Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745
ISSN (PRINT) : 2393 - 8161

AN EFFICIENT IMAGE RETRIVAL BASED ON HSV COLOR SPACE AND KNN CLASSIFER

E.Venkateswarlu [1], D.Lalitha Kumari[2]
[1]M.Tech Student, JNTU College of Engineering &Technology,Anatapur
[2]Assistant Profesort, JNTU College of Engineering &Technology,Anatapur

Abstract—This electronic document is a "live" template. The various components of your paper In this paper, it is proposed an efficient image retrieval scheme to retrieve images. The numbers of images retrieved that are irrelevant are more using HSV Color Space. HSV stands for the Hue, Saturation and Value, provides the perception representation according with human visual feature. The retrieval accuracy is better in KNN Classifier compared to HSV Color Space. The experimental database contains 50 images and comparing the results of the HSV Color Space and KNN Classifier. The KNN Classifier is better performance than the color space.
Keywords*—Retrieving Images, CBIR, histogram intersection KNN classifier*

I. INTRODUCTION

In the recent years, Content-based image retrieval (CBIR) has been an active research area in image processing. The tremendous growth of digital image and video on internet, so the traditional text-based retrievals based on keywords are not sufficient to resolve image retrieval. CBIR can greatly enhance the accuracy and efficiency of retrieving and managing the data of image. k- Nearest Neighbor (KNN) is one of the most popular algorithm and Many researchers have found that the KNN algorithm accomplishes very good performance in their experiments on different data sets.. Chabot system employs text along with color histogram and integrates a relational database to retrieve image. The features of Content-based image such as color, shape, texture and outline are used for image retrieval. Among these features, color is an important feature in CBIR, which is invariant on size, orientation and complexity. In this paper, we proposed the novel feature extraction technique using HSV color space instead of RGB color space and KNN Classifier for image retrieval.

The paper is organized as follows.KNN Classifier is presented in Section II. Section III describes the HSV Color Space.The similarity measure Histogram Intersection is presented in Section IV. Section V describes the conclusion.

II. KNN CLASSIFIER

A generalized algorithm called KNN Classifier, in which a new pattern is classified into the class with the most members present among the K nearest neighbors. To find out the k nearest neighbor samples, all the similarities between the training samples must be calculated. When the number of training samples is less, the KNN classifier is no longer optimal, but if the training set contains a huge number of samples, the KNN classifier needs more time to calculate the similarities [3]. The KNN classification algorithm depends on three aspects .one is calculation complexity due to the usage of all the training samples for classification, second is that the performance is solely dependent on the training set, and third is that there is no weight difference between samples.

The procedure of KNN Classifier is shown below. Select training data set, in which all training samples are relatively similar to the test sample

a. Select Test sample
b. Compute distance metric between test sample and training samples.The value of k determines the number of nearest neighbors to be retrieved.

Consider the majority vote of class labels among the k-nearest neighbors, that determines the class
1) Repeat step 1 to 5 for another test sample

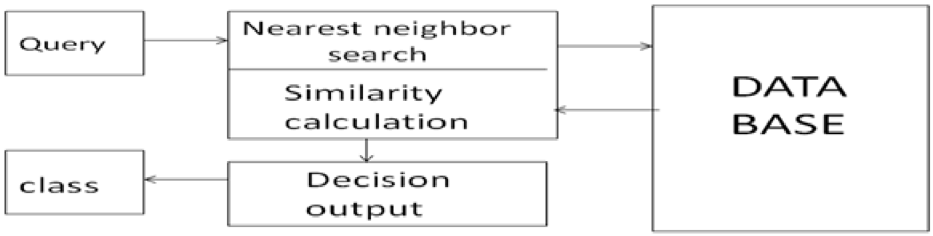

Figure1. KNN Classifier

In order to measure similarity measure between query image and database images, we use some kind of distance function called Euclidean distance, to determine how much two histograms differ from each other. When matching process is completed, results are sorted in ascending order and retrieval images are presented. Euclidean distance is defined in equation.

III. HSV COLOR SPACE

The HSV stands for the Hue, Saturation and Value, provides the perception representation according with human visual feature. The HSV model, defines a color space in terms of three constituent components: Hue, the color type Range from 0 to 360. Saturation the "vibrancy" of the color: Ranges from 0 to 100%, and occasionally is called the "purity". Value the brightness of the color: Ranges from 0 to 100%. To decrease the number of colors used in image retrieval, we quantize the number of colors into several bins. J.R. Smith [3] designs the scheme to quantize the color space into 166 colors. Li [5] design the non-uniform scheme to quantize into 72 colors. We propose the scheme to produce 15 non-uniform colors. The formula that transfers from RGB to HSV is defined as below:

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

$$S = 1 - 3/(R+G+B)(\min(R,G,B)) \qquad \text{----- (2)}$$

$$V = 1/3(R+G+B) \qquad \text{-----(3)}$$

The proposed scheme for HSV color space contains three phases [2]. First of all we resize all images to reduce the size of images and processing time. Secondly we convert each pixel of resized image to quantized color code. Finally we compare the quantized color code between the query image and database image. In conventional schemes[3], they extract the image feature vector from images using Histogram Intersection to measure the similarity of image for matching between a query image and image from database[6].When matching processing is completed, results are sorted in ascending order and retrieval images are presented.

Volume 02, Issue 04, [April – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

IV.FEATURE EXTRACTION

For the feature extraction, we have used the Euclidian distance to measure the similarity between the Query image and database image using KNN classifier. There are several distance formulas for measuring the similarity of Color histograms. The Color distance formulas arrive at a measure of similarity between images based on the perception of Color content. The distance formulas that are used normally for image retrieval are histogram Euclidean distance, histogram intersection distance.

Let h and g represent two Color histograms. The intersection of histograms $\boldsymbol{h}$ and $\boldsymbol{g}$ is given by:

$$d(h,g) = \frac{\sum_A \sum_B \sum_C \min(h(a,b,c), g(a,b,c))}{(mod\ h, mod\ g)} \quad \text{------(4)}$$

The Euclidean distance between two color histograms h and g is given by:

$$D^2(h,g)=\sum_A \sum_B \sum_C (h(a,b,c)-g(a,b,c))^2 \quad \text{-------(5)}$$

Where ***a, b, c*** are color components. When matching processing is completed, results are sorted in ascending order and retrieval images are presented.

V.EXPERIMENTAL RESULTS

This paper attempts to evaluate the performance of the CBIR system on sample datasets of images using KNN classifier. We have purposefully kept 10 relevant images of the particular query image image type in the data base folder along with other relevant images. We have designed a 5X5 matrix for showing the retrieved images in the result window. This project improves the efficiency of CBIR system by adopting KNN classifier, comparing the results obtained from histogram intersection with that of the KNN Classifier. To check the performance of proposed technique we have used precision and recall. The standard definitions of these two measures are given by the following equations.
Precision=No. of relevant images retrieved/Total no. of images retrieved
Recall = No. of relevant images retrieved/Total no. of relevant images in the database
Accuracy=Precision+ recall/2

Table 1. Retrieval accuracy using Histogram intersection

Class	No.of relevant	Total no.of images retrieved	No.of relevant	Precision	Recall	Accuracy

	images in data base		images retrieved			
African people	10	24	10	0.416	1	52.08%
Building	10	24	9	0.375	0.9	63.75%
Buses	10	24	10	0.416	1	70.80%
Dinosaurs	10	19	9	0.473	0.9	68.65%
Elephants	10	24	4	0.166	0.4	28.33%
Flowers	10	24	8	0.333	0.8	56.65%
Horses	10	19	8	0.421	0.8	61.05%
Beach	10	22	9	0.409	0.9	65.45%
Food	10	24	4	0.166	0.4	28.33%
Mountain	10	24	5	0.208	0.5	35.41%
Average						**53.04%**

Table 2. Retrieval accuracy using KNN Classifier

Class	No. of relevant images in data base	Total no. of images retrieved	No. of relevant images retrieved	Precision	Recall	Accuracy
African people	10	9	7	0.777	0.7	73.85%
Building	10	9	6	0.666	0.6	63.30%
Buses	10	9	8	0.888	0.8	84.40%
Dinosaurs	10	9	7	0.777	0.7	73.85%
Elephants	10	9	5	0.555	0.5	52.75%
Flowers	10	9	4	0.444	0.4	42.20%
Horses	10	9	4	0.444	0.4	42.20%
Beach	10	9	3	0.333	0.3	31.65%
Food	10	9	6	0.666	0.6	63.30%
Mountain	10	9	5	0.555	0.5	52.75%
Average						**58.02%**

VI. CONCLUSION

In this project we have illustrated two similarity measures. Select a query image histogram from all histograms of all data base images.KNN classifier can be very effective if an analysis of neighbors is required and improved the retrieval accuracy than the HSV color space quantization(8,2,2) with histogram intersection distance under a uniform size of 20X20.

VII. REFERENCES

[1] *AIRS - Advanced Image Retrieval Service.*
http://www.imageclick.com/airs/sub/aboutAIRs.html (accessedOctober 2008).

[2] KP.Andriole, "The Society for Computer Applications in Radiology Transforming the Radiological Interpretation Process (TRIP) Inititative", November 2005. Last Accessed: March 26, 2007.
[3] Siddiqu Sharmin, "A Wavelet Based Technique for Analysis and Classification of Texture Images", Carleton University, Ottawa, Canada, Proj. Rep. 70.593, April 2002.
[4] SK.Chang, A.Hsu, "Image information systems: Where do we go from here?", IEEE Transactions on Knowledge and Information Engineering, 4(5), pp.644- 658, 1994.
[5] P.Agrain, H.Zhang, D.Petkovic, "Content-based representation and retrieval of visual media: a state of the art review", Multimedia Tools and Applications, 3(3), pp. 179-202, 1997.
[6] Jain, Anil K., and Aditya Vailaya. *Image Retrieval Using Color And Shape*. Great Britain: Elsevier Science Ltd, 1995.
[7] F.-D. Jou, K.-C. Fan, and Y.-L. Chang. Efficient matching of large-size histograms. *Pattern Recognition Letters*,25(3), pp. 277-286, 2004.
[8] Wang Shengjiu, "A Robust CBIR Approach Using Local Color Histograms", Tech. Rep. TR 01-13, October 2001.
[9] J.-K. Kamarainen, V. Kyrki, J. Ilonen, and
H. K¨alvi¨ainen. Improving similarity measures for
histograms using smoothing projections. *Pattern*
Recognition Letters, 24(12), pp. 2009-2019, 2003.
[10] Jeong, Sangoh. *Histogram-Based Color Image Retrieval*. Psych221/EE362 Project Report, Stanford University, 2001.
[11] Seidl Thomas and Kriegel Hans-Peter, "Efficient User-Adaptable Similarity Search in Large Multimedia Databases", in Proceedings of the 23rd International Conference on Very Large Data Bases VLDB'97, Athens, Greece, August 1997.
[12] J.P. Eakins, M.E. Graham, "Content-based image retrieval", Tech. Rep. JTAP-039, JISC Technology Application Program, Newcastle upon Tyne, 2000.
[13] A.Smeulders, M.Worring, S.Santini, A.Gupta, R.Jain, "Content-based image retrieval at the end of the early years", IEEE Trans Pattern Rec. & Machine Intel., 22(12):1349-80, 2000.
[14] Lai.Tring-Sheng, "A Photographic Image Retrieval System", School of Computing, Engineering and Technology University of Sunderland Sunderland, United Kingdom January 2000,
[15] HD. Tagare, CC. Jaffe, J. Duncan, "a content-based retrieval approach", 4(3):184-98, J Am Med Inform Assoc. May-Jun 1997.