International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 24 (2017) pp. 15966-15971

An Efficient Content Based Image Retrieval Using Block Color Histogram and Color Co-occurrence Matrix

Mrs. J. Vanitha[1] and Dr. M. SenthilMurugan [2]

[1]Research Scholar, Research & Development Centre, Bharathiar University, Coimbatore 641046, Tamil Nadu, India.

[1]Associate Professor, Department of MCA, EGS Pillay Engineering College, Nagapattinam, Tamilnadu, India.

[2]Director, AVC College of Engineering, Mayiladuthurai, Tamilnadu, India.

[1]Orcid: 0000-0002-9246-6726

Abstract

In the present days, Images are widely used everywhere and the image retrieval plays a vital role. Content Based Image Retrieval (CBIR) system is a well-known technique for effective image retrieval. The Proposed CBIR system use color and spatial feature to retrieve similar images. Color histogram is the widely used method to extract color features which liberate translation, rotation and scaling of image and avoid spatial feature. The color co-occurrence matrix extracts spatial feature. The Proposed CBIR system uses 3*3 block color histogram with 1:2:1 proportion of segmentation with weight and color co-occurrence matrix to extract color and spatial feature. The images in the database are indexed with feature vectors by using SR-tree algorithm which is used to increase the speed of retrieval. The feature matching process is carried out using Euclidean distance of color histogram and color co-occurrence matrix and find distance metric. Then the images in the database are sorted using distance metric and required number of images will be retrieved.

Keywords: Block Color histogram, Content Based Image Retrieval, Color co-occurrence Matrix, Euclidean distance, HSV Color space

INTRODUCTION

Images are widely used nowadays. Image retrieval employs vital role in Military affairs, education, medical science, agriculture etc. Image retrieval can be classified as context based image retrieval and content based image retrieval. Searching of images using keywords and text instead of image content which is called context based image retrieval, will not give better result. The image contents are color, texture, shape and spatial information. Color Histogram and Color Co-occurrence Matrix are used to extract color and spatial feature from query image and images in database [29].

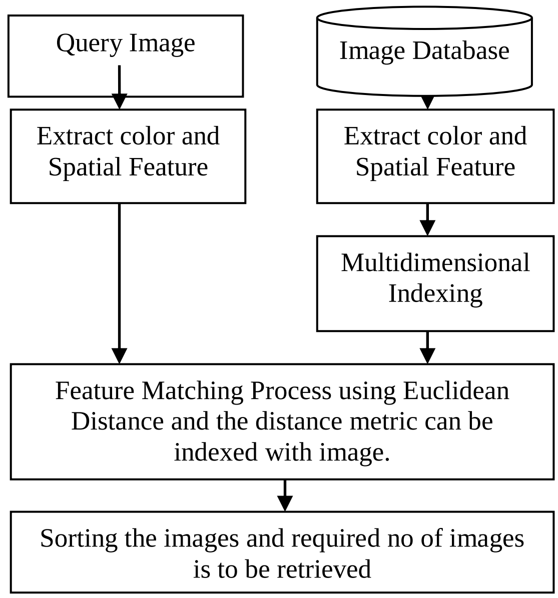

Figure 1: Proposed CBIR Architecture

The multidimensional indexing is used after the process of extracting color and spatial feature and stored the values of Hue, Saturation, Value, Color Histogram and Color co-occurrence matrix with the images for increasing retrieval speed.

The Feature Matching process uses Euclidean distance to sort the images in the database after which fixed numbers of images will be retrieved.

This proposed content based image retrieval system (Figure 1) retrieves more similar images.

The rest of the paper is organized as follows. Review of the literature is summarized in Section II. In Section III, the methods of feature extraction are summarized. In Section IV, Multidimensional indexing is explained .The proposed feature matching process is detailed in Section V. Image retrieval process is explained in Section VI. In Section VII, Application of CBIR is provided. Conclusions are summarized in section VIII. Then the list of references is provided finally.

REVIEW OF LITERATURE

Content Based Image Retrieval (CBIR) System is accessing image in an effective way [1]. The traditional way of an image retrieved using annotated text, lacks the effective description of the image. The main idea of CBIR is to analyze image information by low level features of an image [2], which include color, texture, shape and spatial information of image. CBIR is employed in many areas including military affairs, medical science, education, architectural design etc. CBIR system includes QBIC [3], Photobook [4], VisualSEEk [5], Virage [6], Netra [7] and SIMPLIcity [8] etc. Histogram is the most commonly used technique to describe features of image [13]. The Shape [10], texture [11, 26] and spatial features [12] etc. of the image are taken in account to improve the CBIR. Because of its simplicity and robustness, color is the most effective feature. Color histograms are used to extract color feature [13] for implementation. HSV color space is used to represent color for better human visual perception [14, 15]. Color co-occurrence matrix is formulated only in the HSV color space [9, 16, 27]. Images can be retrieved quickly and accurately by using fused low-level features [13, 28]. In the Feature matching process, the query image is divided into two blocks, and each image in the database is also divided into two blocks and histograms are also calculated separately, comparison was made separately and finally by considering all local histograms a sorted order of best suitable images were generated, final search results were displayed from the sorted order [17, 19]. For Indexing the images in the database, SS-tree indexing is superior for similarity indexing applications Compared with R*-tree indexing [18]. The SR-Tree [24] enhances the disjoint among regions which improves the performance on nearest neighbor queries. The SR-Tree is the best multi dimensional indexing structure among the SS-tree[18,23], the R*-Tree[21,22] and the K-D-B –Tree[20].

Feature extraction

The color and spatial features of image are extracted using color histogram and color co-occurrence matrix respectively.

Color Histogram

Color histogram defined as a set of bins where each bin denotes the probability of pixels in the image being of a particular color**.** There are two types of color histograms, Global color histograms and Local color histogram. A Global Color Histogram represents one whole image (Figure 2) with a single color histogram.

A Local Color Histogram divides an image into fixed blocks and takes the color histogram of each block. The number of block start with two. By observation, 3*3 block color histogram gives good result and it is better than global color histogram from the human visual perception [13]. The 3*3 block color histogram can be classified into two. These are knows as,

Type 1: Equally divided and same weight given for each division

Type 2: Unequally divided and double the weight given for center division (Figure 3).

Figure 2: Original Image

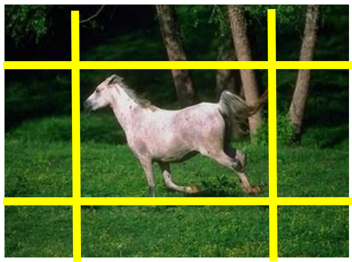

Figure 3: Segmented Image with 1:2:1 Proportion

1	2	1
2	4	2
1	2	1

Figure 4: Weight for each block

In the proposed CBIR System, Type 2 is used with weight (Figure 4). The images in the database and query image are segmented into 3*3 block with 1:2:1 proportion[31]. Color histogram for each block is stored in an array for future use.

For i = 1 to number of images in the database

  For j = 1 to 9

  IHis[i ,j]= ColorHistogram(j)

  Next j

Next i

Where i denotes number of image in database and j denotes number of blocks in an image and IHis[i,j] stores color histogram of j^{th} block of i^{th} image in the database.

Color Co-occurrence Matrix

Color co-occurrence matrix (CCM) includes three dimensional matrixes where the first dimension and second dimension contain colors of any pair of pixel p and Np and the spatial distance between them along the third dimension. CCM is simplified to represent the number of color (hue) pairs between adjacent pixels in the image. For each pixel in the image, 4-neighbors (horizontal and vertical neighbors) are accounted [16]. The CCM is used to extract spatial feature in this proposed CBIR.

It generated in HSV color space. Hue represents the dominant wavelength in light. It is the term for the pure spectrum colors. Hue is expressed from 0° to 360°. It represent hue of red (starts at 0°), yellow(starts at 60°), green(starts at 120°), cyan(starts at 180°), blue(starts at 240°) and magenta(starts at 300°). All hues can be mixed from three basic hues known as primaries.

Saturation represents the dominance of hue in color. It is intensity of color. Highly saturated color is vivid, low saturated color is muted. There is no saturation in image is grey image.

Value describes the brightness or intensity of the color. It is relative lightness or darkness of color.

The RGB color space is converted into HSV color space. It is quantized in to 16*16. The hue value (hl) is 0 to 15. Each component of CCM can be written as

$$CCM\ (i, j)=P_r(hl_p, hl_{Np}, sp(p,Np)) \qquad (1)$$

	Np 90°	
Np 180°	P Pixel	Np 0°
	Np 270°	

Figure 5: Angles for a pixel

The four matrices representing the neighbor at angle 0°, 90°, 180°, 270°(non diagonal points).

Multi dimensional indexing

The SR-Tree identifies a region by the intersection of bounding rectangle and a bounding sphere . It improves the performance on nearest neighbor queries[32]. The organization of the SR-Tree has its base on R-Tree, R*-Tree and of the SS-Tree. Also the traversal algorithm is same as R* Tree and SS Tree. The SR-Tree differs from R* Tree and SS Tree in the approach of calculating the distance from the search point to every region. The performance test shows it is more efficient for fewer uniform data set. The proposed CBIR has the process of the feature vectors are stored with images using the SR–Tree algorithm for increasing similarity searching efficiency.

Feature matching process

The feature matching process is used to match the features between query image and images in the database. The multidimensional indexing is used to retrieve feature vectors quickly for feature matching process.The proposed CBIR is using Euclidean distance for feature matching process[30]. The Euclidean distance value is small which means these are similar images.

The following algorithms are used to find the Euclidean distance value between query image and images in the database.

Algorithm for measuring similarity between query image and the images in database using Euclidean distance of weighted 3*3 block color histogram (EDwbch(I_j))

Step 1: Input Query Image (Q) and Image Database(I_j Where j ranges from 1 to no of images in the database).

Step 2: Read the Query Image (Q) .

Step 3: The image can be segmented into 3*3 block according to the Proportion 1:2:1(Q_i where i ranges from 1 to 9).

Step 4: Calculate the histogram for each block of the Query Image separately.

Step 5: Read the image (I_j) from the given database.

Step 6: Do the step 3 – 5 for all the images in the database and stored as an array for future use.

Step 7: Compare the 1st block color histogram of image in database with 1st block color histogram of query image using Euclidean distance $d_i = \sqrt{\sum_{i=1}^{N}(Ii - Qi)}^2$. The distance is calculated as EDwbch(Ij) = $\sum_{i=1}^{9} Wid_i$ where Wi is the weight of the i^{th} block and d_i is the Euclidean distance of the i^{th} block.

Step 8: The distance marked as EDwbch(Ij) where EDwbch(I_j) is weighted Euclidean distance of the image in the database(I_j).

Step 9: Continue the step 7 and 8 until all the images are to be read from the image database.

By using this algorithm, the color histogram of nine blocks in an image compare with color histogram of nine blocks in query image by using Euclidean distance. The resultant value is known as EDwbch (I_j). This process is followed all the images in database.

Algorithm for similarity measure using Euclidean distance of Color Co-occurrence Matrix of HSV of a pixel (EDccm(I_j)).

Step 1: Load the Image Database(I_j Where j ranges from 1 to number of images in the database).

Step 2: Quantization means reducing the number of color bins in a histogram by taking similar colors put into same bin. Quantize the images in image database for [16, 16].

Step 3: Convert the images from RGB to HSV Color Space.

$$H = \left\{ \frac{COS^{-1}\ 1/1/2\ \left[(R-G)\ +\ (R-B) \right]}{\sqrt{(R-G)^2\ +\ (R-B)\ (G-B)}} \right\} \tag{2}$$

$$S = \frac{1\ -\ 3\left[min(R,\ G,\ B) \right]}{R+G+B} \tag{3}$$

$$V = \frac{1\ (R+G+B)}{3} \tag{4}$$

Step 4: Formulate CCM for all images in the database.

Step 5: Load the Query image(Q) and apply the procedure 2-4 to find CCM values of Query image.

Step 6: Determine the Euclidean distance of Query image with image in the database.

Step 7: The distance of Color co-occurrence matrix is marked as EDccm(Ij) of the image in the database.

By using this algorithm, we determine Euclidean distance between query image and images in the database (EDwbch(Ij)) for all images using color co-occurrence matrix of HSV of a pixel. The images in the database are sorted using the addition of distance values (EDwbch(Ij) + EDccm(Ij)) in the database.

Image retrieval

The images are sorted in the image database in ascending order based on sum of Euclidean distance value of color histogram and color co-occurrence matrix. The Euclidean distance is small which means the image is similar image. The fixed number of images is to be retrieved from the beginning of the database.

Implemention of cbir

The experimental result (Table 1) shows that the proposed CBIR can be used in many image retrieval problems. This experiment is implemented in MATLAB. The images which are retrieved have feature vectors closest to the query image(Figure 8). The image retrieval performance of the system can be measured in terms of Precision (Figure 6) and Recall (Figure 7).

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ no\ of\ images\ to\ be\ retrieved}$$

$$Recall = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ no\ of\ relevant\ images\ in\ database}$$

The work implementation is,

1. Image Database: The Experiment is used Wang Image dataset.

2. Feature Vector: Compute color histogram for each block in the 3*3 block with 1:2:1 segmented image and color co-occurrence matrix and treated as feature vectors.

3. Indexing: Using SR-Tree algorithm for multidimensional indexing the feature vectors for future retrieval quickly.

4. Matching: To calculate distance between query image and database images using Euclidean distance.

5. Retrieval Efficiency: Precision and Recall were calculated to evaluate the retrieval performance by using color images in Wang image database.

Table 1: Precision and Recall comparison between base and proposed system

Category	Base System Result		Proposed system Results	
	Precision	Recall	Precision	Recall
African	68	14	72	16
Beach	54	19	59	20
Monument	56	17	58	19
Buses	89	12	89	12
Dinosaurs	99	10	99	11
Elephants	66	15	70	16
Flowers	89	11	93	13
Horses	80	13	85	14
Mountains	52	21	56	24
Food	73	13	77	15

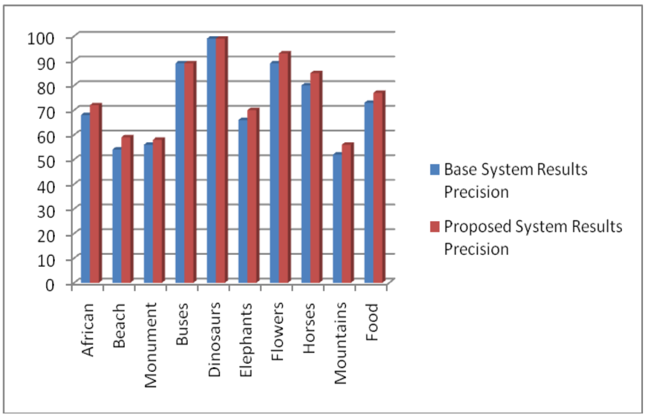

Figure 6: Precision Comparison between Base and Proposed System

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 24 (2017) pp. 15966-15971
© Research India Publications. http://www.ripublication.com

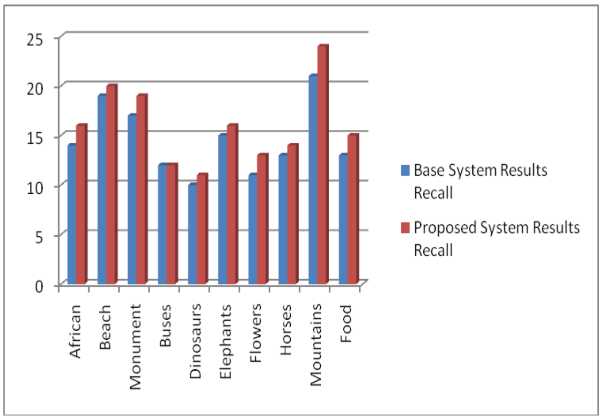

Figure 7: Recall Comparison between Base and Proposed System

The Base system used color and texture feature to retrieve images from the database. The Proposed system is using color and spatial feature.

Figure 8: Sample query (The image on the top left is the query) &. Arranged from left to right and top to bottom are the retrieved images.

CONCLUSION

The Proposed CBIR system is using color and spatial feature. The Feature extraction can be carried out by using 3*3 block color Histogram for color feature and Color Co-occurrence Matrix of HSV of a pixel for spatial feature. The feature matching process is using Euclidean distance after which images are sorted in ascending order and the fixed number of images is to be retrieved. In future, the image classification has to be adding to reduce the images in the search space which helps to increase the speed of retrieval and also to retrieve more similar images.

REFERENCES

[1] P. Muneesawang, L. Guan, An interactive approach for CBIR using a network of radial basis functions, IEEE Transactions on Multimedia 6 (2004) 703–716.

[2] R. Datta, D. Joshi, J. Li, J.Z. Wang, Image retrieval: ideas, influences, and trends of the new age, ACM Computing Surveys 40 (2) (2008) 1–60.

[3] W. Niblack, R. Barber, W. Equitz, et al., The QBIC project: querying images by content using color, texture, and shape, in: SPIE 1908, San Jose, CA, 1993,pp. 173–187.

[4] A. Pentland, R.W. Picard, S. Scarloff, Photobook: tools for content-based manipulation of image databases, in: SPIE 2185, San Jose, CA, 1994, pp. 34–47.

[5] S. Mehrotra, Y. Rui, M. Ortega, et al., Supporting content-based queries over images in MARS, in: Proc. of IEEE Int'l Conf. on Multimedia Computing and Systems'97, Ottawa, Ontario, Canada, 1997, pp. 632–633.

[6] J.R. Bach, C. Fuller, A. Gupta, et al., Virage image search engine: an open framework for image management, in: SPIE 2670, 23, San Jose, CA, 1996, pp. 76–87.

[7] J.R. Smith, Integrated spatial and feature image systems: retrieval, analysis and compression, Ph.D. Dissertation, Columbia University, New York City, 1997.

[8] J.Z. Wang, J. Li, G. Wiederhold, SIMPLIcity: semantics-sensitive integrated matching for picture libraries, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (9) (2001) 947–963.

[9] A.Vadivel,ShamikSural,A.K. Majumdar, An Integrated Color and Intensity Co-occurrence Matrix, Pattern Recognition Letters 28 (2007) 974–983.

[10] G. Gagaudakis, P.L. Rosin, Incorporating shape into histograms for CBIR, Pattern Recognition 35 (2002) 81–91.

[11] P.S. Hiremath, Jagadeesh Pujari, Content based image retrieval using color, texture and shape features, in: 15th International Conference on Advanced Computing and Communications, ADCOM 2007, 2007, pp. 780–784.

[12] Y.K. Chan, C.Y. Chen, Image retrieval system based on color-complexity and color-spatial features, Journal of Systems and Software 71 (2004) 65–70.

[13] Jun Yue, Zhenbo Li , Lu Liu , Zetian Fu, Content based image retrieval using color and texture fused features, Mathematical and Computer Modelling 54 (2011) 1121–1127.

[14] Nishant Shrivastava, Vipin Tyagi, An efficient technique for retrieval of color images in large Databases, Computers and Electrical Engineering 46 (2015) 314–327.

[15] Dibya Jyoti Bora, Anil Kumar Gupta, Fayaz Ahmad Khan, Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation, International Journal of Emerging Technology and Advanced Engineering Volume 5, Issue 2, February 2015.

[16] Seong-O Shim, Tae-Sun Choi, *Senior Member, IEEE*, **I**mage indexing by modified color co-occurrence matrix ,IEEE Xplorer.

[17] Chesti Altaff Hussain, Dr. D. Venkata Rao, T. Praveen, Color histogram based image retrieval, International Journal of Advanced Engineering Technology E-ISSN 0976-3945(2013)

[18] David A.White, Ramesh Jain , Similarity Indexing with SS-tree,IEEE,516-523(1996)

[19] Wenqing Huang, Jiazhe Dai and Qiang Wu , Image Retrieval Based on Weighted Block Color Histogram and Texton Co-occurrence Matrix, International Conference on Information Technology and Management Innovation,848-853(2015).

[20] J. T. Robinson, “The K-D-B-tree: a Search Structure for Large Multidimensional Dynamic Indexes,” Proc. ACM SIGMOD, Ann Arbor, US.4, pp 10-18, Apr.1981.

[21] A. Guttman, “R-Trees: a Dynamic Index Structure for Spatial Searching ,” Proc. ACM SIGMOD, Boston, USA, pp. 47-57, June. 1996.

[22] N. Beckmann, H. P. Kriegel, R. Schneider and B. Seeger,” The R*-tree: an Efficient and Robust Access Method for Points and Rectangles,” Proc. ACM SIGMOD. Atlantic City,US.4,pp.322-331, May 1990.

[23] D. A. White and R.Jain, “ Similarity Indexing: Algorithms and Performance,” Proc. SPIE vol. 2670, San Diego, USA , pp. 62-73,Jan 1996.

[24] Norio Katayama and Shin’ichi Satoh , The SR-Tree: An Index Structure for High-Dimensional Nearest Neighbor Queries Proc. ACM SIGMOD Int. Conf. on Management of Data, pp.13-15, May 1997.

[25] Mona Mahrous Mohammeda, Amr Badr, M.B. Abdelhalim, Image classification and retrieval using optimized Pulse-Coupled Neural Network, Expert Systems with Applications 42 (2015) 4927–4936.

[26] P.W. Huang, S.K. Dai, "Image retrieval by texture similarity," Pattern Recognition, vol. 36,2003, pp. 665 - 679.

[27] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," Image and Vision Computing, vol. 22, 2004, pp. 1211-1220.

[28] C.-H. Lin, R.-T. Chen, Y.-K. Chan, A smart content-based image retrieval system based on color and texture feature, Image and Vision Computing 27 (2009)658–665.

[29] J.Vanitha, M.Senthilmurugan ,A Survey On Content Based Image Retrieval Using Color Histogram And HSV Color Model , International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 10 No.51 ,2015 © Research India Publications

[30] J.Vanitha, M.Senthilmurugan, Content Based Image Retrieval System Using Image Classification, Advances in Natural and Applied Sciences Volume: 10, Number 8 , June 2016 issue Page No: 1-5 Place of Publication: Jordan ISSN: 1995-0772 EISSN: 1998-1090 , AENSI Publication

[31] J.Vanitha, M.Senthilmurugan, Feature Matching Process Using Euclidean Distance Of Weighted Block Color Histogram And Color Co-occurrence Matrix For Content Based Image Retrieval System , International Journal of Emerging Technology and Advanced Engineering (IJETAE), ISSN 2250-2459 ,Vol. 7 , Issue No.6 Page No: 143-147, June 2017

[32] J.Vanitha, M.Senthilmurugan **,**Multidimensional Image Indexing Using Sr-tree Algorithm For Content-based Image Retrieval System published in Emerging Research in Computing, Information, Communication and Applications. ERCICA 2016, EBOOK ISBN: 978-981-10-4741-1, DOI :10.1007/978-981-10-4741-1_8, Page No 81-88, Springer, Singapore