-818, 1997.
2. H. T. Shen, B. C. Ooi, K. L. Tan, "Giving meanings to www images" Proceedings of ACM Multimedia, pp. 39-48, 2000.
3. B S Manjunath, W Y Ma, "Texture feature for browsing and retrieval of image data",IEEE Transaction on PAMI, Vol 18, No. 8, pp.837-842,1996.
4. Y. Rui, C. Alfred, T. S. Huang, "Modified descriptor for shape representation, a practical approach",In: Proc of First Int's workshop on Image Database and Multimedia Search, 1996.
5. Cao LiHua, Liu Wei, and Li GuoHui, "Research and Implementation of an Image Retrieval Algorithm Based on Multiple Dominant Colors", Journal of Computer Research & Development, Vol 36, No. 1,
6. pp.96-100,1999.
7. J. R. Smith, F. S. Chang, "Tools and Techniques for Color Image Retrieval", Symposium on Electronic Imaging: Science and Technology-Storage and Retrieval for Image and Video Database IV, pp.426-237, 1996.
8. Song Mailing, Li Huan, "An Image Retrieval Technology Based on HSV Color Space", Computer Knowledge and Technology, No. 3,pp.200-201, 2007.
9. Shang Lin, Yang YuBin, Wang Liang, Chen ZhaoQian, "An Image Texture Retrieval Algorithm Based on Color Co-occurrence Matrix (MCM)", Journal of NanJing University (Natrual Science). Vol 40, No. 5, pp. 540-547, Sept.2004.
10. YANG Yubin, Chen Shifu, Lin Hui, "A Novel Image Retrieval Method Using Texture Features Based on Color-Connected Regions", ACTA ELECTRONICA SINICA, Vol 33, No. 1, pp. 57-62,Jan. 2005.
11. M.Narayana, Subhash Kulkarni : Content Based Image Retrieval using Sketches. In: Springler India Proceedings of ICAdC, AISC 174, pp. 1117–1123 (2013)
12. Eitz, M., Hildebrand, K., Boubekeur, T., Alexa, M.: An evaluation of descriptors for large-scale image retrieval from sketched feature lines. Computers and Graphics 34, 482– 498 (2010)
13. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Hiang, Q., Dom, B., Gorkani, M.,Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by image and video content: the QBIC system. IEEE Computer 28, 23–32 (2002)
14. Györök, G.: Embedded hybrid controller with programmable analog circuit. In: IEEE 14th International Conference on Intelligent Systems, pp. 59.1–59.4 (May 2010)
15. Hu, R., Barnard, M., Collomosse, J.: Gradient _eld descriptor for sketch based image retrieval and localization. In: International Conference on Image Processing, pp. 1–4 (2010)
16. Jain, A.K., Lee, J.E., Jin, R.: Sketch to photo matching: a feature-based approach. In: Proc. SPIE, Biometric Technology for Human Identification VII, vol. 7667, pp. 766702–766702 (2010)