International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 06 | June-2018 www.irjet.net p-ISSN: 2395-0072

Feature Based Image retrieval based on Color

Purbani Kar[1], Lalita Kumari[2]

[1,2]*Assistant Professor, Dept. of Computer Science & engineering, NIT Agartala, Tripura, India*

Abstract - *Image retrieval is the field of study concerned with searching and retrieving digital images from a collection of database. Content-based image retrieval plays a central role in the application area such as multimedia database systems in recent years. The emergence of multimedia technology and the rapidly expanding image and video collections on the Internet have attracted significant research efforts in providing tools for effective retrieval and management of visual data. The fundamental idea of this approach is to generate automatically image descriptions directly from the image content by analyzing the content of the images. Feature extraction is the basis of content-based image retrieval. This involves extraction of the image features at a distinguishable extent. For color based image retrieval, color features are the most important elements enabling human to recognize images. For categorizing images color features can provide powerful information and they are used for image retrieval, so color based image retrieval is mostly used method. Color features of the images are generally represented by color histograms. Before using color histograms, however, we need to select and quantify a color space model and choose a distance metric.*

Key Words*:Content based image retrieval,Color histogram, Image features, Color space model, Distance metric*

1. INTRODUCTION

With the rapid development of image capture technology and internet, the number of digital image is getting larger and larger. Traditional keyword-base retrieval method cannot work efficiently anymore. It is a process that organized and stored the information according to a certain way, and in accordance with the needs of users to find the interrelated information, it is also called Information Storage and Retrieval.The main method of image files is to establish Keywords or text description of the title as well as some additional information, and then establish a link between storage path and the keywords of the image, which is text-based image retrieval.

However, with the storage capacity of images to start using GB or TB, the own shortcomings of text-based image retrieval technology led to two difficulties in the retrieval: First, it has been impossible to note each image; second, the subjectivity and nonprecision of image annotation may lead to the adaptation in the retrieval process.

In order to overcome these problems, a Great progress has been achieved in the field of Content Based Image Retrieval(CBIR) in recent years. The color feature, texture feature, shape feature and affine invariant features have all been used in image retrieval. Of all image content features, color and texture are two important features and play an important role in image content. Color histogram has the advantages of transform invariant, rotate invariant and scale invariant and has been widely used in image retrieval. In This paper Three dimension color space HSV is used and by using distance metrics a no of images similar to the query images are retrieved. [5]

2. RELATED WORK

Content Based Image Retrieval using Texture, Color and Shape for Image Analysis by Amanbir Sandhu and Aarti Kochhar is a paper which describes content based image retrieval based on features like color, texture and shape. For Texture feature extraxtion model they have considered Gray Level Cooccurrence matrix (GLCM). [1]

Petteri Kerminen and Moncef Gabbouj has developed a image retrieval technique based on color matching.They have discussed three color spaces RGB, L*a*b* and HSV/I.They have mentioned the advantages and disadvantages of these three color spaces.And also they mentioned the color quantization and color histogram.In their system the RGB components are analyzed pixel by pixel and their real combination is the most important thing (i.e. if a point has the values 150, 10,150 it is light violet).[2]

In Color Based Image Retrieval System by Pawandeep Kaur1, Sakshi Thakral2, Mandeep Singh3 they have implemented a system quite similar to us.They have been implemented on the image database and one query image is chosen for getting images having almost same histogram.They used the HSV color model, Minkowski-form distance metrics etc in their system. They have mentioned four requirements of color based image retrieval system. Such as Technique to obtain the metadata, having primitive features of images, Users query demands evaluated by interfaces used, methods to compare the similar or different images, Efficient indexing and metadata storage techniques.[3]

Avneet Kaur, V. K. Banga mentioned Color moments, color histograms, color coherence vector, color correlogram for their system.Color moments are the statistical moments of the probability distributions of colors. Color moments used especially when image contain just the objects. [4] The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be effective and efficent in representing color distribution of images.

Research of Image Retrieval Based on Color by Bai Xue, Liu Wanjun discussed two methods of feature extraction Color histogram and Color Moment.In view of the defects of the two extraction method of color feature color histogram and color moment method, they select an integrated approach of the two methods to extract the color feature in this paper in order to improve the retrieval accuracy and reform the ranking. [5]

3. TECHNOLOGIES USED

Color features of the images are generally represented by color histograms. Before using color histograms, however, we need to select and quantify a color space model and choose a distance metric.

3.1 Color Model:

There are many color models to express color such as the RGB color model, YUV color model and the HSV color model. In this system we have used the HSV color space.

The computer can only identify the RGB color component of an image, in which R represents the red component, G represents the green component, B represents the blue component[5].

Therefore, we need the following formula for the image conversion from RGB color space to HSV color

$$h' \models \begin{cases} \frac{(g-b)}{\delta} & \text{if r=max } h = h' * 60 \\ \frac{2+(b-r)}{\delta} & \text{if g=max } s = \frac{max-min}{max} \\ \frac{4+(r-g)}{\delta} & \text{if b=max } v = \frac{r+g+b}{3} \end{cases}$$

3.2 Three-Dimensional Color Histogram:

3.2.1 Quantization of HSV Color Vectors:

A (16:4:4) non-uniform quantization method is adopted in which H vector is divided into 16 values and S, V are divided into 4 values separately. The S, V and H values after quantization are shown below.

$$S = \begin{cases} 0.075 & s \in [0, 0.15] \\ 0.275 & s \subset [0.15, 0.4] \\ 0.575 & s \subset [0.4, 0.75] \\ 0.875 & s \in [0.75, 1] \end{cases}$$

$$V = \begin{cases} 0.075 & v \in [0, 0.15] \\ 0.275 & v \in [0.15, 0.4] \\ 0.575 & v \in [0.4, 0.75] \\ 0.875 & v \in [0.75, 1] \end{cases}$$

$$H = \begin{cases} 0 & h \in [345, 360] \text{or } (0, 15) \\ 20 & h \in [16, 25] \\ 35 & h \in [26, 45] \\ 50 & h \in [46, 55] \\ 68 & h \in [56, 80] \\ 94 & h \in [81, 108] \\ 124 & h \in [109, 140] \\ 153 & h \in [141, 165] \\ 178 & h \in [166, 190] \\ 205 & h \in [191, 220] \\ 238 & h \in [221, 255] \\ 265 & h \in [256, 275] \\ 283 & h \in [276, 290] \\ 303 & h \in [291, 315] \\ 323 & h \in [316, 330] \\ 338 & h \in [331, 345] \end{cases}$$

3.2.2 The Construction of Three-Dimensional Color Histogram:

In the image after quantization, its H vector has 16 values and S, V has 4 values separately. So this paper defines an array of 16*4*4 size which is T to calculate the ratio of pixels of each color to the overall pixels. The elements in T is defined in

$$T(i, j, k) = \frac{N_{i,j,k}}{M}, (1 \le i \le 16, 1 \le j, k \le 4)$$

Where T(i, j, k) means the ratio of pixels whose color value are the i^{th} value in H, the j^{th} value in S and the k^{th} value in V to the overall pixels of the image.$N_{i;j;k}$ is the number of pixels whose color values are values mentioned above. M is the overall pixel number of the image.

3.2.3 Similarity measures:

Define sample image as I and the image to be matched is J and their three dimensional color histograms are F_I and F_J. Histogram intersection is performed to determine similarity of these two images.

$$S(I, J) = \frac{\sum_{i=1}^{16} \sum_{j=1}^{4} \sum_{k=1}^{4} min(F_I(i, j, k), F_J(i, j, k))}{2M}$$

Where M is the overall number of pixels in the image,min () is a function that can give the smallest value. The value range of S (I,J) is 0 to 1. The more similar the two images are, the bigger the value of S is. For the same images, S is 1. Many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features in recent years. We denote D (I, J) as the distance measure between the query image I and the image J in the database; and f_i (I) as the number of pixels in bin i of I.

3.2.4 Minkowski-Form Distance:

Minkowski-form distance is the most widely used metric for image retrieval. Minkowski-form distance metrics compare only the same bins between color histograms. .If each dimension of image feature vector is independent of each other and is of equal importance, the Minkowski-form distance Lp is appropriate.

$$D(I,J) = (\sum_i |f_i(I) - f_i(J)|p)^{1/p}$$

When p=1, 2, and, D (I, J) is the L_1, L_2 (also called Euclidean distance), and L distance respectively. The Histogram intersection can be taken as a special case of L_1 distance, which is used by Swain and Ballard to compute the similarity between color images.

$$S(I,J) = \frac{\sum_{i=1}^{N} min(f_i(I), f_i(J))}{\sum_{i=1}^{N} f_i(J)}$$

Where A= [a_{ij}] is a similarity matrix, and a_{ij} denotes the similarity between bin F_i and F_j. and are vectors that list all the entries in and Quadratic form distance can lead to perceptually more desirable results than Euclidean distance.

3.2.5 User Interaction

Two main features of user interaction in image retrieval system are as follows:

3.2.6 Query Specification

Various types of queries can be listed as simple visual feature query, feature combination query, localized feature query, query by example, user-defined attribute query, object relationship query, and concept queries.

User can have two ways to make distinction. In first way user looks for category search and in second method he searches for target search.

3.2.7 Relevance Feedback

The iterative and automatic refinement of a query is known as relevance feedback in information retrieval literature.

Relevance feedback can be seen as a form of supervised learning to adjust the subsequent queries using the information gathered from the user's feedback.

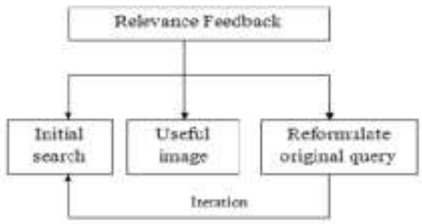

Fig-1: Flow of relevance feedback

4. STEPS OF THE RETRIEVAL SYSTEM

4.1Training

Almost 700 images have been used for populating the database. For each image a 3-D histogram of its HSV values is computed. At the end of the training stage, all 3D HSV histograms are stored in the same .mat file.

4.2Query

In order to retrieve M (user-defined) query results, the following steps are executed:

1. The 3D (HSV) histogram of the query image is computed. Then, the number of bins in each direction (i.e., HSV space) is duplicated by means of interpolation.

2. for each image i in the database:

- Load its histogram Hist (i).
- Use interpolation for duplicating the number of bins in each direction.
- For each 3-D hist bin, compute the distance (D) between the hist of the query image and the i-th database image.
- Keep only distances (D_2) for which, the respective hist bins of the query image are larger than a predefined threshold T (let L_2 the number of these distances).
- Use a 2nd threshold: find the distance (D_3) values which are smaller than T_2, and let L_3 be the number of such values.
- The similarity measure is defined as: S (i) = L_2 * average (D_3) / (L_{32}).

3. Sort the similarity vector and prompt the user with the images that have the M smaller S values.

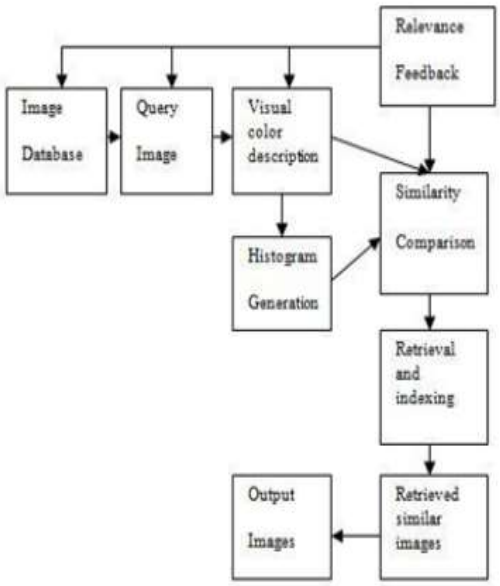

Fig-2: Image retrieval system based on color

5. RESULTS AND DISCUSSION

The methodology discussed in the earlier chapter has been implemented on the image database and one query image is chosen for getting images having almost same histogram. HSV histogram is used for comparison.

Here steps of implementation have been shown in 5.1 and 5.2.The experiments are performed under computer of Intel Core i5 CPU, 2.30Hz, 4G RAM and with operating system of Windows 7.

The size of images is 250*225. For an image set of 606 images the average time is 4 second which is very fast.

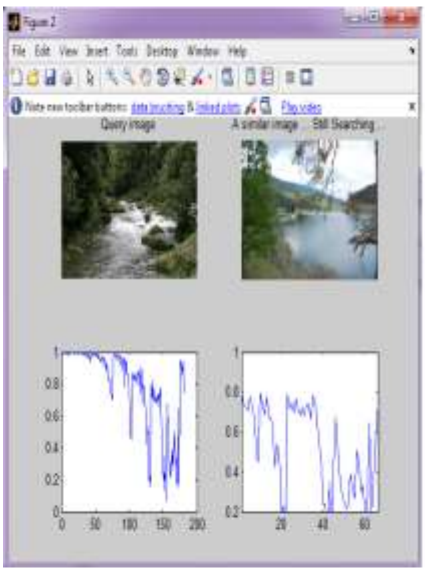

Fig-3:Images Get Compared And Their Equivalent Hsv Histograms

Fig-4: Eleven most resembling images with the query image

6. CONCLUSION AND FUTURE SCOPE

Content Based Image Retrieval is now became a burning topic in computer science. It is getting more and more importance to practical applications.This paper introduces an effective image retrieval method which is based on the color feature.

Color is usually represented by the color histogram, color correlogram, color coherence vector, and color moment under a certain color space. Here, in this paper 3D-HSV color histogram has been used.Up to now; the Minkowski and Quadratic form distance are the most commonly used distances for image retrieval. To set up an indexing scheme, dimension reduction is usually performed to reduce the dimensionality of the visual feature vector. Query results can be refined through the relevance feedback of users.

Although color-based image retrieval provides an intelligent and automatic solution the majority of current techniques are based on low level features. In general, each of these low level features tends to capture only one aspect of an image property. Neither a single feature nor a combination of multiple features has explicit semantic meaning. Although relevance feedback provides a way of filling the gap between semantic searching and low-level data processing, this problem remains unsolved and more research is required.

REFERENCES

[1] Content Based Image Retrieval using Texture, Color and Shape for Image Analysis,Amanbir Sandhu ,Rayat Bahra College of Engg & Nanotechnology for Women,Hoshiarpur (Pb), Aarti Kochhar, Council for Innovative Research International Journal of Computers & Technology www.ijctonline.com ISSN: 2277-3061 Volume 3, No. 1, AUG, 2012

[2] IMAGE RETRIEVAL BASED ON COLOR MATCHING, Petteri Kerminen, Moncef Gabbouj, Pori School of Technology and economica, TUT(Tampere University of Technology), Information Technology Signal

Processing Laboratory P.P Box 30, P.O Box 553 28601 Pori 33101 Tampere.

[3] Color Based Image Retrieval System Pawandeep Kaur, Sakshi Thakral, Mandeep Singh,IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Volume 1, Issue 5 (May-June 2012), PP 01-05 www.iosrjournals.org

[4] Color Based Image Retrieval Avneet Kaur, V. K. Banga

[5] Research of Image Retrieval Based on Color Bai Xue, Liu Wanjun School of Software, Liaoning Technical University, Hulidao, 125105, China, 2009 International Forum on Computer Science-Technology and Applications xiaopanghuamao@163.com

[6] Color Image Retrieval Based on Color and Texture Features, Xiuxin Chen, Kebin Jia, Xiaoqin Lian, and Shiang Wei , D. Jin and S. Lin (Eds.): Advances in CSIE, Vol. 1, AISC 168, pp. 669674. Springerlink.com Springer-Verlag Berlin Heidelberg 2012

[7] Efficient CBIR Using Color Histogram Processing, Neetu Sharma, Paresh Rawat and jaikaran Singh ,Signal & Image Processing : An International Journal(SIPIJ) Vol.2, No.1, March 2011

[8] Color-based retrieval Nicu Sebe *, Michael S.Lew,Leiden Institute of Advanced Computer Science, Department of Computer Science, Niels Bohrweg 1, Leiden, CA 2333, The Netherlands Received 21 February 2000; received in revised form 14 July 2000

[9] Dr. Fuhui Long, Dr. Hongjiang Zhang, Prof. David Dagan Feng, Fundamentals of contentbased Image retrieval, Chapter 1.

[10] Swain MJ and Balland DH. Color indexing of Computer Vision [J], 1991-7(1):11-32.