Image Retrieval Using Histogram Based Contents of an Image

Arun JB,
Teacher's Training Center
Govt. Polytechnic Collage
Jodhpur,India

Reshu Choudhary
Electronics and Communication Engineering
M.B.M. Engineering Collage,J.N.V.U.
Jodhpur,India

Abstract

The advances in data storage and image acquisition techniques have enabled the creation of huge image databases. Retrieving desired images from such large and varied collection is a challenging problem, to solve this problem, text-based and content-based are the two techniques adopted for search and retrieval in an image database. The Content-based image retrieval system has been an active research area due to its effectiveness. The system applies image contents of an image correspond to objects. This paper presents a novel image feature representation method, namely histogram based technique, for image retrieval to enhance indexing and retrieval performance. The color histogram for an image is constructed by counting the number of pixels of each color. For computation of histogram we have to count number of pixels that correspond to a specific color in quantized color space RGB. Finally we get efficient images matching to our query image using histogram comparison technique. In order to compare histograms of two images, we first need to generate specific codes for all histogram bins. In this technique database images are in any position either flipped (horizontal/ vertical) or rotated to any angle can be successfully retrieved with high precision and recall rates. This system has also a function to measure retrieval effectiveness is precision vs. recall graph for an image regarding histogram-based retrieval method.

1. Introduction

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common method of image retrieval is to adding metadata such as captioning, keywords. Manual image annotation is time-consuming, laborious and expensive; to address this using a Content Based Image Retrieval (CBIR), images can be analyzed and indexed automatically which depends on their objective visual content. This paper presents a histogram based technique for image retrieval. A color histogram is a representation of the distribution of colors in an image. The color histogram can be built for any kind of color space, although the term is more often used for three dimensional spaces like RGB or HSV [1], [2].The histogram provides a compact summarization of the distribution of data in an image [3-6]. For digital images, a color histogram represents the number of pixels that have colors in each of a fixed list of color ranges; that span the image's color space, the set of all possible colors. The color histogram of an image is relatively invariant with translation and rotation about the viewing axis, and varies only slowly with the angle of view. By comparing histograms of two images and matching the color content of one image with the other, the color histogram is particularly well suited for the problem of recognizing an object of unknown position and rotation within a scene.

Another approach to representative color image content is two-dimensional color histogram [7-10]. A two-dimensional color histogram considers the relation between the pixel pair colors (not only the lighting component). A two-dimensional color histogram is a two-dimensional array. The size of each dimension is the number of colors that were used in the phase of color quantization. These arrays are treated as matrices, each element of which stores a normalized count of pixel pairs, with each color corresponding to the index of an element in each pixel neighborhood. While creating a set of final images, the comparison of two-dimensional color histograms should be arranged. In this technique, a histogram is computed from all of the pixels in the image, and the peaks and valleys in the histogram are used to locate the clusters in the image. Color or intensity can be used as the measure.

2. Literature Review

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of images. The first microcomputer-based image database retrieval system was developed at MIT, in the 1990's, by Banireddy Prasaad, Amar Gupta, Hoo-min Toong, and Stuart Madnick.

Some software are presented to store and retrieve primitive data and as well as complex data like images. The image type, size, color and texture characteristics are extracted from the images and

IJERT

stored into the database [11-15]. This type of software has its own content based Retrieval module that allows users to build content based visual queries to the image level. All these are done by some background programming of the system. The image features are extracted from the original image and stored as metadata. And for the query image the same features are extracted from the image and compared with each other.

Many approaches are available based on histogram extraction technique, [16], [17] but the color coherence vector approach gave a new blow to previous histogram based approaches. This color coherence vector extracts not only the color distribution of pixels in images like color histogram, but also extracts the spatial information of pixels in the images. It gives us a more sophisticated approach towards histogram refinement. Use of multiple color coherence vectors gives much better efficiency than single one though it has higher computational complexity [4]. Some efficient work on histogram is done to detect image copy also. In this scheme multi resolution histogram is used. It is almost same like the plain color histogram method. But it adds some extra feature like Encoding of spatial information directly.

In image enhancement, the image is to form better suited for analysis by a human or machine processes consist of a collection of techniques that seek to improve the visual appearance of an image or to convert the image to a form better suited for analysis by a human or machine.

IJERT

2.1 Histogram Equalization

This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values [18].

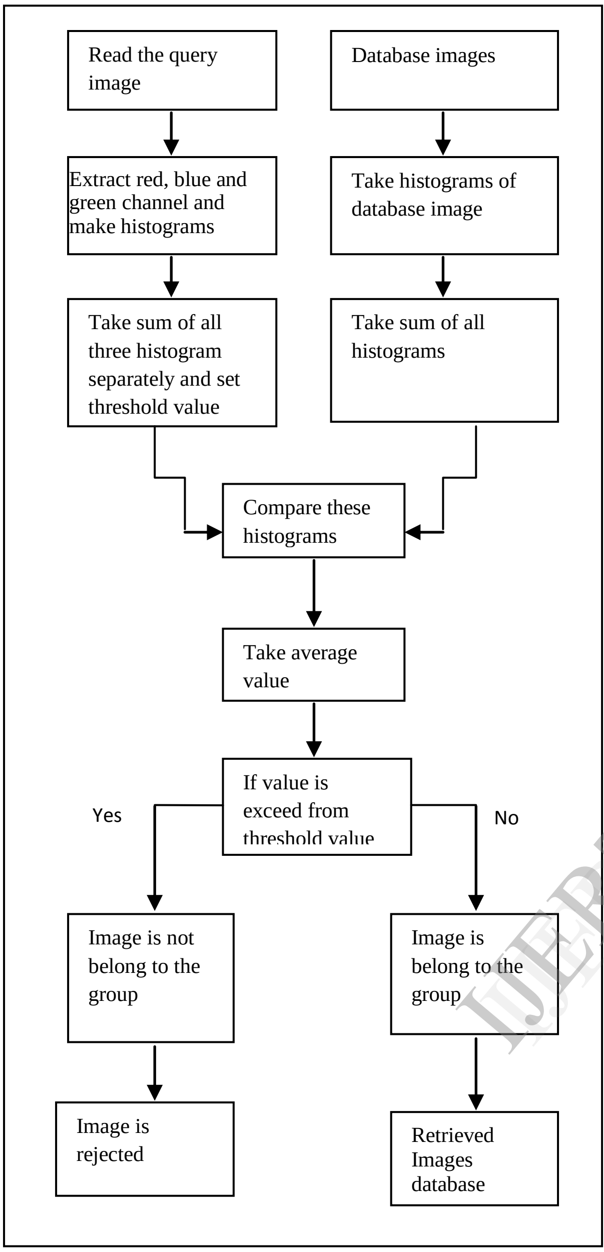
Histogram equalization automatically determines a transformation function seeking to produce an output image with a uniform Histogram. Histogram equalization is a method in image processing of contrast adjustment using the image histogram. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values [19]. Through this adjustment, the intensities can be better distributed on the histogram. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. Histogram equalization automatically determines a transformation function seeking to produce an output image with a uniform histogram.

2.2 Geometric Histogram

Geometric histogram generalized the color spatial distribution by computing the color histogram with specific geometric relationship between pixels of each color histogram bucket. The concept is a unification of some existing techniques such as color density maps, color correlogram and color tuples. A color density map is a set of pixels with the same color that considered as a geometric subset of the 2-D plane. The centroid and the radius of the subset are calculated and the number of pixels in each of the regions is computed to form a vector called the density map of color. The map is arranged of all colors into a matrix by making the density vector of each color as a row [20], [22]. On the other hand, color correlogram [1] is a vector of three indices. Geometric histogram is almost the same with region based algorithm.

3. Proposed Method

This proposed methodology involves study of image histogram. An image is a function of two variables. It is simply represented by a 2D matrix after the digitization is done. If a pixel in that image is black then the pixel value is assigned to ‘0’ and in case of a white pixel the pixel value is assigned to ‘255’. For other pixels some other values in between 255 and zero is assigned. Color histogram of an image is a type of bar graph and these acts as a graphical representation of the tonal distribution in a digital image. The number of pixels is plotted for each tonal value. An image viewer can understand the entire tonal distribution of an image by looking at the histogram bar graph. Proposed algorithm is as follows:



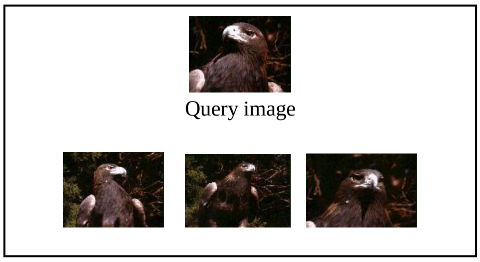


Fig.2 Retrieve Results for Eagle

The results of the above algorithm are shown in table 1. The parameter 1, parameter 2 and parameter 3 are the histogram, centers and maximum value respectively. The histogram data are not processed for any treatment. These three parameters are evaluated over selected database images. We use 158 images from our database out of them 151 images are not matched and 7 images are matched to our query image.

Images in a database are in any position can retrieved. Take an example of eagle image as query image. Images are present vertically/horizontal flipped and rotated in 90º, 180º and 270º in database.

IJERT

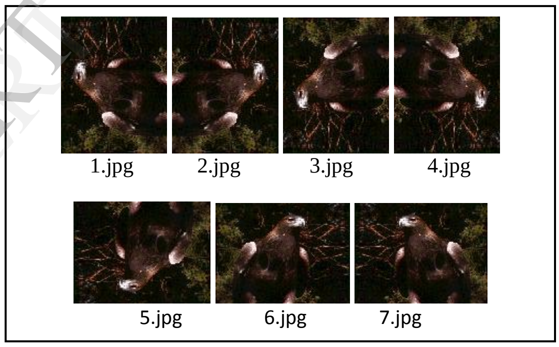


Fig.3 Rotated and Flipped Retrieved Images in database

3.1. Experimental Results

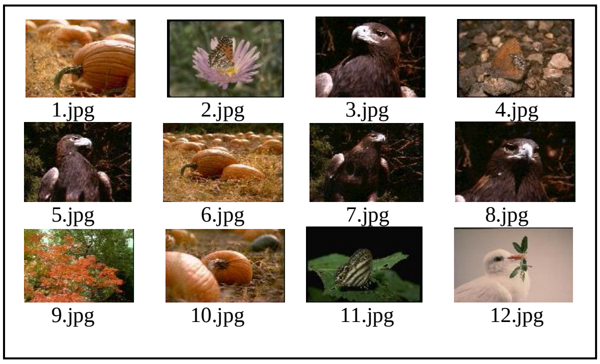


Fig.1 Sample of Image Database

These all images are retrieved with 100% accuracy rate for query image.

3.2 Performance Evaluation

The performance of retrieval of the system can be measured in terms of its recall and precision. Recall measures the ability of the system to retrieve all the models that are relevant, while precision measures the ability of the system to retrieve only the models that are relevant. It has been reported that the histogram gives the best performance through recall and precision value [21],[22]. They are defined as:

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images}}$$

TABLE I. PRECISION AND RECALL VALUES

Parameter 1		Parameter 2		Parameter 3	
Precision	*Recall*	*Precision*	*Recall*	*Precision*	*Recall*
1	0.143	1	0.143	1	0.143
1	0.428	1	0.143	1	0.286
1	0.57	1	0.428	0.8	0.57
0.833	0.714	0.75	0.428	0.45	0.714
0.625	0.714	0.714	0.714	0.385	0.714
0.5	0.857	0.625	0.714	0.357	0.714
0.31	0.857	0.55	0.714	0.33	0.714
0.269	1	0.46	0.857	0.286	0.857
0.23	1	0.315	0.857	0.27	0.857

The precision and recall graphs are shown in Fig. 3, 4 and 5 for parameter 1, 2 and 3 respectively.

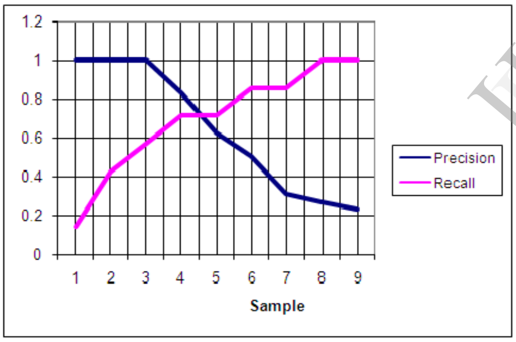


Fig.4 Graph for Parameter 1

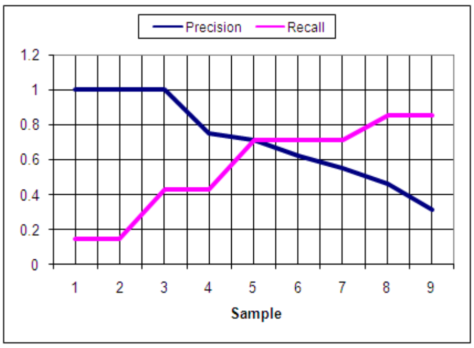


Fig.5 Graph for Parameter 2

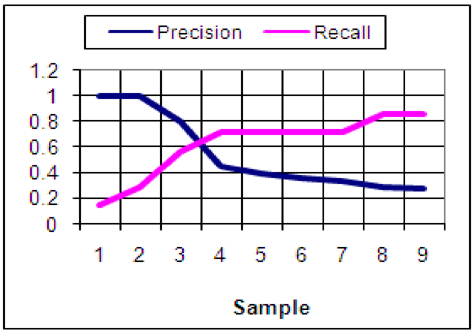


Fig.6 Graph for Parameter 3

4. Conclusion

An approach to image retrieval is proposed but it has endless scope to work with. Endless discussion can be done with it. By different approaches it has been seen that the result and retrieval ratio depends upon the image class, for some images it gives good retrieval ratio while poor for some other ones. In this paper, we presented a novel approach for image retrieval by histogram technique without applying any processing techniques. The experimental result shows that the proposed method outperforms the other retrieval methods in terms of Average Precision. We can say that for our implemented system, the image classes used by us give very good image retrieval accuracy. Our system performance is quite reasonable as per the accuracy graph shown above. This approach concentrates only on retrieving of image files. There are scopes for future that the values of thresholds can be set according to the type of image to be retrieval.

IJERT

5. References

[1] J. Huang and S.K. Ravi, "Image Indexing Using Color Correlograms", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico, June 1997.
[2] G. Pass and R. Zabih, "Refinement Histogram for Content-Based Image Rtrieval", IEEE Workshop on Application of Computer Vision, pp. 96-102.,1996.
[3] Jozsef Vass, Jia Yao, Anupam Joshi, Kannappan Palaniappn and Xinhua Zhuang, "Interactive Image Retrieval over the Internet", Seventeenth IEEE Symposium on Reliable Distributed Systems, pp. 461-466, 20-23 October 1998.
[4] J.C. Baillie, "Certainty Color Maps Compared to Histograms", International Conference on Image Processing, Vol. 3, 24-28 June, 2002
[5] Ryszard S. Choras, "Image Feature Extraction Techniques and their Applications for CBIR and Biometrics Systems", International Journal Of Biology And Biomedical Engineering, Vol. 1, Issue 1, 2007.
[6] Baochang Zhang, Shiguang Shan, Xilin Chen, and Wen Gao, "Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition", IEEE Transactions on Image Processing, Vol. 16, No. 1, January 2007.
[7] Qi Chen, Bo-li Xiong, Jun Lu and Gang-yao Kuang, "An Approach on Analyzing Histogram and Selecting Threshold", International Conference on Computer Science and Software Engineering, Vol. 6, 12-14 December 2008.

[8] Wang Shoujue and Qin Hong, "A Study of Order-Based Block Color Feature Image Retrieval Compared with Cumulative Color Histogram Method", Sixth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD'09, Vol. 1, 14-16 August 2009

[9] H. B. Kekre and Kavita Sonavane, "CBIR Using Kekre's Transform Over Row Column Mean And Variance Vector", International Journal on Computer Science and Engineering (IJCSE), Vol. 02, No. 05, 1609-1614, 2010.

[10] Ramesh Babu Durai, "A Generic Approach To Content Based Image Retrieval Using Dct And Classification Techniques", International Journal on Computer Science and Engineering (IJCSE), Vol. 02, No. 06, 2022-2024, 2010.

[11] G. Raghupathi, R.S. Anand, and M.L Dewal, "Color and Texture Features for Content Based Image Retrieval", Second International Conference on Multimedia and Content Based Image Retrieval, 21-23 July 2010.

[12] Aliaa.A.A Yousif, A.A Darwish and R.A. Mohammad, "Content Based Medical Image Retrieval Based on Pyramid Structure Wavelet", IJCSNS International Journal of Computer Science and Network Security, Vol.10, No.3, March 2010.

[13] Swati V. Sakhare and Vrushali G. Nasre, "Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture", International Journal of Computer Science and Informatics, Vol. I, Issue II, 2011.

[14] P.L. Mazzeo, L. Giove, G.M. Moramarco and P. Spagnolo, "HSV and RGB Color Histograms Comparing for Objects Tracking Among Non Overlapping FOVs, using CBTF", 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS), Augu 30 - September. 2, 2011.

[15] Latika Pinjarkar, Manisha Sharma and Kamal Mehta, "Comparison and Analysis of Content Based Image Retrieval Systems Based On Relevance Feedback", Journal of Emerging Trends in Computing and Information Sciences, Vol. 3, No. 6, July 2012.

[16] Poulami Haldar and Joydeep Mukherjee, "Content Based Image Retrieval Using Histogram,Color and Edge", International Journal of Computer Application, Vol. 48, No.11, June 2012.

[17] R. Venkata Ramana Chary, D. Rajya Lakshmi and K.V.N Sunitha, "Image Retrieval Techniques for Color based Images from Large Set of Database", International Journal of Computer Applications, Vol. 40, No. 4, February, 2012.

[18] Vivek Kapur, P.T. Karule and M.M. Raghuwanshi, "Content Based Image Retrieval Using Soft Computing Technique", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2013), March 17-18, 2013.

[19] Hazem Hatamleh, "Image Processing of Two Identical and Similar Photos", Journal of Information Engineering and Applications, Vol.3, No.1, 2013.

[20] S. Gorji Kandi and K. Ansari, "Color Scene Transform Between Images Using Rosenfeld-Kak Histogram Matching Method", Progess Color Colorants Coatings, Vol. 6, pp.17-24, 2013.

[21] Shilpa P. Pant, "Content Based Image Retrieval Using Color Feature", International Journal of Engineering Research and Technology (IJERT), Vol. 2, Issue 4, April, 2013.

[22] E. Kostopoulou, E. Zacharia and D. Maroulis, "An Effective Approach for Detection and Segmentation of Protein Spots on 2D-Gel Images", DOI.10.1109/JBHI,2259208,2013.

IJERT