A Novel Image Retrieval Method Using Segmentation and Color Moments

T.V. Saikrishna[1], Dr.A.Yesubabu[2], Dr.A.Anandarao[3], T.Sudha Rani[4]

[1]Assoc. Professor, Computer Science Department, QIS College of Engineering & Technology, Ongole, A.P., India.
[2]HOD & Professor, Computer Science Department, SIR C R Reddy College of Engineering, Eluru, A.P., India.
[3]Principal & Professor, Computer Science Department, JNTU College of Engineering, Anantapur, A.P., India.
[4]Sr. Asst. Professor, Computer Science Department, Aditya Engineering College, Kakinada, A.P., India.

ABSTRACT

Large image databases are becoming popular due to the ease with which images are being created/digitized and stored. Content Based Image Retrieval (CBIR) has therefore evolved into a necessity. It is a challenging task to design an effective and efficient CBIR system. Current research works attempt to obtain the semantics or meaning of the image to perform better retrieval. In this paper, we are going to propose a color image retrieval method based on the primitives of color moments. First, an image is divided into four segments. Then, the color moments of all segments are extracted and clustered into four classes. In this paper mean moments of each class are considered as a primitive of the image. All primitives are used as features and each class mean combined into single class mean. The distance between query image mean with the corresponding database images are calculated by using SAD method. Experiment results prove that the CBIR using our new measure has better performance than the existing methods.

KEYWORDS

Image, CBIR, SAD, color moments, segmentation & similarity measure.

1. INTRODUCTION

Since 1970's Image database management and retrieval has been an active research area [2]. With the speedy increase in computer speed and decrease in memory cost, image databases containing thousands or even millions of images which are used in many application areas such as medicine, satellite imaging, and biometric databases, where it is important to maintain a high degree of precision. Manual annotation becomes infeasible both time and cost-wise, with the growth in the number of images. Content-Based Image Retrieval (CBIR) is a powerful tool since it searches the image database by employing visual cues alone. CBIR systems extract features from the raw images themselves and calculate an association measure (*similarity* or *dissimilarity/distance*) between the query image and database images based on these features. Because of the high demand for searching image databases of ever-growing size, CBIR is becoming very popular. Since speed and precision are important, we need to develop a system for retrieving images that is efficient.

In CBIR systems [3-11], images automatically indexed by summarizing their visual features. A feature that can capture a certain visual property of an image either globally for the entire image or locally for regions or objects [7]. Color, texture and shape are commonly used features in

CBIR systems. Mapping the image pixels into the feature space is known as feature extraction. Feature extraction is a major function in any CBIR system. Those Extracted features are used to represent images for searching, indexing and browsing images in an image database. Use of feature space is more effective in terms of storage and computation. Feature space is represented as a feature vector in many CBIR systems. The similarity between images can be determined through features which are represented as vector. Different techniques are used to measure similarity in CBIR systems.

In retrieval stage, query image is also represented as a feature vector and the similarities between the query vector and stored feature vectors are computed [10]. The distance between the query image and stored images is determined through similarity measure. After that images are ordered according to the distance and retrieved. Visual indexing techniques can be categorized into two groups. Pixel domain and compressed domain techniques. Pixel domain techniques are used to index visual features such as color, texture and shape [9]. We have considered color feature which is extracted in Pixel domain.

1.1. Primitives of color moments extraction

The first order (mean), the second order (variance) and the third order (skewness) color moments have been proved to be efficient in representing color distribution of images [12]. The first three moments are defined as:

$$u_i = \frac{1}{N} \sum_{j=1}^{N} f_{ij} \qquad \qquad \dots (1)$$

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (f_{ij} - u_i)^2} \qquad \dots (2)$$

$$s_i = \sqrt[3]{\frac{1}{N} \sum_{j=1}^{N} (f_{ij} - u_i)^3} \qquad \dots (3)$$

Where f_{ij} is the value of the i-th color component of the image pixel j, and N is the number of pixels in the image.

Since only 9 (three moments for each of the three color components) [12] numbers are used to represent the color content of each image, color moments are a very compact representation compared to other color features. Due to this compactness, it may also lower the discrimination power.

2.Methodology: Color image retrieval

Content-based image retrieval (CBIR) is a new but widely adopted method for finding images from vast image databases. In CBIR images are indexed on the basis of low-level features, such as color, texture, and shape that can automatically be derived from the visual content of the images [1]. Here we propose an efficient approach for image retrieval based on color feature. Similar to most CBIR systems, we need to index images by extracting their features in an offline process. We then submit a query image and find similar images to that query based on a matching criterion. We first start with feature extraction. Figure1 represents proposed scheme architecture for this step. Firstly a database is prepared for different type of images. After this, analysis is performed on database. Analysis represents assessment of different descriptors used in this approach. Database is indexed according to values of different images. Finally the database is arranged on the basis of measures. When a user query is submitted for similarity matching, the step of analysis and feature selection is repeated as performed with image database. Now the value of query image is compared with the values of different images stored in database. As a result, the images having closest values compared to query image color values are extracted from database.
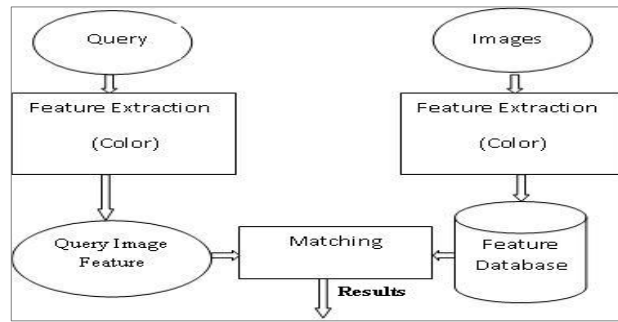
Fig 1: Proposed system architecture

According to figure1 when a query image is submitted for image retrieval, its color features are extracted and matching operation is performed between query image features and the image features stored in database. The result close to the query image is then retrieved from the database.
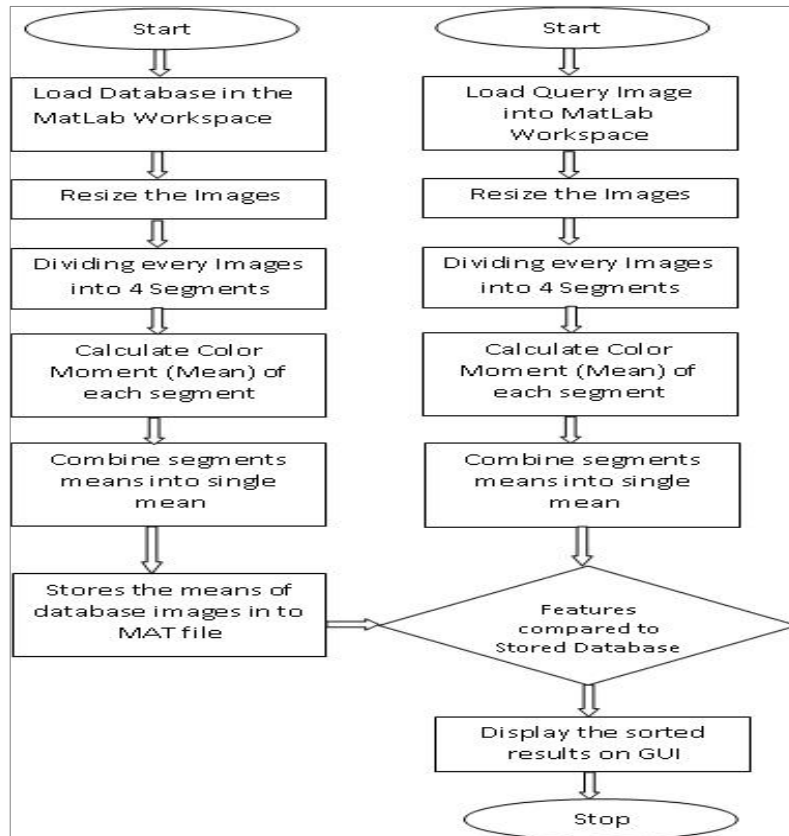

Fig 2: Flow chart for proposed method.

2.1. Algorithm for proposed Scheme

Step 1: Load image database in the MatLab workspace.
Step 2: Resize the image for [256, 256].
Step 3: partitioning database images into 4 segments.
Step 4: Calculate color moment (Mean) of each segment of the image.

Step 5: Combines each segment means into a single mean.
Step 6: Store the means of database images into MAT file.
Step 7: Load the Query image.
Step 8: Apply the procedure 2-6 to find mean of test image.
Step 9: Determine the Sum of Absolute Differences (SAD) of query image with stored feature of database.
Step 10: sort the Sum of Absolute Differences (SAD) values to perform indexing.
Step 11: Display the result on GUI.

2.2. Similarity Measure

Eight similarity measurements have been proposed. In this work, we use Sum-of-Absolute Differences (SAD). Several researchers have reported that this method is the most effective and efficient. Equation given below defines SAD respectively [2].

$$SAD(f_q, f_t) = \sum_{i=0}^{n-1}(1f_q[i] - f_t[i]) \qquad \dots (4)$$

Where *fq, ft* represent query feature vector and database feature vectors and n is the number of features in each vector.

3.Experiment Results & Analysis

A general purpose image database consists of 1000 images which used for experimentation. The database consists of different categories such as Africans and villages, Beaches, Buildings, Buses, Dinosaurs, Animals, Flowers, Horses, Food and Natural scenes. All the categories are used for retrieval. These images are stored in JPEG format with size 256x256 and each image is represented with RGB color space. In order to measure retrieval effectiveness for an image retrieval system, precision and recall values are used, the ratio of relevant retrieved images to the total number of retrieved images (precision) and the ratio of retrieved relevant images to the total number of relevant images in the database (Recall). Table1 summarizes the experiment results compared with the existing local histogram method.

Precision = (the no. of retrieved images that are relevant) / (The no. of retrieved images)
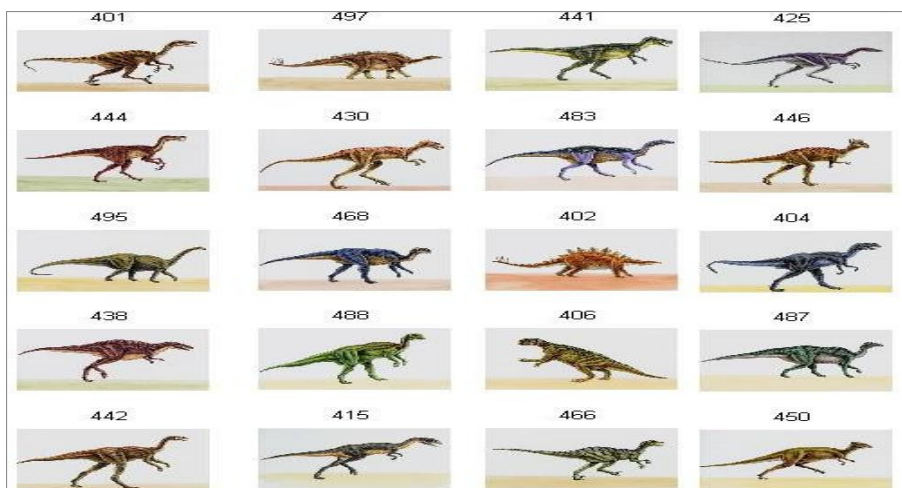Recall = (the no. of retrieved images that are relevant) / (The total no. of relevant images).

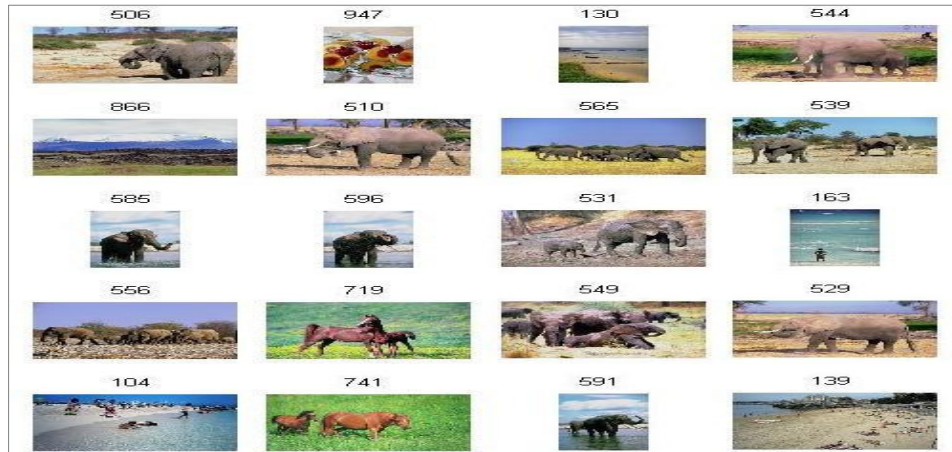Fig 3: Image Retrieval Results for Query image 401.jpg

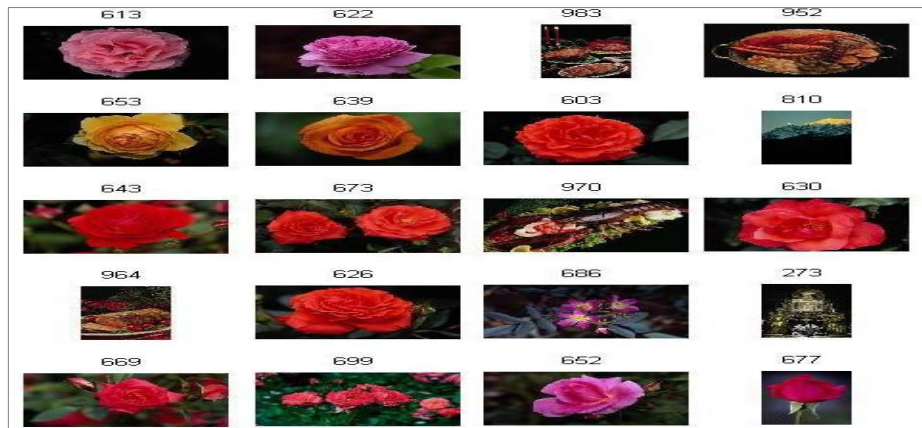Fig 4: Image Retrieval Results for Query image 506.jpg

Fig 5: Image Retrieval Results for Query image 613.jpg

Table 1: Precision and Recall Results of Local Histogram and Proposed Method

| Category of Images	Local Histogram Method		Mean Segment Method	
	precision	Recall	Precision	Recall
Africans &Villages	0.2505	0.0509	0.133	0.0266
Beaches	0.146	0.0292	0.2615	0.0523
Buildings	0.1525	0.0305	0.1105	0.0221
Busses	0.1785	0.0357	0.1725	0.0345
Dinosaurs	0.8625	0.1725	0.9615	0.1923
Elephants	0.223	0.0446	0.349	0.0698
Flowers	0.187	0.0374	0.495	0.099
Horses	0.1715	0.0343	0.208	0.0416
Nature	0.161	0.0322	0.259	0.0518
Food	0.2135	0.0427	0.156	0.0312
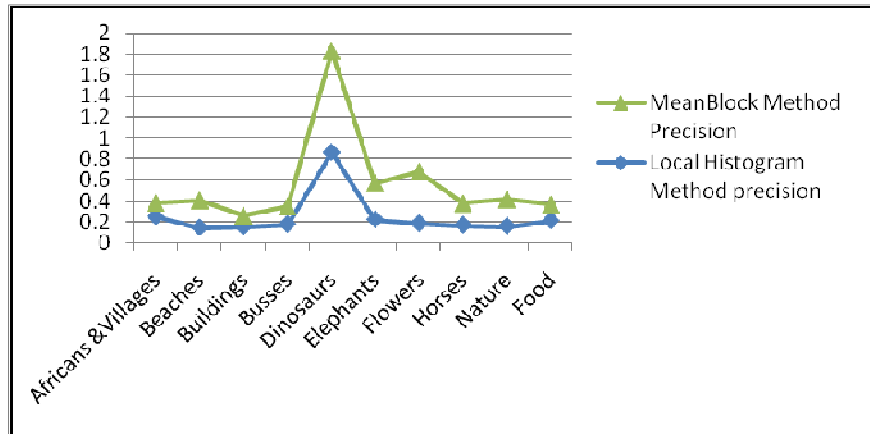
Fig 6: Comparison chart Local Histogram & Proposed Method Precision values

CONCLUSION

In this paper, a CBIR method has been proposed which uses the color moments and the proposed method gives higher average precision and recall. An experimental comparison of a number of different color descriptors for content-based image retrieval was performed. Color histograms and color moments are considered for retrieval. The retrieval efficiency of the color descriptors was investigated by means of recall and precision. In most of the images categories color moment's shows better performance than the local histogram method.

REFERENCES

[1] Texture Based Image Indexing and Retrieval, Dr. V Vijaya Kumar N Gnaneswara Rao, V Venkata Krishna IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.5, May 2009.

[2] T. Huang and Y. Rui. Image retrieval: Past, present, and future. In Proc. of the Intl. Symposium on Multimedia Information Processing, pages 1–23, 1997.

[3] Chang S.K., and Hsu A. (1992), "Image information systems: where do we go from here?" IEEE Trans. On Knowledge and Data Engineering, Vol. 5, No. 5, pp. 431-442.

[4] Tamura, H. and Yokoya N. (1984), "Image database systems: A survey," Pattern Recognition, Vol. 17, No. 1, pp. 29-43.

[5] Zachary J. M., Jr. and Sitharama S. I. Content Based Image Retrieval Systems, Journal of the American Society for Information Science and Technology, 2001.

[6] Goodrum A. A. (2000), Image Information Retrieval: An overview of Current Research.

[7] Smulders A.M.W., Worring M., Santini, A. S., Gupta, and R. Jain (2000), "Content-based image retrieval at the end of the early years," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.22, No. 12, pp. 1349-1380.

[8] Faloutsos C. et. al. (1994), "Efficient and effective querying by image content, "Journal of intelligent information systems, Vol.3, pp.231-262.

[9] Fuhui L., Hongjiang Z. and David D. F. (2003), Multimedia Information retrieval and Management.

[10] Rui Y. and Huang T. S (1997), Image retrieval: Past, Present, and Future, Journal of Visual Communication and Image Representation.

[11] Huang J., Kumar S. R., Mitra M., Wei-Jing Z., and Zabih R. (1997), Image indexing using color correlograms. In IEEE Conference on Computer Vision and Pattern Recognition, pages 762—768.

[12] Hu M.K. (1962), Visual pattern recognition by moment invariants, computer methods in image analysis. IRE Transactions on Information Theory, Vol. 8.