Statistical Moments Extracted from Eight Bins Formed by CG Partitioning of Histogram Modified using Linear Equations

[1]Dr. H. B. Kekre, [2]Kavita Sonawane

[1]Sr. Professor, Computer Engineering Department, MPSTME, NMIMS, Mumbai, India

[2]Ph.D Research Scholar, Computer Engineering Department, MPSTME, NMIMS, Mumbai, India

Abstract

This paper explores the novel idea for feature extraction based on bins approach for CBIR. This work has fulfilled all the criterion of efficient feature extraction method like dimensionality reduction, fast extraction and efficient retrieval. Main idea used in feature extraction is based on the histogram and the linear functions used to modify it. Each BMP image in database is separated in R, G and B planes. We have calculated the histogram for each of them which are modified using linear equations. These modified histograms are partitioned using CG so that mass of intensities of the image pixels will be distributed uniformly in two parts. This CG partitioning of three image planes leads to generate the eight bins. Information extracted from these bins is in the form of statistical first four absolute moments namely MEAN (MEAN), Standard deviation (STD), Skewness (SKEW), and Kurtosis (KURTO). Each of these moments are computed separately for R, G and B colors. This generates four types of feature vectors of dimension eight. Database of 2000 BMP images is used for the experimentation. Multiple feature vector databases are prepared as part of preprocessing work of this CBIR system. Comparison of query and database feature vector is carried out using three similarity measures namely Euclidean distance (ED), Absolute distance (AD) and Cosine correlation distance (CD). To evaluate the retrieval efficiency of this system we have used three parameters, Precision Recall Cross over Point (PRCP), Longest String (LS) and Length of String to Retrieve all Relevant (LSRR).

Keywords: *Bins, CBIR, Histogram modification, Statistical moments, MEAN, STD, SKEW, KURTO, ED, AD, CD, PRCP, LS, LSRR.*

1. Introduction

CBIR system explained in this paper is based on bins approach. The bins designed in this system are formed by partitioning of the histogram. The histogram is modified before partitioning. The modification is done using the linear equations proposed in this paper to shift the intensity levels to higher side. Focus of many CBIR system discussed in the literature is on feature extraction process. Image features are broadly classified into two types local and global image feature [1-4]. These features can further be classified in variant and invariant feature vectors. The image contents prominently visible to human eyes are color, texture and shape. [5-9].

We are using the color image feature which is invariant in nature; i.e it will not vary even after the image translation or rotation. We are not using color content directly from the image as it is. We are focusing on the color histograms of the image obtained for three image planes (R, G and B) separately [10]-[13]. Various techniques discussed in the literature are using the histograms based approaches for CBIR system such as Histogram Intersection (HI), Histogram Euclidean Distance (HED) and Histogram Quadratic Distance Measures (HQDM) Integrated Histogram Bin Matching (IHBM) etc [14-19].

Image histogram is the tool that counts the no of pixels having same intensity and keeps them in separate bin for each intensity level. In MATLAB we know that to handle 256 intensity levels of image, it represents the histogram of size 256 bins. Many approaches proposed in the literature are using the 256 bins of histogram for comparing database and query image as it is. It gives high dimensional feature vector of size 256 bins. It also increases the computational complexity and the execution time required to compare the images. In our work we have overcome this problem by reducing the feature vector size to just eight bins. This greatly reduces the execution time required to compare the images. The process of modifying and partitioning the histogram along with the formation of bins is explained in detail in the following sections.

Statistical moments extracted from eight bins are used as feature vectors to represent the image [20-23]. Based on this information we have different types of feature vectors as we are computing these moments for R, G and B colors separately. We can modify the histogram using three linear equations proposed in this paper to increase the intensity values. By considering all these variations in the methods used to extract the image features we have prepared total 36 (3 colors, 3 linear equations and 4 moments) feature vector databases for 2000 BMP images. In all CBIR systems it is necessary to compare the query and database image vectors so that resultant set of images similar to query can be generated. This is done by means of similarity measure, which gives the degree of similarity between two objects. We have used three similarity measures to facilitate the comparison process [21], [24]. Measures used this CBIR are, Euclidean distance (ED), Absolute distance (AD) and Cosine correlation distance (CD). Each of the 36 feature vector databases are tested using the same set of 200 query images selected randomly from image database. Retrieval results are obtained with respect to each similarity measure used. All results are then evaluated and compared using the three performance evaluation parameters namely Precision Recall Cross over Point (PRCP), Longest String (LS) and Length of String to Retrieve all Relevant (LSRR)[25][26]. Presentation of the paper is as follows:

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 2, September 2012
ISSN (Online): 1694-0814
www.IJCSI.org

Section II describes the Linear equations used to modify the histogram. Section III elaborates on the bins formation and feature extraction process. Section IV explains the application of similarity measures and performance evaluation parameters. Section V gives the Experimental results and discussion along with the performance analysis.

2. Histogram Modification using Linear Equation

2.1 Linear Equation

Linear equation is used as histogram specification given in Equation1. It is used in two parts, p1 to p2 and p2 to p3 as shown in Figure1.

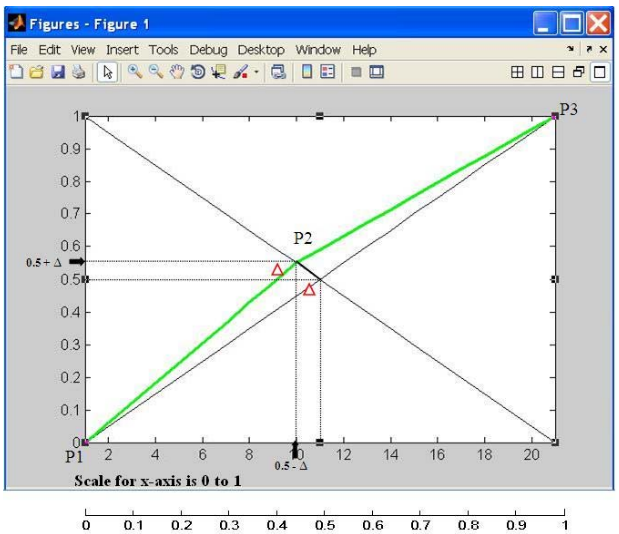


Figure 1. Histogram Specification Linear Equation for P1P2 and P2 P3

Let , linear Equation for first part is

$$y = mx + c \qquad (1)$$

Where $m = \dfrac{0.5 - \Delta}{0.5 + \Delta}$ and $c = 0$

Where Δ is the variable as shown in Figure1 use to give increment to the original intensity. Similarly for second part

Let,

$$y = mx + c \qquad (2)$$

Where $m = \left(\dfrac{1-(0.5-\Delta)}{1-(0.5+\Delta)}\right)$ and $C = \dfrac{-2\Delta}{0.5-\Delta}$

Using three different Δ, we get three different histogram specifications obtained using above equations and are shown in Figure 2. It shows the three specifications in red, green and blue colors obtained to modify the histogram using equations 1 and 2 along with three different Δ i.e. 0.05, 0.1 and 0.15 respectively.

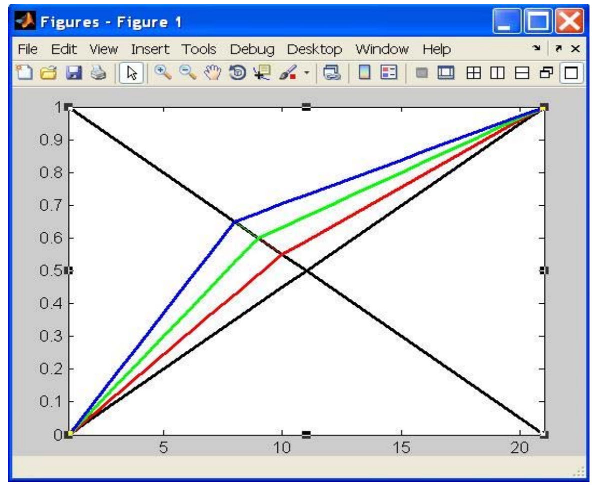
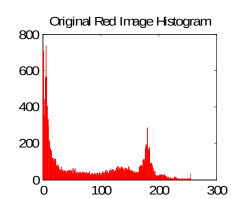


Figure 2. Histogram Specification with respect to three different Δ replaced in Eq.1 and 2.

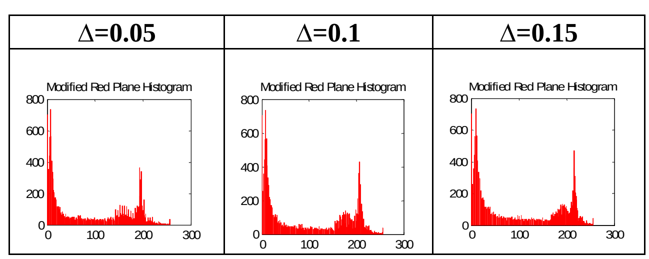
2.2 Application of Linear Equation over Histogram

The curves shown in Figure 2 are clearly showing that how the intensities will change. Effect after applying these functions over the histogram of R plane is shown in following Figure 3, through a, b and c.

Figure 3. a. Bus Image with R, G and B planes



b. Histogram of Red Plane of Bus Image



c. Red Plane Histogram Modifed using Eq 1 and 2 with respective Δ.

Above Figure 3 a. shows Bus image with its R, G and B planes. Figure3. b. shows the original histogram of red plane.

Figure 3. c. shows histogram modified using equations 1 and 2 with three different Δ values. We see in Figure 3 c. that the intensities are pushed high by some amount Δ.

3. CG Partitioning and Bins Formation

3.1 CG Partitioning

Processes explain in section II part 2 is applied to all three planes R, G and B. Each modified histogram is then partitioned into two parts using CG given in equation 3.

$$CG = \left[\frac{(L_1 W_1 + L_2 W_2 + ... + L_n W_n)}{\sum_{i=1}^{n} W_i} \right] \quad (3)$$

Where L_i is intensity Level and W_i is no of pixels at L_i

After computing the CG we get the histogram partitions such that the moment of each part around CG is of same magnitude.

Modified histogram partitioned using CG with part ids 0 and 1 as shown in Figure 4.

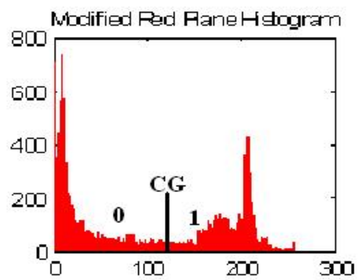


Figure 4. CG Based Partitioning of modified Histogram

3.2 Bins Formation: 000 to 111

Each image for which the feature vector will be extracted will be ready with the modified histogram with two parts 0 and 1. Three planes with two partitions generate $\mathbf{2^3 = 8}$ **c**ombinations; from 000 to 111 these are the addresses of 8 bins. These bins initially are holding the count of pixels falling in particular partition of R, G, B histograms for their R, G and B intensities respectively. e.g. Pixel's R, G and B values are falling in Part 1, part 0 and part 1of R, G and B histogram respectively then that pixel will be counted in bin no.'101' i.e. bin5. This process is applied to all pixels of the image one by one. All pixel count is distributed into these 8 bins according to the intensities they have. Feature Extraction and Feature Vector Databases

4. Feature Extraction In terms of Statistical Parameters

Each image under feature extraction will follow the process explain in section 2 and 3. We have seen that the image is ready with 8 bins having count of pixels. These bins are further directed to hold the detail information of these pixels counted in each bin. We are considering the R, G and B intensities of each pixel and computing the statistical moments for them. We are calculating the first four absolute centralize moments of the R, G B color intensities for the pixels counted into each bin for each image. These 8 bins are representing that image and called feature vectors of that image. Each statistical moment is representing one type of feature vector. This way we have obtained four types of feature vectors for each image. The statistical moments calculated over the image information available in each bin using the following equations 4, 5, 6 and 7.

Mean
$$\overline{R} = \frac{1}{N} \sum_{i=1}^{N} R_i \quad (4)$$

Standard deviation
$$R_{SD} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} (R - \overline{R})^2} \quad (5)$$

Skewness
$$R_{SK} = \frac{1}{N} \sqrt[3]{\sum_{i=1}^{N} \left| R - \overline{R} \right|^3} \quad (6)$$

Kurtosis
$$R_{KU} = \frac{1}{N} \sqrt[4]{\sum_{I=1}^{N} \left| R - \overline{R} \right|^4} \quad (7)$$

First two moments provides some information on the appearance of the distribution of the image data i.e. no of pixels along with the specific intensity each has. Next two moments are also providing the information on the shape of the distribution.

Calculation of moment is done separately for each color intensities associated with pixels counted into each bin. This gives multiple feature vectors for each image and in turn generates multiple feature vector databases.

4.1 Feature Vector Databases:

Moments are calculated for the pixels counted into each bin for R, G, and B colors separately. We have maintained the separate feature vector database for each type of feature vector based on the type of moment calculated for three color intensities and we have obtained total 3; (colors) x 4(moment) x 3 (linear equations used to modify the histogram) = 36 feature vector databases as shown in Figure 5.

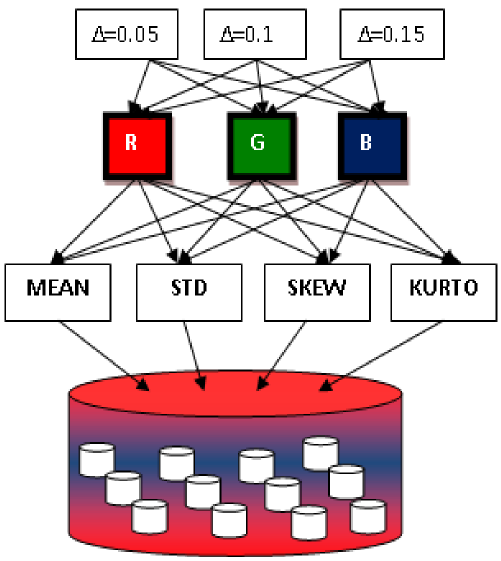


Figure 5. Formation of (36) Multiple feature vector Databases

5. Feature vector Comparison and Performance Evaluation

Once the feature vector databases are ready, system goes through two steps. First is to face the query image and second, generate the retrieval result. Between these two steps it performs the comparison of query and database image feature vectors by means of three similarity measures explained in part A. Similarity measure calculates the distance between the query and all database images. These distances are then sorted in ascending order. Database images which are at minimum distance from query image are selected for retrieval result. To analyze the behavior, response time and overall performance of the proposed CBIR system we have used three performance evaluation parameters namely PRCP, Longest String and LSRR. Part B elaborates these three parameters with equations.

5.1 Similarity Measures: ED, AD and CD

In our system the query and database image feature vectors are compared using three similarity measures Euclidean, Absolute and Cosine correlation distance given in equations 8, 9 and 10.

Euclidean Distance: $D_{QI} = \sqrt{\sum_{i=1}^{n} \left|(FQ_i - FI_i)\right|^2}$	(8)
Absolute Distance: $D_{QI} = \sum_{1}^{n} \left|(FQ_I - FI_I)\right|$	(9)
Cosine Correlation Distance: $\frac{\langle D(n) \mid Q(n)\rangle}{\sqrt{[\langle D(n) \mid D(n)\rangle \langle Q(n) \mid Q(n)\rangle]}}$ Where D(n) and Q(n) are Database and Query feature Vectors resp.	(10)

5.2 Performance Evaluation: PRCP, Longest String, LSRR

5.2.1 PRCP

Precision Recall Cross over Point is based on the theory of two traditional parameters precision and recall and are given equation 11 and 12.

$$\text{Precision} = \frac{\text{Relevant Retreived Images}}{\text{All Retreived Images}} \tag{11}$$

$$Recall = \frac{Relevant\ Retreived\ Images}{All\ Relevant\ In\ Database} \tag{12}$$

All CBIR system users expect that precision and recall both should be as high as possible. PRCP is the cross over point where both will have same value. PRCP, 1 indicates the ideal performance of the system and 0 indicates the worst case performance of the system. This parameters gives indication of how long we are from the ideal system [25][26].

5.2.2 Longest String

This parameter measures the no of relevant images retrieved continuously from the database images sorted in ascending order according to the distances sorted from minimum to maximum.

5.2.3 LSRR:

It measures the length of string to retrieve all relevant images from database or to make recall 1. This gives the indication of the system's strength to give maximum retrieval at faster rate. This LSRR measured in percentage shows the traversal done by the system through the sorted distances to collect all images relevant to query.

6. Results and Discussion

Proposed system explained in this paper is experimented with following query and database details.

6.1 Image Database and Query Image

Database: Database used for the experimentation of this work consist of 2000 BMP images containing 20 classes, where each class contains 100 images. The image classes used are Flower, Sunset, Mountain, Building, Bus, Dinosaur, Elephant, Barbie, Mickey, Horse, Kingfisher, Dove, Crow, Rainbow rose, Pyramid, Plate, Car, Trees, Ship and waterfall. Sample image from each class of database is shown in Figure 6.

Query Image: All approaches designed and implemented in this work are tested using the query image given as example. Set of 200 query images is used for testing and evaluating the performance of the system. It includes 10 sample query images selected randomly from each of the 20 classes available in database.

Figure 6. 20 Sample Images from database of 2000 BMP images

6.2 Results with detail Analysis

System is executed for all approaches based on the linear equations with three variations of Δ, four moments extracted from 8bins of R, G and B intensities. All the feature vector databases are fired with same set of 200 query images. Results obtained for each of three similarity measures ED, AD and CD are evaluated using the three evaluation parameters mentioned in section V.

6.2.1 PRCP for : Moments Mean, Standard deviation, Skewness and Kurtosis

Table 1: PRCP obtained for Mean of R, G and B with ED, AD and CD

PRCP : MEAN												
SM	***R***				***G***				***B***			
	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3
ED	5573	5792	5720	5346	5448	5833	5675	5491	5309	5997	5874	5632
AD	5749	5987	6036	5767	5480	6002	5935	5728	5421	6014	6050	5944
CD	5722	5874	5759	5353	5802	6392	6161	5826	5236	6352	5211	5891

Table 2: PRCP obtained for Standard Deviation of R, G and B with ED, AD and CD

PRCP : STD												
SM	***R***				***G***				***B***			
	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3
ED	6021	6196	6200	6073	6263	6294	6149	5784	5634	5544	5644	5552
AD	6285	6440	6407	6313	6439	6460	6436	6102	5880	6014	6127	5984
CD	5626	5696	5685	5634	6140	6171	5961	5620	5480	5634	5618	5419

Table 3: PRCP obtained for Skewness of R, G and B with ED, AD and CD

PRCP : SKEWNESS												
SM	***R***				***G***				***B***			
	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3
ED	4564	4720	4931	5047	4969	5051	4930	4814	4775	4980	4977	4808
AD	4957	5058	5172	5177	5230	5342	5168	5088	5066	5393	5386	5175
CD	4408	4466	4544	4581	4825	4791	4736	4687	4627	4900	4806	4819

Table 4: PRCP obtained for Kurtosis of R, G and B with ED, AD and CD

PRCP : KURTOSIS												
SM	***R***				***G***				***B***			
	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3	ORG	MOD1	MOD2	MOD3
ED	6074	6239	6237	6145	6703	6766	6537	6094	6032	6105	6002	5773
AD	6305	6440	6411	6335	6847	6429	6723	6337	6190	6372	6313	6139
CD	5810	5863	5879	5756	6340	6393	6104	5801	5739	5903	5869	5693

Table 1, 2, 3 and 4 are showing the PRCP results obtained for first four moments. Each row in the tables is indicating the results obtained with respect to each similarity measure separately. Each column in all the tables we can observe that, are showing results obtained for original as well as histograms modified with three different Δ (MOD1, 2 and 3), replaced in linear equations 1 and 2 for R, G and B colors separately. After observing these results we can clearly state that modified histograms are far better than original histogram in all four tables. For Mean MOD1 is better in 7 cases out of 9, for STD and SKEW it is better at 5 places out of 9, for KURTO it is better at 7 places out of 9 as compared to other modifications and the original histogram. Means we can say that MOD 1 is performing better in all cases for all moments. We observed the highest PRCP value obtained for MOD1 is for KURTO parameter for green color, i.e 6766 out of 20,000(Total retrieval of 200 query images, each out of 100). It means in these results precision and recall reached to 0.33 as an average of 200 query images. To improve these results further we thought of combining the R, G and B color results and refined and narrow down the result to single instead of separate three results. To do this, we have applied criterion OR over the R, G, and B results. i.e final retrieval set = Results of, (R' OR 'G' OR' B'). Results obtained after applying criterion OR for all the above results of all four moments with respect to each similarity measure are shown in Table 5. In Table 5 we can observe that the results are improved greatly. The highest PRCP value obtained now is 10872 for feature vector Kurtosis with AD measure for MOD1. Almost all results have crossed 10,000 after the OR criterion. The best value now, as an average result of 200 query images for precision and recall is reached to **0.55.**

Table 5: PRCP obtained For : R 'OR' G 'OR' B over Results of all four moments for All Modifications

	ED				***AD***				***CD***			
CRITERION 'OR'	**ORG**	**MOD1**	**MOD2**	**MOD3**	**ORG**	**MOD1**	**MOD2**	**MOD3**	**ORG**	**MOD1**	**MOD2**	**MOD3**
MEAN	8911	9003	8815	8398	9141	9302	9282	9098	8978	9617	9307	8784
STD	10228	10365	10365	10114	10397	10591	10602	10356	9988	10140	9973	9733
SKEW	9925	9434	9421	9448	9548	9734	9733	9650	8888	8981	9016	8949
KURTO	10649	10777	10580	10242	10696	**10872**	9650	10420	10230	10289	10134	9840

6.2.2 Longest String

Longest string parameter is meant for user to compare the performance of different CBIR systems. This string of continuous images may appear anywhere in the set of images sorted according to the distances sorted in ascending order. This parameter is used in our system to evaluate the performance of all approaches used with respect to four moments and three Δ values.

Charts 1, 2, 3 and 4 are showing the longest string results obtained for four moments MEAN, STD, SKEW and KURTO respectively. We can observe in these four charts Barbie class is giving best performance among all other classes for all other parameters. The best values for this class have reached to 95, 64, 53 and 54 for MEAN, STD, SKEW and KURTO respectively. These are obtained for similarity measure AD and all are for modified histograms. Means here also we can say that modified histograms are performing better than original one. One more observation is that Average Longest string i.e (average of 20 queries results as max. longest) for MEAN is 20 for almost all cases. Similarly for STD, SKEW and KURTO the average longest string has reached in between 15 to 20.

Along with these results one more analysis is done for the performance of each color (R, G and B). First, we have obtained the longest string for the 10 queries from each class and for each color R, G and B separately. Out of these 30 results, maximum longest is selected and plotted in the graph. While doing this we kept track that the max value selected is belongs to which of the three colors. At the end we have taken the count, that from which color we have got the maximum results. According to this we found red and green color is dominating over blue for all the cases.

6.2.3 LSRR

Measuring the length of string to retrieve all relevant images from database measures the efficiency and strength of the system. It determines the traversal time and length of sorted string of images traversed by the system to collect all relevant images. That is why the LSRR for ideal CBIR should be as low as possible. LSRR parameter measured for all approaches with respect to each moment, similarity measure and the histogram modifications are shown in charts 5, 6, 7 and 8. As explained in previous section, we are selecting the minimum traversal required out of the 30 results as LSRR and it belongs to which color feature vector is also noted. Considering the LSRR results of 20 queries i.e one from each class we have taken the average LSRR in percentage. All charts showing LSRR results plotted for average of 20 queries for all factors. Analysis done for the LSRR with respect to color factor we found red is performing better in all cases as compared to G and B color.

Chart 1. Maximum Longest String for MEAN

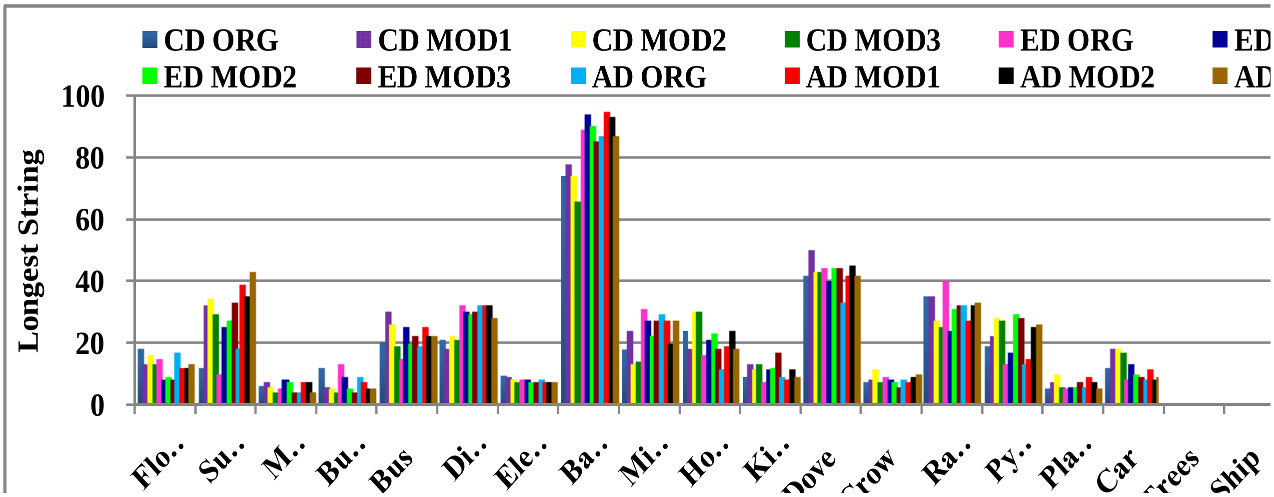
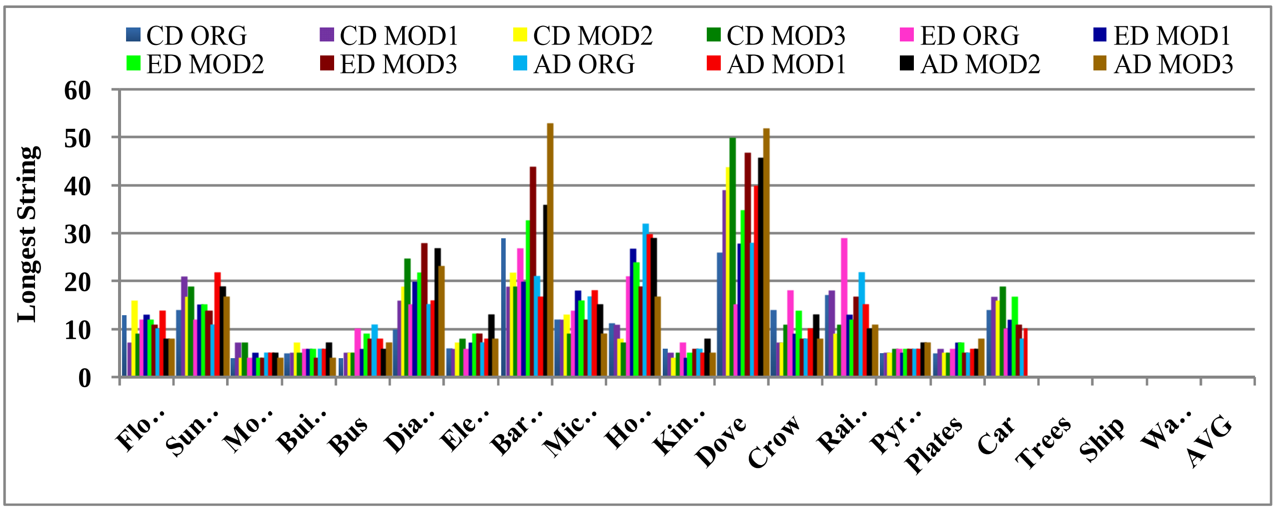


Chart 2 Maximum Longest String for Standard DEVIATION: STD



Char 3 Maximum Longest String for Skewness: SKEW

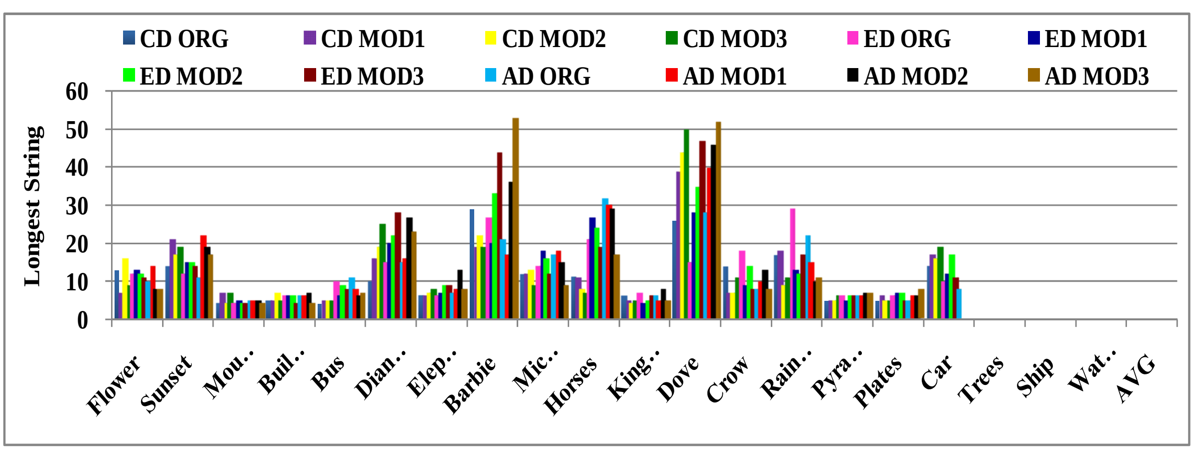
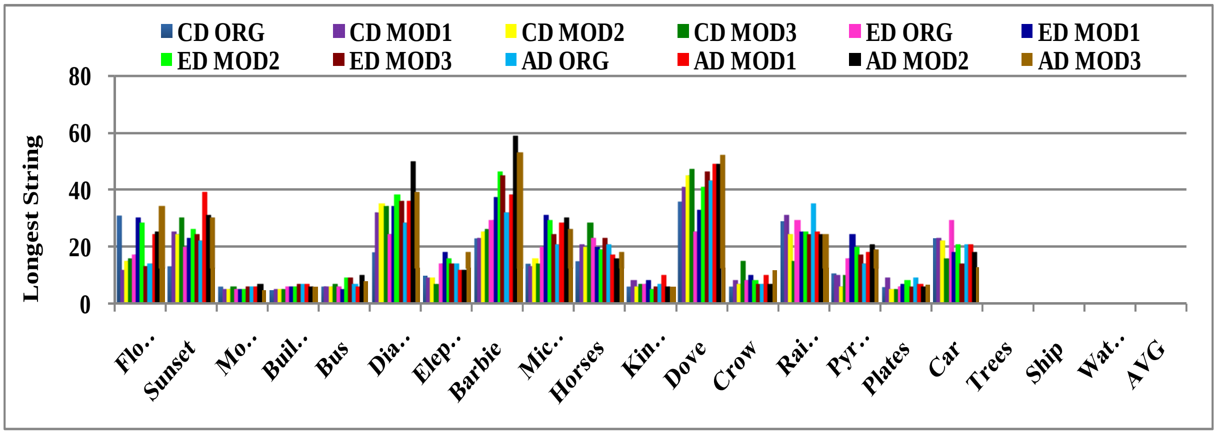
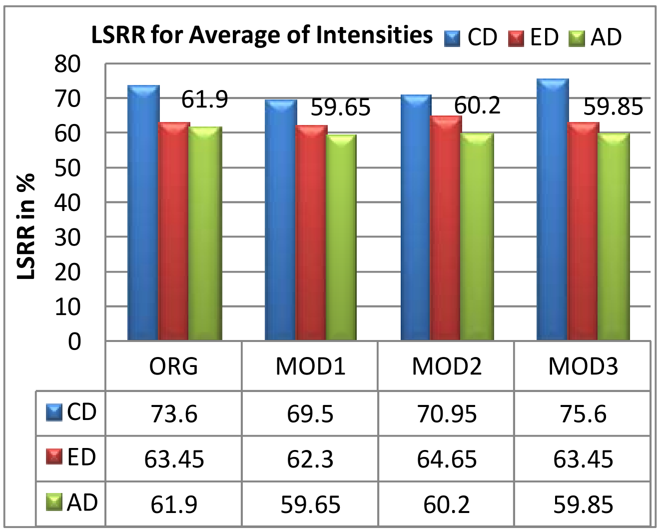


Chart 4. Maximum Longest String for Kurtosis: KURTO



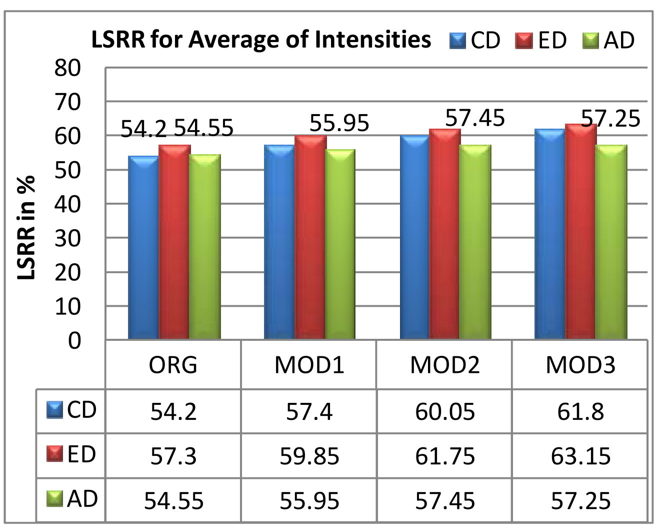
LSRR

Chart 5. Minimum LSRR for MEAN



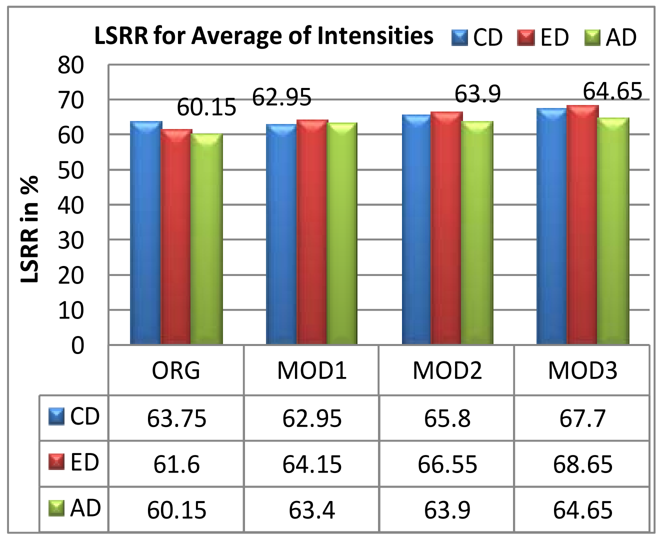
	ORG	MOD1	MOD2	MOD3
CD	73.6	69.5	70.95	75.6
ED	63.45	62.3	64.65	63.45
AD	61.9	59.65	60.2	59.85

Chart 6. Minimum LSRR for Standard Deviation: STD



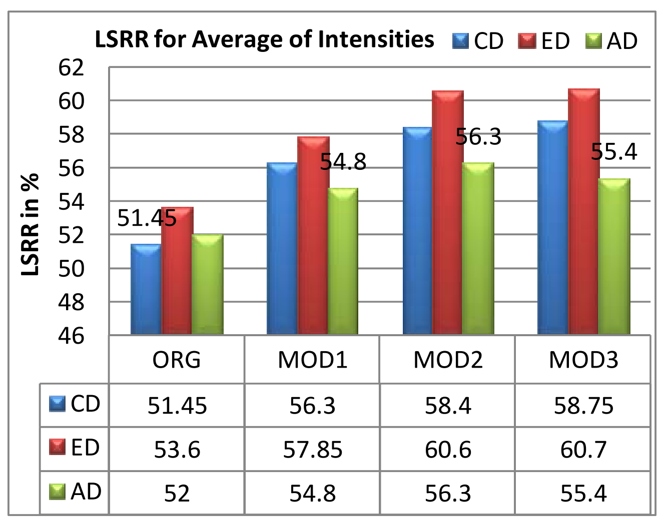
	ORG	MOD1	MOD2	MOD3
CD	54.2	57.4	60.05	61.8
ED	57.3	59.85	61.75	63.15
AD	54.55	55.95	57.45	57.25

Chart 7. Minimum LSRR for Skewness: SKEW



	ORG	MOD1	MOD2	MOD3
CD	63.75	62.95	65.8	67.7
ED	61.6	64.15	66.55	68.65
AD	60.15	63.4	63.9	64.65

Chart 8. Minimum LSRR for Kurtosis: KURTO



	ORG	MOD1	MOD2	MOD3
CD	51.45	56.3	58.4	58.75
ED	53.6	57.85	60.6	60.7
AD	52	54.8	56.3	55.4

We expect that % LSRR should be as low as possible to prove the system's strength and efficiency to make recall 1. We can observe in above Charts 5 to 8 that, % LSRR obtained for all of them are not exceeding 75% i.e traversal less than 75% itself generates recall 1. Among these results second, third and fourth moments are taking less than 65% in many cases. Performance given by modified histograms is better than original histogram results for MEAN and few cases in other moments. The best result obtained for average % LSRR is 51% for ORG histogram for Kurtosis with respect to CD measure.

7. CONCLUSION

Work presented in this paper is highlighting the use of linear equations with three different Δ. These are used to modify the histogram so that lower intensities will be pushed up and improved results can be obtained for the CBIR as compared to original histogram.

Feature extraction approaches worked out in this paper are based on bins formed by partitioning the histogram. It greatly reduces the size of the feature vector to just 8 bins. It saves the execution time and reduces the computational complexity to good extent as compared to all 256 histogram bins used as it is without partitioning.

Initially the bins are holding the count of pixels falling within particular range of intensities [11][13]. In this work we have computed the first four statistical central moments for the R, G and B intensities of each bin separately. These statistical moments have given better results as compared to previous work.[22]

Performance given by Cosine correlation and absolute distance measures is far better than Euclidean distance which is used by many other CBIR systems. We can recommend the use of these two measures for CBIR which has given positive improvement.

Accuracy, Completeness, Efficiency and Strength of the system has been proved by using three different performance evaluation parameters PRCP, Longest String and LSRR. Maximum longest string obtained is 95 and many are within 50 to 60 is good achievement for the system. Even LSRR parameter proved its best by giving the average for all 20 queries from 20 different classes that traversal below 75% is giving recall 1.

Overall precision and recall means accuracy and completeness of the proposed CBIR system is **0.55,** which indicates very good performance of this system as an average of 200 queries as compared to other CBIR systems.

References

[1] Dimitri A. Lisin, Marwan A. Mattar, “Combining local and global image features for object class recognition”, Intelligente System, Fachbereich Informatik, TUDarmstadt, Germany.

[2] Younes Raoui, El Houssine, Michel DEVY, “Global and local image descriptors for Content Based Image Retrieval and object recognition”, Applied Mathematical Sciences, Vol. 5, no. 42, 2109 – 2136, 2011.

[3] Yong Rui and Thomas S. Huang “Image Retrieval: Current Techniques, Promising Directions, and Open Issues”, Journal of Visual Communication and Image Representation **10**, 39–62 (1999)

[4] Seongtaek Jo, Kyhyun Um, “A Signature Representation and Indexing Scheme of Color-Spatial Information for Similar Image Retrieval” 0-7695-0577-5/00 $10.00 0 2000 IEEE.

[5] Ki-Hyun Kang, Yong-In Yoon, Jong-Soo Choi, “Additive Texture Information Extraction Using Color Coherence Vector” Proceedings of the 7th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 15-17, 2007.

[6] SubrahmanyamMurala, Anil BalajiGonde, R. P. Maheshwari, “Color and Texture Features for Image Indexing and Retrieval”, 2009 IEEE International Advance Computing Conference (IACC 2009), Patiala, India, 6-7 March 2009.

[7] Nidhi Singhai, Prof. Shishir K. Shandilya (July 2010): A Survey On: Content Based Image Retrieval Systems , International Journal of Computer Applications (0975 – 8887)Volume 4 – No.2.

[8] Egon L. van den Broeka, Peter M. F. Kistersb, and Louis G. Vuurpijla (2000)The utilization of human color categorization for content-based image retrieval : http://www.icsi.berkeley.edu/wcs/

[9] Ying Liua, Dengsheng Zhanga, Guojun Lua,Wei-Ying Mab,Y. Liu.(2007) : A survey of content-based image retrieval with high-level semantics / Pattern Recognition 40 (2007) 262– 282.

[10] Nam Yee Kim, Kang Soo You, Gi-Hyoung Yoo, Hoon Sung Kwak, An Efficient Histogram Algorithm for Retrieval from Lighting Changed-Images, 978-0-7695-3546-3/08 $25.00 © 2008 IEEE DOI 10.1109/FGCNS.2008.30. Second International Conference on Future Generation Communication and Networking Symposia.

[11] H. B. Kekre, Kavita Sonawane,“ Image Retrieval Using Histogram Based Bins of Pixel Counts and Average of Intensities”, (IJCSIS) International Journal of Computer Science and Information Security,Vol. 10, No.1, 2010.

[12] Wei-Min Zheng1, Zhe-Ming, “Color Image Retrieval Schemes using index histograms based on various spatial-domain vector quantizes” International Journal of Innovative Computing, Information and Control ICIC International ISSN 1349-4198 Volume 2, Number 6, December 2006.

[13] H. B. Kekre, Kavita Patil, Bin Pixel Count, Mean and Total of Intensities Extracted From Partitioned Equalized Histogram for CBIR, International Journal of Engineering Science and Technology (IJEST) ISSN : 0975-5462 Vol. 4 No.03 March 2012.

[14] Fan-Di Jou, Kuo-Chin Fan, Yang-Lang Chang ,”Efficient matching of large-size histograms” Institute of Computer Science and Information Engineering, National Central University, Chung-Li 32054, Taiwan, ROC Pattern Recognition Letters 25 (2004) 277–286

[15] Hafner, J., Sawhney, H.S., Equitz, W., Flickner, M., Niblack,W., 1995. Efficient color histogram indexing for quadratic form distance functions. IEEE Trans. Pattern Anal.Machine Intel. 17 (7), 729–736.Jou, F.D., 2003.

[16] Vito Di Gesu and Valery Starovoitov. Distance-Based Functions for Image Comparison. Patter Recognition Letters, 20:207–214, 1999.

[17] J.R.Smith and S.-F Chang,Integrated spatial and feature image query, Multimedia Syst. 7(2) ,129-140 (1999).

[18] J. R. Smith. Integrated Spatial and Feature Image Systems: Retrieval, Analysis and compression PhD thesis, Columbia University, New York, NY, February 1997.

[19] John R. Smith and Shih-Fu Chang,”Tools and Techniques for Color Image Retrieval” Columbia University Department of Electrical Engineering and Center for Telecommunications Research, New York, N.Y. 10027

[20] H. B. Kekre, KavitaSonawane,“Feature Extraction in the form of Statistical Moments Extracted to Bins formed using Partitioned Equalized Histogram for CBIR”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-3, February 2012.

[21] H. B. Kekre , Kavita Sonawane, Bins Formation using CG based Partitioning of Histogram Modified Using Proposed Polynomial Transform ‘Y=2X-X2’for CBIR, www.ijacsa.thesai.org, Vol. 3, No. 5, 2011.

[22] H. B. Kekre and Kavita Sonawane, “Feature extraction in bins using global and local thresholding of images for CBIR”, (IJ-CA-ETS) , ISSN: 0974-3596, October ’09 – March ’10, Volume 2, Issue 2.

[23] H. B. Kekre and Kavita Sonawane, “Bins Approach to Image Retrieval using Statistical Parameters Based on Histogram Partitioning of R, G, B Planes”, International Journal of Advances in Engineering & Technology, Jan 2012.

[24] H. B. Kekre and Kavita Sonawane, “Effect of Similarity Measures for CBIR Using Bins Approach”, International Journal of Image Processing (IJIP), Volume (6) : Issue (3) : 2012.

[25] Dr. H. B. Kekre, KavitaSonawane, “Retrieval of Images Using DCT and DCT Wavelet Over Image Blocks”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011.

[26] H. B. Kekre, Kavita Sonawane(2011) : Query based Image Retrieval using Kekre’s, DCT and Hybrid wavelet Transform over 1st and 2nd Moment, International Journal of Computer Applications (0975 – 8887),Volume 32– No.4, October 2011.

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg.from Jabalpur University in 1958,M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S. Engg. (Electrical Engg.)from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. For last 13 years worked as a Professor in Department of Computer Engg. atThadomalShahani Engineering College, Mumbai. He is currently Senior Professor working with Mukesh Patel School of Technology Management and Engineering, SVKM’s NMIMS University, Vile Parle(w), Mumbai, INDIA. He has guided 17 Ph.D.s, 150 M.E./M.Tech Projects and several B.E./B.Tech Projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networks. He has more than 450 papers in National / International Conferences / Journals to his credit. Recently twelve students working under his guidance have received best paper awards. Five of his students have been awarded Ph. D. of NMIMS University. Currently he is guiding eight Ph.D. students. He is member of ISTE and IETE.

Ms. Kavita V. Sonawane has received M.E (Computer Engineering) degree from Mumbai University in 2008, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engg, SVKM’s NMIMS University, Vile-Parle (w), Mumbai, INDIA. She has more than 8 years of experience in teaching. Currently working as a Assistant professor in Department of Computer Engineering at St. Francis Institute of Technology Mumbai. Her area of interest is Image Processing, Data structures and Computer Architecture. She has 16 papers in National/ International conferences / Journals to her credit. She is member of ISTE.