Effective Image Retrieval Based on an Experimental Combination of Texture Features and Comparison of Different Histogram Quantizations in the DCT Domain

Fazal Malik and Baharum Baharudin
Computer and Information Sciences Department, Universiti Teknologi PETRONAS, Malaysia

Abstract: *The compressed domain is appealing for the image retrieval because of the direct efficient feature extraction; moreover, currently almost all the images are available in a compressed format using the Discrete Cosine Transformation (DCT). In this paper, the quantized histogram statistical texture features are extracted from the DCT blocks using the significant energy of the DC and the first three AC coefficients of the blocks and are used for the retrieval of the similar images. The effectiveness of the image retrieval is analyzed by performing an experimental comparison of the different combinations of the texture features to get an optimum combination and the comparison of the different quantization bins by using the optimum combinations of the features. The proposed approach is tested by using the corel image database and the experimental results show that the proposed approach has a robust image retrieval using the combinations of the features with the different histogram quantization bins in the frequency domain.*

Keywords: *Compressed domain, feature extraction, DCT, statistical texture features, quantized histogram.*

Received July 10, 2012; accepted January 16, 2013; published online April 4, 2013

1. Introduction

The Internet and the other image capturing devices provide a huge number of the images due to which an efficient and an effective retrieval system is needed to retrieve these images based on the contents of the images like color, texture and shape. This system is called the Content Based Image Retrieval (CBIR). CBIR is an intensive and difficult area of the research [9].

Feature extraction and similarity measurement are the two steps to be performed by the CBIR system. In the feature extraction, the contents (features) of the image are extracted and stored in the form of a Feature Vector (FV) to create a feature database. In the similarity measurement, the user query image FV is compared with all the Feature Vectors (FVs*)* in the database for the similarities to retrieve the most similar images to the query image from the database [13, 22].

The storage space and manipulation of the images are the two problems caused by the availability of the huge number of images due to the advance development in the image capture devices, Internet, and computer hardware. To achieve the solutions of these issues, currently most of the images are represented in a compressed format like the JPEG and MPEG [9, 13]. The low level features of the image can be extracted directly in the compressed domain without decoding to the pixel domain to reduce the computation cost and increase the processing efficiency [10]. Therefore, in the compressed domain, the features can be extracted from the images by using the Discrete Cosine Transformation (DCT) which is also involved in the process of the image compression. During the compression process, some significant information is lost by the DCT transformation and some perceptually crucial information is left behind, which has an important role in the retrieval of the similar images. This information is utilized for the best image retrieval by quantizing the histograms of the DCT blocks using a proper quantization factor and a number of DCT blocks [24].

The nine FVs are constructed by using only the DC and the AC coefficients in the nine different directions of the 8×8 DCT transformed blocks of the grayscale level distribution in an image. This method gives the good performance in terms of the similar image retrieval [3]. The image in the YUV color space is divided into four blocks and in each block only the Y channel is converted into the DCT blocks to get the horizontal, vertical and diagonal texture features in all the blocks [20].

The energy histograms are constructed by using only the DC and some of the AC coefficients in the certain directions. These histograms are combined to create the FVs for the retrieval of the similar images. This approach provides a high performance in terms of the retrieval using a medium sized database for the experiment [8]. The DC coefficients vector is combined with the AC coefficients in the nine directions to get the FV of the texture features of the JPEG format images. The AC coefficients describe the texture information [16]. By computing the mean and the standard deviation in the DCT blocks, the coefficients are represented as the statistical features in the compressed domain to obtain a good performance in terms of efficiency and effectiveness. The results of this approach show robustness to the rotation and translation of the images [5]. The DCT coefficients are used to calculate the central moments of the second and the third order as the texture features to retrieve the similar images [21]. The DC and the first three AC coefficients of all the 8×8 DCT blocks are picked up in a zigzag order to construct the quantized histograms of 32 bins to be used as the FVs for the retrieval. This approach is tested by using the animal dataset of the Corel database [11].

An experimental comparison of the statistical features such as the color expectancy, skewness, color variance and cross correlation is performed using a color histogram, which concludes that the cross correlation has the good performance in terms of the retrieval [17].

For the last three decades, the study of the texture features such as the MPEG7 edge histogram, relational invariant feature histogram, global texture features, Gabor features and Tamura texture histogram have not fully described the texture properties of the images. To overcome this issue, various texture features are combined in different combinations to get a high performance in terms of the image retrieval [4].

The histograms are quantized in bins to reduce the computation cost of the feature extraction. The RGB color image consists of the three channels of 256×256×256 = 16777216 bins and it can be quantized into 8×8×8 = 512 bins to reduce the computations. This computation will increase the performance in terms of efficiency and it will be a trade-off between the efficiency and the computation time [6].

In this paper, our main contribution is to show the performance of the image retrieval by the combination of the different texture features to get an optimum combination of the texture features in terms of the precision and the comparison of the different histogram quantization bins using the optimum combinations of the texture features to get a optimum quantization in terms of the precision in the DCT domain. To the best of our knowledge, such a combination of the statistical texture features and the comparisons among the different quantization bins have not been reported in the literature for the statistical quantized histogram texture features of the grayscale images in the DCT domain. In our approach, we start with the non-overlapping 8×8 DCT block transformation of a grayscale image. The histograms of the DC and the first three AC coefficients are constructed. The statistical texture features: mean, standard deviation, skewness, kurtosis, energy, entropy and smoothness are calculated in the different combinations to get the optimum combination by using the probability distribution of the intensity levels in the histogram bins of all the blocks. The optimum combination of the features is selected on the basis of the retrieval of the images. Next another comparison of the different histogram quantization bins is performed using the optimum combination of the features for the analysis to get the optimum quantization for the effective image retrieval. In this paper, we demonstrate the combination of the statistical quantized histogram texture features and the comparison of the quantization bins, which give the best performance in terms of the image retrieval in the DCT domain.

The rest of the paper is organized as such that section 2 presents the DCT block transformation. Section 3 describes the construction and the quantization of the histograms in the DCT blocks. The proposed features and their extraction are elaborated in section 4. Section 5 describes the similarity measurement of the image retrieval. Section 6 analyzes the experimental results of the combination of various texture features and the comparison of the quantization bins in terms of the retrieval precision. Finally, section 7 concludes this paper.

2. DCT Block Transformation

In our proposed approach, first the input RGB color image is converted into the grayscale image as shown in Figure 1. To reduce the computations, because it consists of only a single plane while the RGB image consists of the three planes: Red, Green and Blue [2], the grayscale image is divided into the simple non-overlapping 8×8 blocks. Then, all these blocks are transformed into the DCT blocks in the frequency domain. Each block is in a 2-dimensional matrix. The 2-dimensional DCT block of the size $N \times N$ for $i, j = 1, 2, \ldots, N$ can be calculated as:

$$F(u,v) = \frac{1}{\sqrt{2N}} c(u) c(v) \sum_{x=1}^{N} \sum_{y=1}^{N} f(x,y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \times \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (1)$$

$$c(u) = \begin{cases} \frac{1}{\sqrt{2}} & if u = 0 \\ 1 & if u > 0 \end{cases}$$

Where $F(u, v)$ is the transformed block, $f(x, y)$ is the element of the block and N is the size of the block.

The first uppermost DCT coefficient in the DCT block is $F(0, 0)$ in equation 1; it is the average intensity value of a block and it is also called the DC coefficient or energy of the block. The other coefficients of the DCT blocks are named as AC coefficients, which are corresponding to the different frequencies (co-sinusoidal).

After the DCT transformation, the DC coefficients of all the blocks and the first three AC coefficients (AC1, AC2 and AC3) are selected in a zigzag order as shown in Figure 2, all these DC and AC coefficients will be used to construct the histograms.

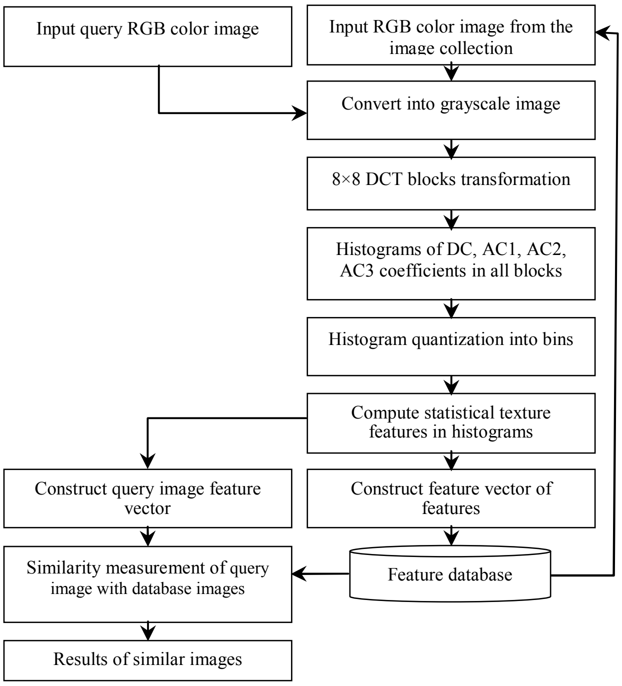


Figure 1. Block diagram of the proposed approach.

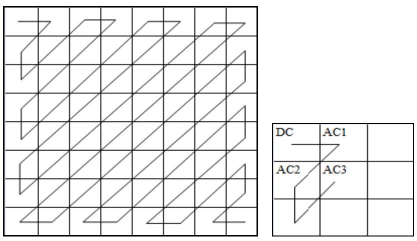


Figure 2. Coefficients selection in 8×8 DCT block in zigzag order.

3. Histogram Quantization

The histogram of the DC is defined as the frequencies of the occurrences of the DC coefficients. The histogram of the AC coefficients is the frequencies of the occurrences of the AC coefficients in all the blocks. The DC histogram is then quantized into L bins and calculated as:

$$H_{DC} = \{h(b_1)_{DC}, h(b_2)_{DC} \ldots . h(b_L)_{DC}\} \tag{2}$$

Where $h(b_i)_{DC}$ is the frequency of the DC coefficients in bin b_i and H_{DC} is the histogram of the L bins.

The histograms of the AC coefficients are also quantized into L bins and constructed as:

$$H_{AC1} = \{h(b_1)_{AC1}, h(b_2)_{AC1}, \ldots, h(b_L)_{AC1}\} \tag{3}$$

$$H_{AC2} = \{h(b_1)_{AC2}, h(b_2)_{AC2}, \ldots, h(b_L)_{AC2}\} \tag{4}$$

$$H_{AC3} = \{h(b_1)_{AC3}, h(b_2)_{AC3}, \ldots, h(b_L)_{AC3}\} \tag{5}$$

Where $h(b_i)_{AC1}$, $h(b_i)_{AC2}$ and $h(b_i)_{AC3}$ are the frequencies and H_{AC1}, H_{AC2} and H_{AC3} are the histograms of AC1, AC2 and AC3 coefficients using the quantization of the L bins.

4. Feature Extraction

The statistical texture features are considered useful for the classification and retrieval of the similar images. These texture features provide the information about the properties of the intensity level distribution in the image like the uniformity, smoothness, flatness, contrast and brightness. The statistical texture features mean, standard deviation, skewness, kurtosis, energy, entropy and smoothness are calculated by using the probability distribution of the intensity levels in the bins of the DC, AC1, AC2 and AC3 histograms.

Let $P(b)$ be the probability distribution of a bin b in each of the histograms and is calculated for $b = 1, 2, \ldots, L$ bins as:

$$P(b) = \frac{H(b)}{M} \tag{6}$$

where M is the total number of blocks in image I. The first texture feature is the mean which represents something about the brightness of the image. The mean measures the average value of the intensity values. If the mean value is high, then it shows that the image is bright and if the value is low, then the image is dark. The mean can be calculated [14, 15] as:

$$mean = \sum_{b=1}^{L} bP(b) \tag{7}$$

The standard deviation is the second order moment and it shows the contrast of the gray level intensities. The low value of the standard deviation indicates the low contrast and the high value shows the high contrast of the image and can be computed [14, 19] as:

$$stddev = \sqrt{\sum_{b=1}^{L} (b - mean)^2 P(b)} \tag{8}$$

The third order moment is the skewness and it shows the skewness of the intensity values. It is the measurement of the inequality of the intensity level distribution about the mean. The calculated value of the skewness will be positive or negative. The negative value indicates that a large number of intensity values are laying on the right side of the mean and the skewness of the tail of the intensity values is towards the left side of the distribution or the tail on the left side is longer than the right side. The positive value indicates that a large number of the intensity values are on the left side of the mean and the skewness of the tail of the intensity values is towards the right side of the distribution or the tail on the right side is longer than the left side. The zero value indicates that the distribution of the intensity values is relatively equal on both sides of the mean. The skewness [7, 14, 18] can be defined as:

$$SKEW = \frac{1}{(stddev)^3} \sum_{b=1}^{L} (b - mean)^3 P(b) \tag{9}$$

The fourth order moment is the kurtosis and is used to measure the peak of the distribution of the intensity values around the mean. The high value of the kurtosis indicates that the peak of the distribution is sharp and the tail is longer and fat. The low value of the kurtosis indicates that the peak of the distribution is rounded and the tail is shorter and thinner. The kurtosis [7, 14, 18] can be defined as:

$$kurtosis = \frac{1}{(stddev)^4} \sum_{b=1}^{L} (b - mean)^4 P(b) \tag{10}$$

The energy feature measures the uniformity of the intensity level distribution. If the value is high, then the distribution is to a small number of intensity levels. The energy can be defined [14, 15] as:

$$ENERGY = \sum_{b=1}^{L} [P(b)]^2 \tag{11}$$

The entropy measures the randomness of the distribution of the coefficient values over intensity levels. If the value of entropy is high, then the distribution is among more intensity levels in the image. This measurement is the inverse of the energy. A simple image has low entropy while a complex image has high entropy. The entropy can be computed [14, 15] as:

$$ENTROPY = -\sum_{b=1}^{L} P(b) \log_2[P(b)] \tag{12}$$

The smoothness texture is measured by using the standard deviation value. It can be defined [19] as:

$$SM = 1 - \frac{1}{1 + (stddev)^2} \tag{13}$$

After the calculation of these texture features, the *fv* of these values is constructed as:

$$fv = \{mean, stddev, SKEW, kurtosis, ENERGY, ENTROPY, SM\} \tag{14}$$

The FVs for all the histograms are calculated such as fv_{HDC} for H_{DC}, fv_{HAC1} for H_{AC1}, fv_{HAC2} for H_{AC2} and fv_{HAC3} for H_{AC3} histograms. The four FVs are combined together into a single *FV* such as:

$$FV = [fv_{HDC}, fv_{HAC1}, fv_{HAC2}, fv_{HAC3}] \tag{15}$$

The *FVs* of all the images are constructed and stored to create a feature database. The FV of the user query is also constructed in the same way and compared with the FVs of the database for the similarity and retrieval of the relevant images. The block diagram of the proposed approach is shown in Figure 1.

5. Similarity Measurements

Once the feature database of the images is created with the FVs using equations 2 to 15 in the first phase of the approach as shown in Figure 1, then the user can give an image as a query to retrieve the similar images from the database. The FV of the query image is computed by using equations 2 to 15 as the second phase of the same approach as shown in Figure 1.

To measure the similarity between the query image and the database images, the difference is calculated between the query FV and the database FVs by using the distance metric. The small difference between the two FVs indicates a large similarity and a small distance. The vectors of the images with a small distance are most similar to the query image. The distance metric, which we have included in this work, is the Euclidean distance. This distance metric is most commonly used for the similarity measurement in the image retrieval because of its efficiency and effectiveness. It measures the distance between the two vectors of the images by calculating the square root of the sum of the squared absolute differences.

Let the query FV be represented by Q and the database FV by D to calculate the difference between the two vectors for $i = 1, 2, ..., n$ features using the Euclidean distance as:

$$\Delta d = \sqrt{\sum_{i=1}^{n} \left(|Q_i - D_i|\right)^2} \tag{16}$$

Where n is the number of features. Both the images are the same for $\Delta d = 0$ and the small value of Δd shows the most relevant image to the query image.

6. Experimental Results

The proposed approach is tested by using the Corel database of images, which is freely available for the researchers [23]. The database consists of 1000 images having 10 categories and each of which has 100 images. The categories are people, beaches, buildings, buses, dinosaurs, elephants, roses, horses, mountains, and food. All these categories are used for the experiments. All the images are in the RGB color space. They are in the JPEG format with the sizes of 256×384 and 384×256 pixels.

6.1. Phases of the Proposed Approach

The approach is performed in two phases:

- *Phase 1:* In the first phase, all the images are acquired one after another, from the collection of the images for the feature extraction. The extracted features are stored in a database in the form of FVs using equation 15 to create a feature database as shown in Figure 1.
- *Phase 2:* In the second phase, the user is asked to input the query image to retrieve the relevant images from the feature database by using the same approach. The FV is constructed using equation 15 and compared with the FVs of the database by computing the similarities using the distance metric in equation 16. The relevant images are displayed to

the user according to the query image as shown in Figure 1.

In the experiment, the two phases of approach are performed for the single and the different combinations of the texture features and for all the histogram quantization bins, one by one, using all the images as the query images of database. The average precision is calculated in all the experiments and the results are analyzed and compared to obtain the effective image retrieval.

6.2. Performance Analysis of the Combination of the Texture Features

To obtain the effective retrieval of the similar images, the comparison of the different combinations of the statistical texture features is performed since the single texture feature does not give a complete description of the image to present. The two steps in Figure 1 are performed for the single and the different combinations of the features using the 10 image categories of the image database. The features which are extracted from the images are represented into the FV form using equation 15. These FVs are used in matching of the query image with the database images for the similarity measurement. The size of the FVs varies due to the different combinations of the features, for example by considering the mean and standard deviation, then each FV of the DC and the first three AC coefficients will have two features using equation 14 and these four FVs with the two features are combined to get a single FV of eight features using equation 15. The results of the different combinations of the texture features in the quantization of histograms in the 32 bins in the DCT domain using the JPEG images are shown in Tables 1, 2 and 3.

Table 1. Average precision of the combination of the seven featues (C32) in percenatge for all image categories.

Categories	Precision
Dinosaurs	100
Roses	100
Horses	93
People	92
Buses	89
Beaches	86
Elephants	76
Buildings	68
Foods	56
Mountains	47
Average	81

Table 1 shows the average precision of all the image categories for the C32 combination of the seven texture features using the quantized histograms in the DCT domain. It can be seen that the dinosaurs, roses, horses, people and buses give better results as compared to the other categories. All the images of each category are used as the query images. For C32 combination, the histograms are quantized into 32 bins. The overall average precision is 81%, which shows good image retrieval.

Table 2. Average precision of the single and the different combinations of the features in percentage using all the image categories.

Code	Single Features (SF)	Precision
C01	Mean	61
C02	Standard Deviation	67
C03	Skewness	62
C04	Kurtosis	69
C05	Energy	61
C06	Entropy	71
C07	Smoothness	33
	Average	61
	Two Features Combination (2FC)	
C08	Mean + Standard deviation	71
C09	Skewness + Kurtosis	72
C10	Energy + Entropy	61
C11	Kurtosis + Energy	69
C12	Kurtosis + Entropy	70
	Average	69
	Three Features Combination (3FC)	
C13	Energy + Entropy + Smoothness	51
C14	Mean + Standard Deviation + Skewness	74
C15	Kurtosis + Energy + Entropy	69
C16	Mean+ Standard Deviation +Kurtosis	79
C17	Mean+ Standard Deviation +Energy	71
C18	Mean+ Standard Deviation +Entropy	74
C19	Mean+ Standard Deviation +Smoothness	71
	Average	70
	Four Features Combination (4FC)	
C20	Mean + Standard Deviation + Skewness + Kurtosis	80
C21	Mean + Standard Deviation + Skewness + Energy	74
C22	Mean + Standard Deviation + Skewness + Entropy	76
C23	Mean +Standard Deviation + Skewness+ Smoothness	74
C24	Skewness + Kurtosis + Energy + Entropy	73
C25	Mean + Standard Deviation + Kurtosis + Energy	79
C26	Mean + Standard Deviation + Kurtosis + Entropy	80
C27	Mean + Standard Deviation + Kurtosis + Smoothness	79
C28	Mean + Standard Deviation + Energy + Entropy	74
C29	Mean + Standard Deviation + Energy + Smoothness	71
C30	Mean + Standard Deviation + Entropy + Smoothness	74
C31	Kurtosis + Energy + Entropy + Smoothness	72
	Average	76
	All Seven Features Combination	
C32	Mean + Standard Deviation + Skewness + Kurtosis + Energy + Entropy + Smoothness	81

Table 2 and Figure 3 show the results of the single and the different combinations of texture features using all the images of all the categories and the histogram quantization of the 32 bins. It can be seen that the performance of the texture features in terms of precision increases as the combination of the features also increases. The result of using single statistical features gives a low performance of 61% average precision. The two texture feature combination gives a comparatively good average precision of 69%. Three feature random combinations give an average precision of 70%. Four feature random combinations provide improved performance in terms of an average precision of 76%. Finally the combination C32 of all the seven features gives the best results in terms of average precision of 81%. It has been noted that the combination of mean and the standard deviation with the other features give the best results, especially with the skewness, kurtosis and entropy. While the combination of the other features likes smoothness, energy, entropy, skewness and kurtosis with each other in different combinations like the two and the three feature combinations give relatively low performance in terms of the precision. However the combination of all the seven features gives a high performance.

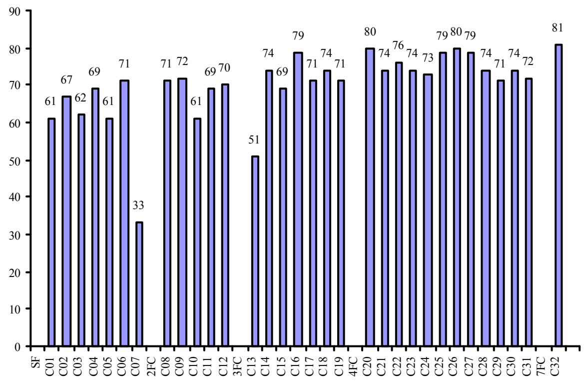


Figure 3. Average precision of the single and the different combinations of texture features for all the image categories.

Table 3. Average precision of the categories of the cobinations.

Feature Combination	Average Precision
Single Feature	61
Two Features Combination	69
Three Features Combination	70
Four Features Combination	76
All Seven Features Combination	81

Table 3 shows the overall results in the average precision of the different combinations. It can be seen that the precision increases from the single features alone, towards all the seven features combination. The four feature and the seven features combinations give a high performance in terms of precision such as 76% and 81%. Therefore, the combination of the texture features of the quantized histograms in the DCT domain using the 32 bin quantization gives good performance in terms of precision in the retrieval of the similar images instead of using a single feature or a combination of texture features less than three features.

Figure 4 shows the average time taken by the different combinations of the texture feature for the creation of the feature database using all the image categories.

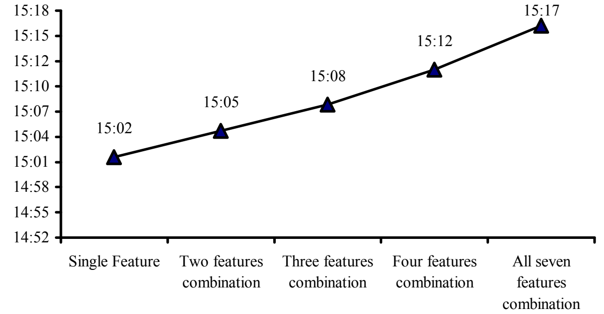


Figure 4. Average time (in minutes) taken by the different combinations of features for creation of the feature database.

It can be seen that the combination of all the seven features takes more time with less differences of time in seconds only than other combinations but the retrieval performance is higher than the other combinations.

6.3. Performance Analysis of the Different Histogram Quantization Bins

The best performance in terms of precision is 81% and 80%, given by the C32 and C20 combinations of the seven and the four features as shown in Table 2. We use these two features combinations to get the optimum quantization of bins for the effective retrieval of the similar images in the DCT domain.

The two steps in Figure 1 are performed for the different number of quantization bins using the images of all the categories. The histograms are quantized into 4, 8, 16, 32, 64, 128 and 256 bins for the both combinations C20 and C32 of the texture features. The histograms are constructed using equations 2 to 5 for the DC and the first three AC coefficients of the DCT blocks. The size of all the four histograms varies due to the quantization of histograms into the different number of bins. Then, the texture features are extracted from the quantized histograms and are combined in a FV form using equation 15. These features vectors are used in matching of the query image with the database images for the similarity measurement. The results of the different quantization bins using the two best combinations of the texture features in the DCT domain for the JPEG images are shown in Tables 4 and 5.

Table 4. Comparison of the proposed approach with the other approaches in terms of precision using the Corel image database.

Categories	Alnihoud[1]	Mohamed *et al.* [11]	Murala *et al.* [12]	Proposed Algorithm
Dinosaurs	100	99	100	100
Elephants	50	70	62	76
Horses	85	40	91	93
Roses	97	N/A	80	100
People	87	N/A	76	92
Buses	84	N/A	52	89
Beaches	68	N/A	50	86
Buildings	70	N/A	47	68
Foods	63	N/A	63	56
Mountains	32	N/A	28	47
Average	74	70	65	81

Table 4 shows the average precision of the different quantization bins using all the categories of images for the C32 combination of the texture features in the DCT domain. The average precision is calculated for each quantization bin from 4 to 256 bins. It can be seen that the dinosaurs, roses, horses, people and buses give better results as compared to the other categories. The images of these categories have good contrast, brightness, textures and smoothness. All the images of each category are used as query images in all the quantization bins for the retrieval of the similar images.

Figure 5 shows the results of the different quantization bins for all the categories of images. It can be seen that the best performance of the quantization histogram bins in terms of precision increases as the number of bins also increases from 4 to 32 bins. But after the 32 bin quantization, the precision decreases as the number of bins continues to increase from 64 to 256 quantization bins. This shows that the 32 bin quantization not only increases the efficiency but also provides good energy in the DC and in the first three AC coefficients in the DCT blocks in the frequency domain for the best retrieval of the JPEG images in terms of an average 81% precision. The quantization of 16 and 64 bins also give the better results. But the quantization of 4, 8 128 and 256 bins give the lower results as compared to the 32 bins. This shows that the distribution of intensity levels in the minimum and maximum quantization bins gives lower results while the results of the quantization bins in the middle like 16, 32 and 64 bins, are higher using all the seven features.

Table 5. Average precision of the different hitogram quantization bins using the combination C32 of all the seven texture features for all the image categories.

Categories	Bins							Average
	4	8	16	32	64	128	256	
Dinosaurs	100	100	100	100	100	100	100	100
Roses	76	62	93	100	100	100	100	90
People	100	100	100	92	78	67	61	85
Horses	62	93	93	93	83	76	71	82
Buses	57	70	86	89	92	84	82	80
Beaches	70	66	86	86	84	84	73	78
Elephants	63	70	73	76	78	77	80	74
Buildings	66	50	63	68	54	48	49	57
Foods	41	47	43	56	54	56	58	51
Mountains	56	46	46	47	46	53	58	50
Average	67	69	78	81	77	74	73	75

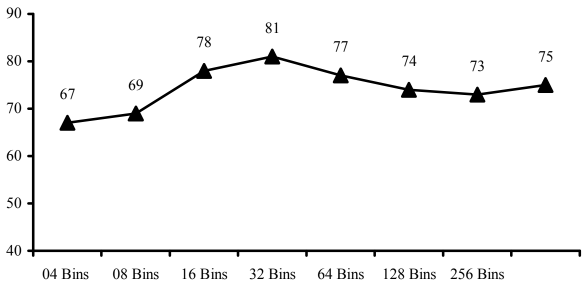


Figure 5. Average precision of the combination C32 using 32 bins quantization.

Table 5 shows the average precision of the different quantization bins using all the categories of the images for the C20 combination of the four texture features in the DCT domain. The average precision is calculated for each quantization bin for all the categories of the images. It can be seen that the dinosaurs, roses, horses, people and buses give better results as compared to the other categories.

Figure 6 shows that the best performance of the quantization histogram bins in terms of the precision is given by the 32 bins quantization as compared to the other quantization bins. This shows that the 32 bins quantization not only increases the efficiency of the feature extraction but also provides a good energy in the DC and in the first three AC coefficients in the DCT blocks in the frequency domain for the best retrieval of the JPEG images in terms of an average 80% precision. The quantization of 16 and 64 bins also give better results.

The result of the C20 combination of the texture features is almost the same as the combination C32, for all the quantization bins. The precision of the 32 bins quantization is the best for both the combinations of C20 and C32 of the texture features.

Figure 7 shows the time taken by the different number of quantization bins for the creation of the feature database using all the image categories. The incremental line shows the time is also increasing when the number of histogram quantization bins is increasing.

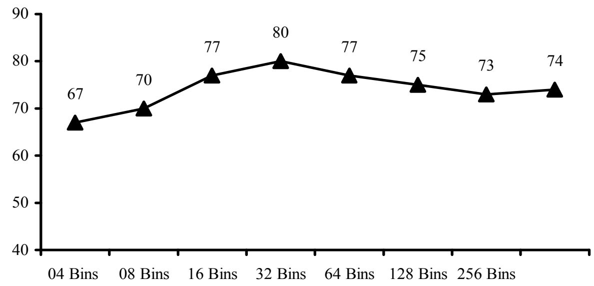


Figure 6. Average precision of the combination C20 of the four features using 32 bins quantization.

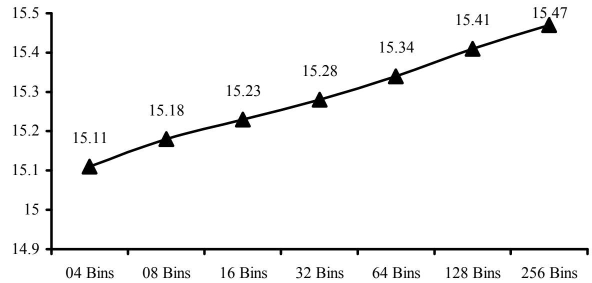


Figure 7. The time (in minutes) taken by the different number of quantization bins for the creation of the feature database.

Since the middle 32 bins gives the better performance in the retrieval, therefore it is considered best quantization even though it takes more time in seconds than the other bins like 4, 8 and 16 bins, while less than 64, 128 and 256 bins.

6.4. Performance Analysis of the Proposed Approach with the other Approaches

Our approach is compared with the approaches of [1, 11, 12] as shown in Table 6. In approach of [1], the color and shape features are extracted, based on the Self-Organizing Map (SOM). Fuzzy Color Histogram (FCH) and subtractive fuzzy clustering algorithms are used to get the color features and the object Model Algorithm is used to get the edge of objects and then shape features like area, centroid, major axis length, minor axis length, eccentricity and orientation are computed to get the performance in terms of an average precision of 74%. In approach of [11], the image is divided into the non-overlapping 8×8 blocks and then these blocks are transformed into the DCT domain. The DC and the first three AC coefficients of each block are picked up in zigzag order. These coefficients are used to construct the quantized histograms of the 32 bins. These histograms are used to construct a FV for the image retrieval. The approach was tested by using the animal dataset of 300 images of Corel database and got an average precision of 70%. In approach of [12], the color and texture features are combined for the image retrieval. For color, the mean and the standard deviation are computed in the histogram of the 32 bins for each channel of the RGB color image, to get the total 192 features. For the texture features, the mean and the standard deviation are computed in the sub bands of the Gabor Wavelet Transform image with the three scales and four orientations to get a FV of 48 features. The performance of this approach was measured in terms of an average precision of 65%.

Table 6. Average precision of the combination C20 of the four texture features using 32 bins quantization.

Categories	**Bins**							**Average**
	4	**8**	**16**	**32**	**64**	**128**	**256**	
Dinosaurs	100	100	100	100	100	100	100	100
Roses	76	62	94	100	100	100	100	90
People	100	100	100	92	78	66	61	85
Horses	62	93	93	93	84	77	71	82
Buses	58	70	86	90	92	84	82	80
Beaches	71	66	83	84	83	84	73	78
Elephants	40	64	72	76	79	77	80	70
Buildings	63	49	61	67	56	49	50	56
Foods	56	46	46	47	46	53	58	50
Mountains	41	46	39	56	53	56	58	50
Average	67	70	77	80	77	75	73	74

In our approach, we start with the non-overlapping 8×8 DCT block transformation of a grayscale image. The histograms of the DC and the first three AC coefficients are constructed. The statistical texture features mean, standard deviation, skewness, kurtosis, energy, entropy and smoothness are computed in different combinations to get the optimum combination by using the probability distribution of the intensity levels in the histograms of 32 bins of all the blocks. After several experiments for different combination we get the best average precision of 81% for C32 combination of the seven texture features. Then best combination is used for the analysis of the quantization of the histograms into the different number of bins and gets 81% average precision for 32 bins while 75% of all quantization bins. We have also used the same Corel dataset which was used by the other approaches. Our proposed approach has better results in terms of precision in the DCT domain using the texture features, especially for the combination of the seven texture features in the 32 bins quantization as shown in Figures 8 and 9.

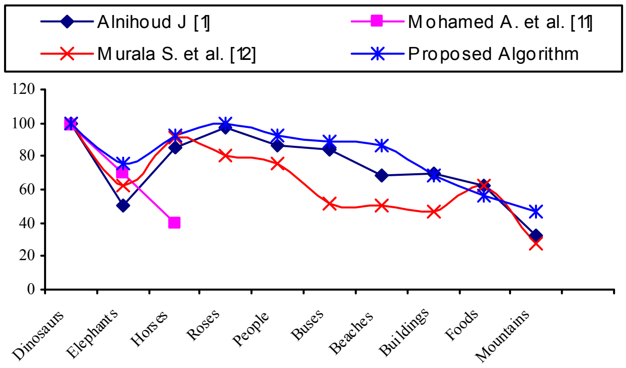


Figure 8. Comparison of the proposed approach with other approaches based on precision in percentage, category wise.

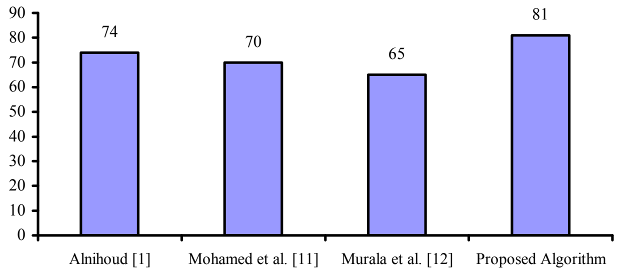
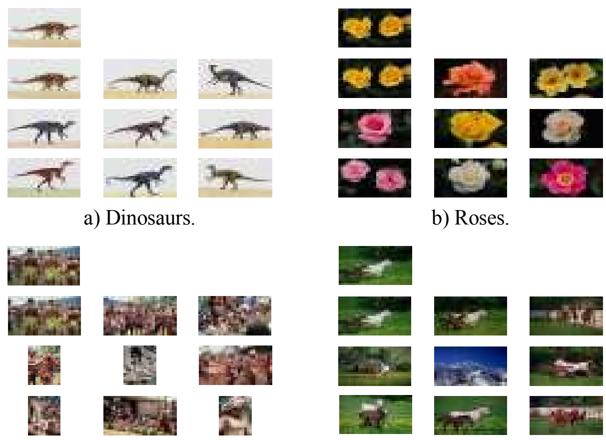


Figure 9. Comparison of the proposed approach with other approaches based on average precision.

The proposed approach is invariant to the rotation and illumination. By changing the rotation of the image at different angles: 0°, 45°, 75°, 90°,135°,180° and 235° and by clipping the illumination of the images at the rate from 10% to 90%, the results of the retrieval were not so affected.

Figure 10 (a-d) shows the results of the query images and in each result the top single image is the query image and the below nine images are the relevant retrieved images. The results show that our proposed approach has good retrieval accuracy.



a) Dinosaurs. b) Roses.

c) People. d) Horses using combination C32 and 32bins quantization.

Figure 10. Query results.

7. Conclusions

In this paper an approach is proposed for CBIR in which the statistical texture features are extracted from the quantized histograms in the DCT domain. Only the DC and the first three AC coefficients having more significant energy are selected in each DCT block to get the statistical quantized histogram texture features. The approach is performed in two steps: In the first step, the analysis of the results of the different combinations of the statistical quantized histogram texture features is performed for the optimum feature combination. Experimental results show that the combination of more features gives better results as compared to a single feature or a small number of the texture features combination. In the second step, the comparison of the different quantization bins is performed using the best and the optimum texture features combinations.

The quantization in 32 bins for the four and the seven texture features combinations give the best results in terms of precision as compared to the low and high bin quantization. In the experiments, all the images of the database are used as query images. We conclude that the combination of more statistical texture features and the quantization of 32 bins for the optimum combination of features give good performance in terms of precision in the DCT domain for the compressed JPEG images.

References

[1] Alnihoud J., "Content-Based Image Retrieval System Based on Self Organizing Map, Fuzzy Color Histogram and Subtractive Fuzzy Clustering," *the International Arab Journal of Information Technology*, vol. 9, no. 5, pp. 452-458, 2012.

[2] Anjum M. and Javed M., "Multiresolution and Varying Expressions Analysis of Face Images for Recognition," *Information Technology Journal*, vol. 6, no. 1, pp. 57-65, 2007.

[3] Bae H. and Jung S., "Image Retrieval Using Texture Based on DCT," *in Proceedings of the International Conference on Information, Communications and Signal Processing*, pp. 1065-1068, 1997.

[4] Deselaers T., Keysers D., and Ney H., "Features for Image Retrieval: An Experimental Comparison," *Information Retrieval*, vol. 11, no. 2, pp. 77-107, 2007.

[5] Feng G. and Jiang J., "JPEG Compressed Image Retrieval via Statistical Features," *Pattern Recognition*, vol. 36, no. 4, pp. 977-985, 2003.

[6] Jeong S., "Histogram-Based Color Image Retrieval," available at: http://scien.stanford.edu/pages/absite/2002/psych221/projects/02/sojeong/#dct, last visited 2001.

[7] Kekre H. and Sonawan K., "Statistical Parameters Based Feature Extraction Using Bins with Polynomial Transform of Histogram," *the International Journal of Advances in Engineering & Technology*, vol. 4, no. 1, pp. 510-524, 2012.

[8] Lay J. and Guan L., "Image Retrieval Based on Energy Histograms of the Low Frequency DCT Coefficients," *in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, USA, vol. 6, pp. 3009-3012, 1999.

[9] Liu Y., Zhang D., Lu G., and Ma W., "A Survey of Content-Based Image Retrieval with High-Level Semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.

[10] Mandal M., Idris F., and Panchanathan S., "A Critical Evaluation of Image and Video Indexing Techniques in the Compressed Domain," *Image and Vision Computing*, vol. 17, no. 7, pp. 513-529, 1999.

[11] Mohamed A., Khellfi F., Weng Y., Jiang J., and Ipson S., "An Efficient Image Retrieval through DCT Histogram Quantization," *in Proceedings of the International Conference on Cyber Worlds*, Bradford, USA, pp. 237-240, 2009.

[12] Murala S., Gonde A., and Maheshwari R., "Color and Texture Features for Image Indexing and Retrieval," *in Proceedings of IEEE International Advance Computing Conference*, Patiala, India, pp. 1411-1416, 2009.

[13] Nezamabadi-Pour H. and Saryazdi S., "Object Based Image Indexing and Retrieval in DCT Domain Using Clustering Techniques," *in Proceedings of World Academy of Science Engineering and Technology*, vol. 3, pp. 207-210, 2005.

[14] Selvarajah S. and Kodituwakku S., "Analysis and Comparison of Texture Features for Content Based Image Retrieval," *the International Journal of Latest Trends in Computing*, vol. 2, no. 1, pp. 108-113, 2011.

[15] Sergy´an S., "Color Histogram Features Based Image Classification in Content-Based Image Retrieval Systems," *in Proceedings of the 6th International Symposium on Applied Machine Intelligence and Informatics*, Herlany, Slovakia, pp. 221-224, 2008.

[16] Shan Z. and Liu Y., "Research on JPEG Image Retrieval," *in Proceedings of the 5th International Conference on Wireless Communications, Networking and Mobile Computing*, Beijing, China, pp. 1-4, 2009.

[17] Sharma N., Rawat P., and Singh J., "Efficient CBIR Using Color Histogram Processing," *Signal & Image Processing: An International Journal*, vol. 2, no. 1, pp. 94-112, 2011.

[18] Suresh P., Sundaram R., and Arumugam A., "Feature Extraction in Compressed Domain for Content Based Image Retrieval," *in Proceedings of IEEE International Conference on Advanced*

Computer Theory and Engineering, Phuket, Thailand, pp. 190-194, 2008.

[19] Thawari P. and Janwe N., "CBIR Based on Color and Texture," *the International Journal of Information Technology and Knowledge Management*, vol. 4, no. 1, pp. 129-132, 2011.

[20] Tsai T., Huang Y., and Chiang T., "Dominant Feature Extraction in Block-DCT Domain," *in Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, Taipei, China, vol. 5, pp. 3623-3628, 2006.

[21] Vailaya A., Jain A., and Zhang H., "On Image Classification: City Images vs. Landscapes," *Pattern Recognition*, vol. 31, no. 12, pp. 1921-1935, 1998.

[22] Veltkamp R. and Tanase M., "Content Based Image Retrieval Systems: A Survey," *Technical Report*, Utrecht University, 2002.

[23] Wang J., Li J., and Wiederhold G., "Simplicity: Semantics-Sensitive Integrated Matching for Picture Libraries," *in Proceedings of the 4th International Conference on Advances in Visual Information Systems*, Berlin, Germany, pp. 360-371, 2000.

[24] Zhong D. and Defee I., "DCT Histogram Optimization for Image Database Retrieval," *Pattern Recognition Letters*, vol. 26, no. 14, pp. 2272-2281, 2005.

Fazal Malik received his Bachelor and Master of computer science degrees from University of Peshawar, Pakistan. He is currently a PhD student at the Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Malaysia. His current research interests include image retrieval and indexing.

Baharum Baharudin received his Bachelor and Masters of science degrees from Central Michigan University, USA, and his PhD degree from the University of Bradford, UK. He is currently an associate professor at the Department of Computer and Information Sciences, Universiti Teknologi PETRONAS Malaysia. His research interests lies in image processing, data mining and knowledge management.