Computers and Electrical Engineering 000 (2017) 1–14

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

An efficient framework for image retrieval using color, texture and edge features☆

L.K. Pavithra*, T. Sree Sharmila

Department of Information Technology, SSN College of Engineering, Kalavakkam, Chennai, India

A R T I C L E I N F O

Article history:
Received 17 November 2016
Revised 24 August 2017
Accepted 24 August 2017
Available online xxx

Keywords:
Canny edge detection
Color moments
Content-Based Image Retrieval (CBIR)
Local Binary Pattern (LBP)
Manhattan similarity measure

A B S T R A C T

This paper proposes a new hybrid framework for Content-Based Image Retrieval (CBIR) system to address the accuracy issues associated with the traditional image retrieval systems. The proposed framework initially selects pertinent images from a large database using color moment information. Subsequently, Local Binary Pattern (LBP) and Canny edge detection methods are used to extract the texture and edge features respectively, from the query and resultant images of the initial stage of this framework. Then, the Manhattan distance information about these two features corresponding to the query and selected images are calculated and combined, and then sorted using bubble sort algorithm. Wang's, Corel-5K and Corel-10K are the three databases used for evaluating the performance of the proposed hybrid framework using precision and recall measures. The average precision measured on these three databases gives approximately 11.8%–22.315%, 8.025%–18.935% and 10.755%–32.221% higher accuracy than the state-of-the-art techniques.

1. Introduction

Nowadays digital imaging has become an indispensable segment in many applications such as medical imaging, remote sensing, crime prevention, education, multimedia, data mining, etc. These applications require digital images as a source for various processes like segmentation, object recognition, tracking, and others. To index and search suitable images from the rapidly increasing digital image collections, an image retrieval system is used. In image databases, for fast and efficient retrieval, images are indexed and searched in two ways, namely, using keywords and using visual contents of an image. The traditional Text-Based Image Retrieval (TBIR) system uses keywords for semantic image retrieval. But this system fails to handle large database because of its non-automatic keyword generation scheme and it fully relies on the perception of human experts who are employed in the keyword generation task. This often leads to inappropriate keyword generation for an image. Whereas, in a Content-Based Image Retrieval (CBIR) system, the visual contents i.e., color, texture, shape, etc., of an image are utilized to give plausible results from the large database [1]. Among the visual contents, color plays a vital role in human perception by giving a pleasant view of the environment. It can also be used to identify an object and distinguish one object from another. Hence, in CBIR, color can be considered as the leading descriptor due to its simplicity in calculation and invariant behavior towards translation, rotation and change in the viewing angle [2]. Color information can be extracted from the image by both global and local techniques. Most traditional global color extraction methods use color histogram

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Area Editor Dr. E. Cabal-Yepez.

* Corresponding author.
E-mail addresses: pavithralk@ssn.edu.in (L.K. Pavithra), sreesharmilat@ssn.edu.in (T.S. Sharmila).

http://dx.doi.org/10.1016/j.compeleceng.2017.08.030

which gives the total count of each color pixel present in an image. Even though color histogram is prominently used to label the color information in an image, it has flaws in representing the spatial information since it gives same histogram representation for entirely different images if their color distributions are identical [3]. Like color, texture can also be a robust descriptor in image retrieval system. In today's systems, most of the commonly used texture extraction methods follow the local template structure for mining the texture information [4]. The texture information is often estimated locally from the gray-level representation of an image [5,6]. The shape of the object is a salient part in recognizing similar images from the database. This shape details can be extracted from the Fourier transformed images [7]. The CBIR system based on mapping between any one of the visual features and the relevant image has a large semantic gap. In order to account for this, a system based on the association between more features has been evolved. The most familiar combination in feature level fusion is the color-texture fusion. Furthermore, texture-shape and shape-color based systems are more ingeniously developed. But still, these coupled feature based systems have failed to meet the user requirements, mainly due to the inability of the feature fusion techniques to relate specific information present in the query image with the database image. Hence, unification of all visual content based retrieval systems is introduced.

In recent years, combined feature based image retrieval is an active research area in CBIR. Because of issues that exist in finding suitable feature extraction methods and their combinations, precision of the retrieval system is still scanty. The proposed framework mainly concentrates on improving the retrieval accuracy of the CBIR system by integrating the low-level features such as color, texture and edge of an image in a multilevel fashion. The main contributions of this work in the CBIR field are as follows: 1) Propose a new framework for image retrieval system which is in contrast to the existing system that uses unification of different features in the same level. 2) Reduce the search space of the similar image retrieval system with the help of color based selection process and hence the system distance measure for color feature need not be taken into account. 3) Individual and proposed integrated feature-based relevant image retrieval systems are extensively tested over the standard database.

The remaining section of this paper is framed as follows: Section 2 describes related works on the field of combined feature based CBIR system. The novel features integration technique is explained in Section 3. Section 4 gives the experimental results and discussion. Conclusion and future scope of this work are discussed in Section 5.

2. Related work

This section presents the state-of-the-art methods in CBIR system based on the combination of multiple features. Color histogram and Gray Level Co-occurrence Matrix (GLCM) features are fused by Yue et al. [8] to give more accurate retrieval results in CBIR system. Here, global and local histograms are evaluated over HSV color space image. Local histogram ends up giving better performance than global histogram and then GLCM is used to extract the texture feature of a gray level image. Local histogram and GLCM features are fused by giving equal weight to color and texture features. Color Co-occurrence Matrix (CCM), Difference Between Pixels of Scan Pattern (DBPSP) and Color Histogram K-Mean (CHKM) image features are fused [9] to get the highly similar image results. Among the three features, two features (CCM and DBPSP) are used to extract the texture information and the third feature (CHKM) is used to extract only the color information.

Multi-scale edge field method for multimedia retrieval [10] uses Canny edge extraction as a part of process to obtain the object boundaries in different scales. Agarwal et al. [11] have applied Canny edge detection on the luminance channel of the YCbCr color image in order to improve performance of the image retrieval system. Liu and Yang [12] have proposed Color Difference Histogram (CDH) on Lab color space that is completely different from long established color histogram method. Lab color space is preferred for estimating CDH because it uses the color difference between color and edge orientation texture details of the image. Subsequently, Canberra distance metric is used to measure the similarity between query and database images. Furthermore, texture and color features based retrieval is obtained by local extrema peak valley pattern and RGB colour histogram [13]. The texture and shape features are mined using Local Ternary Pattern (LTP) and geometric moments to pull off large amount of relevant images [14].

Recently, neural network structure [15] is proposed to reduce the semantic gap in image retrieval. The network is trained from the Bags Of Images (BOI) which have the third level decomposed wavelet packet tree information and the mean eigenvector of each Gabor filter response image. Then, Pearson correlation co-efficient is used to find the similarity between the feature vectors. Finally, the outputs are refined with the help of relevance feedback mechanism. But, the training phase complexity and convergence time of this approach are high.

Color, texture and shape features are integrated by Wang et al. [16] to give efficient retrieval by CBIR system. Color feature is calculated via slightly modified Dominant Color Descriptor (DCD). Here, the image is segmented into eight coarse parts, and then mean value of each segment acts as its quantized color. Next, the clustering is used to merge the nearest colors. Finally, five dominant colors and its percentages are obtained. The texture feature is extracted by taking convolution between image and band pass filter with 4-directions namely horizontal, vertical, 45° and -45°. Also, Pseudo- Zernike moments are used to compute the shape feature of the image because this moment is invariant to rotation and insensitive towards noise. The color and texture [17] details are taken from Lab color space using modified CDH then additionally Angular Radial Transform (ART) based invariant shape features are extracted. The extracted features are normalized using min-max and Z-score normalization methods. Among the two normalization methods, min-max based values gave highly accurate images.

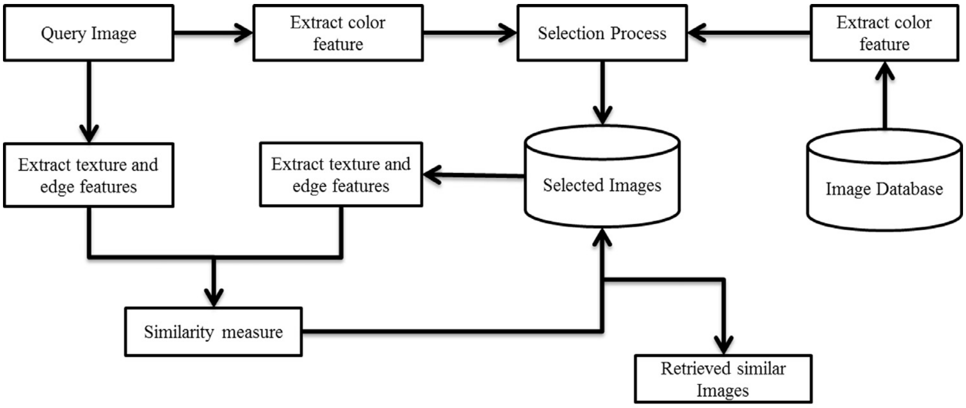

Fig. 1. Block diagram of color, texture and edge based CBIR.

Walia et al. [18] evolved a new flow in image retrieval through merging color, texture and shape feature of an image to achieve a semantic retrieval. The first 3 color moments are calculated from 100×100 HSV image in the initial stage of the retrieval process to get the top 30 sorted images. After that, for every sorted image Edge Histogram Descriptor (EHD) and ART are explored respectively for texture and shape feature extraction. Then, these features are fused along with the help of weighing factor to obtain the relevant images.

In the works mentioned above, all low-level features are extracted from the entire image database and query image. The distance between each of those features is calculated, combined and is used as a metric to identify similar images. Since, these approaches involve feature extraction and distance measure irrespective of their category, it results in increased computational load for the system. The proposed work aims to minimize the computational load of the conventional system by reducing the search space of similar images without compromising the retrieval accuracy of the CBIR system. It uses the image selection module as the first stage of the retrieval process. Hence, feature descriptor used in relevant image selection process must satisfy the following properties i) Invariance to image rotation, translation and noise ii) Less complexity in feature extraction process iii) Compact feature representation to increase the speed and correctness of the retrieval process. In this work, statistical measures such as mean and standard deviation from each color channel of the query image are taken as the initial parameter to achieve desired reduction in search space during retrieval process. Images that are similar to the query image in the database are selected for the next level processing (see Section 4.1). In the second level, low-level features such as texture and edge details are extracted from these reduced set of images. Thus, the computational complexity is comparatively less than conventional approaches (see Section 4.3).

3. Proposed work

The typical feature descriptors used in the proposed work to extract the local features like color, texture and edge are illustrated in this section. The generalized architecture of the proposed framework is shown in Fig. 1.

3.1. Color descriptor

In the proposed work, for minimizing the complexity and improving the effectiveness of the CBIR, the global color descriptor is used in the first level of retrieval. The review of these color descriptors claims an extensive use in many applications due to their consistent behavior. Additionally, they have a fast response for a specific request, which is comparatively higher than the local color extraction techniques used in the image. Hence, color moments (statistical measure) are chosen to represent the color details of the image. It gives the pixel distribution information of the image in two compact forms. The first order moment gives average information about the pixel distribution of a given image and the closeness of the pixel distribution about mean color is estimated by second order moment. Moreover, the usage of color feature in the proposed work is different from the conventional CBIR system although both employ color moments.

In the first stage of the retrieval process, average color information (mean) and the quantity of the amount of pixels that differs from the mean (standard deviation) of the query image are estimated globally from the three color channels (Red, Green, and Blue) of the RGB color space using Eqs. (1) and (2). If the pixels present in the image are close to the average value, it decreases value of standard deviation. A high standard deviation indicates that the huge amount of color pixel is not close to mean value.

$$Mean(Ic) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} P_{cij}, \;\; c = \{R, G, B\} \tag{1}$$

$$Std(Ic) = \left(\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(P_{cij} - Mean(Ic) \right)^2 \right)^{\frac{1}{2}}, \quad c = \{R, G, B\} \tag{2}$$

where Ic holds the color channel information of an image. M and N are the row and column size of an image. P_{cij} indicates the value of image pixel in the ith row and jth column of the particular color channel.

Only the mean value of R, G and B color channels is required for the database images to obtain the reduced search space. The mean value of color image gives the average color information of the image. It is not confirmed that the mean value is one of the pixel information of the given particular image. Illumination conditions and blurring effects of the image may affect the average color information. Therefore, the similar images with above mentioned effects direct its mean color information to spin around the average color information of the query image. Meanwhile, standard deviation of the image is also important to give details about the distribution of image pixel around the average information. This information acts as a lower and upper bound for the mean value and allows taking values between them. Hence, standard deviation details are added and subtracted with the mean. The results of these operations give two threshold values for each channel. They are represented as Low-Threshold (LT) and High-Threshold (HT) which is given as follows:

$$LT(Ic) = Mean(Ic) - Std(Ic), \quad c = \{R, G, B\} \tag{3}$$

$$HT(Ic) = Mean(Ic) + Std(Ic), \quad c = \{R, G, B\} \tag{4}$$

If the first order color moment of each channel (R, G and B) of the database images lies in between the two thresholds (including two threshold values) then those images are selected for the next stage of feature extraction process. The threshold values of the R, G and B color channels are combined by the logical AND (&&) operator. Here, first stage of the proposed work behaves as a filter that takes all images from the database and passes the images which satisfy the well-defined rule in this level. At the end of this first stage, selected images collectively form a subset from the original database. Subsequent level of the retrieval system uses this subset of images instead of using the original database for image retrieval.

3.2. Texture descriptor

Texture is another leading descriptor in a CBIR system. Due to its implementation, simplicity and energetic performance in the field of texture analysis nowadays LBP based texture extraction is extensively used in various application domains like face recognition, biometric application, etc. The proposed system also makes use of a simple LBP [19] to extract the texture of the image. This extraction algorithm is performed over the subset of images selected from the first level of the retrieval process. Before exploring LBP on the selected images, RGB to gray scale transformation is carried out as a pre-processing step on these images. For each iteration, it takes 3×3 overlapping gray scale image as input. The pixel value available in the Center Position (CP) of the 3×3 sub block acts as a threshold value for its neighboring pixels. Using this threshold value, binary representation of that sub block is created. Then, the LBP value of the 3×3 sub block is evaluated in the counter clockwise direction. Finally, the LBP value is updated in the center pixel position of that block in the image. First iteration of the LBP is illustrated in Fig. 2 and Eq. (5) shows the estimation of LBP [19] for a 3×3 block representation:

$$LBP_N = \sum_{i=0}^{N-1} f(P_i - CP)2^i; \quad f(p) = \begin{cases} 1; P \geq 0 \\ 0; P < 0 \end{cases} \tag{5}$$

where N denotes the total number of neighboring pixels for the center pixel in the 3×3 sub block, here $N = 8$. P_i holds the value of neighboring pixels, $i = \{0, 1, 2, ..., 7\}$. CP is the center pixel value of the sub block.

In Fig. 2(b), the first 3×3 sub block is taken from the simple 5×5 size gray image. In this, the center pixel value 21 is the threshold value for its 8 neighbors. The difference between the value of every neighbor pixel and the center pixel value is calculated. If the difference value is greater than or equal to 0 then that particular pixel value is turned into 1, otherwise 0 is updated in its place. Subsequently, these 8 bit binary values are converted to decimal values and rehabilitated in the place of the center pixel as shown in Fig. 2(c). After obtaining the LBP for the whole image, histogram of LBP values are computed which gives the representation of texture feature of an image.

3.3. Edge descriptor

Normally, edges are formed by the abrupt change in the intensity value of the image which is captured by the edge detection algorithms and it holds the boundary representation of the objects present in an image [10]. Low -level visual content of an image can also be expressed and preserved in the form of edges. Human perception is highly sensitive to edges [11]. Since Canny edge detection is used to represent the shape of the object, the proposed work uses Canny edge extraction on the selected images at the end of the first stage. In RGB color space, each color channel is highly correlated with other

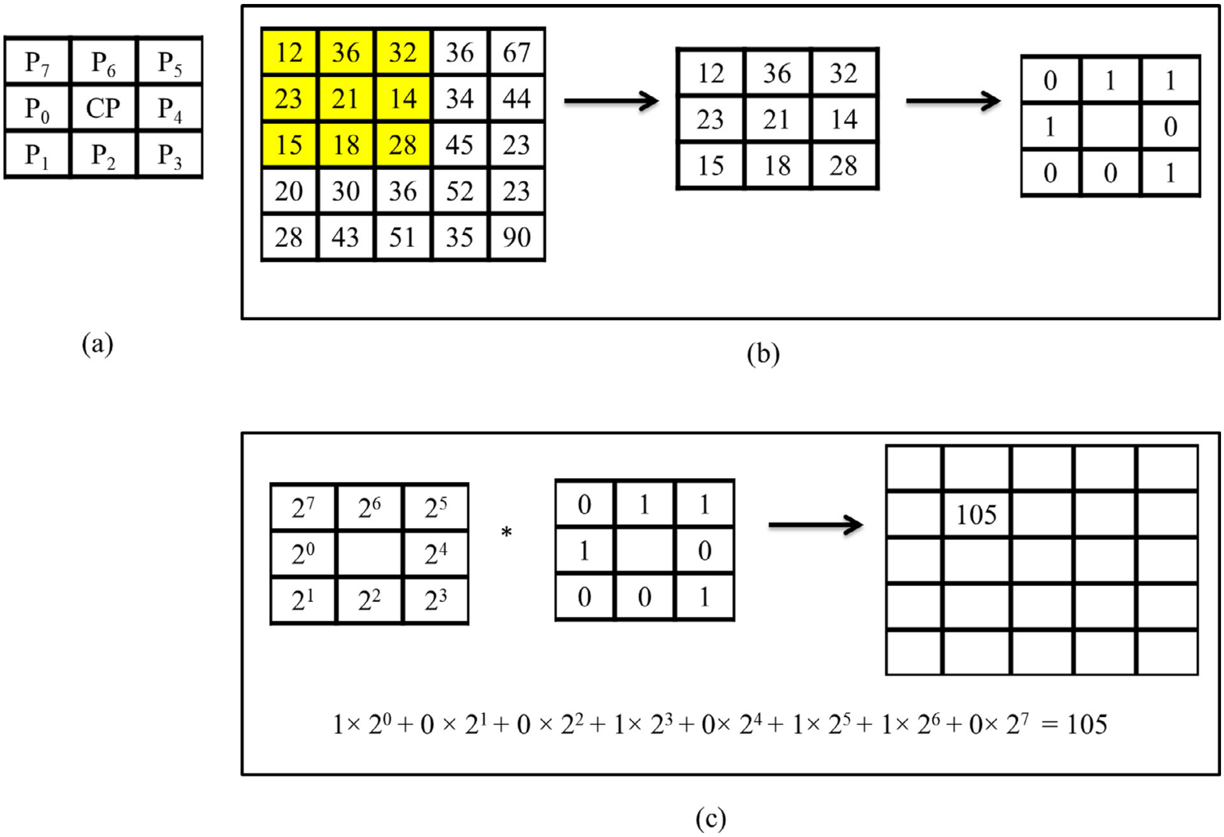

Fig. 2. First iteration of LBP: (a) Pixel position of 3×3 block image (b) Binary value creation using threshold value (c) LBP value generation through the obtained binary values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

color channels such that the splitting of chrominance and luminance information is impossible and is perceptually non-uniform to human perception. Gray scale information is enough to mark edges in images but gray to RGB conversion is not possible to produce color image. Hence, color space transformation has to be performed to obtain the edge details from the intensity plane of an image. This is the first step in edge feature extraction process. RGB to HSV conversion [20] takes place as per Eqs. (6)–(8).

$$H = \begin{cases} 60 \times \left(\frac{(G-B)}{\delta}\right), & R = \max(R, G, B) \\ 60 \times \left(2 + \left(\frac{B-R}{\delta}\right)\right), & G = \max(R, G, B) \\ 60 \times \left(4 + \left(\frac{R-G}{\delta}\right)\right), & B = \max(R, G, B) \end{cases} \tag{6}$$

where $\delta = \max(R, G, B) - \min(R, G, B)$

$$S = \frac{\delta}{\max(R, G, B)} \tag{7}$$

$$V = \max(R, G, B) \tag{8}$$

where H and S carry the chrominance details of the given image. V channel holds the intensity distribution of that image. Canny edge detection algorithm is run over the V channel. Then, the edge extracted V channel is combined with un-modified H and S channel and transformed back to RGB color space. After that, these color edge features are estimated through the histogram of R, G and B channels separately. Over all processing procedure of the proposed work is depicted in the following algorithm.

Please cite this article as: L.K. Pavithra, T.S. Sharmila, An efficient framework for image retrieval using color, texture and edge features, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.08.030

Algorithm: CTEBIR (Q_{IMAGE}, $DB_{NIMAGES}$) – Color, Texture and Edge Based Image Retrieval.

Input: **Query image from the user is denoted by** Q_{IMAGE}. **Image Database (DB) contains** N **number of images denoted by** $DB_{NIMAGES}$. **N denotes the total number of images.**

Output: K **number of images** $IM_1, IM_2, \ldots, IM_K \in DB_{NIMAGES}$ **retrieved as similar images for the given query image** Q_{IMAGE}

1. Separate the R, G and B color channel information from Q_{IMAGE} and $DB_{NIMAGES}$
2. Calculate the 1st and 2nd order color moments for the Q_{IMAGE} and form low and high threshold using Eqs. (3) and (4)
3. Calculate the 1st order color moment for images $\in DB_{NIMAGES}$
4. Count = 0;
5. for $i = 1$ to N do
   a. if the 1st order CM of $DB_{iIMAGES}$ lies between the low and high threshold of the Q_{IMAGE} then select the particular image (IM) and store that in New Database ($NewDB$)
   b. Count = Count + 1;
   c. $NewDB_{COUNTIMAGES} = IM$;
   d. end if

   end
6. Calculate LBP texture feature for Q_{IMAGE} and $NewDB_{COUNTIMAGES}$
7. Calculate Canny edge feature for Q_{IMAGE} and $NewDB_{COUNTIMAGES}$
8. for each image in $NewDB_{COUNTIMAGES}$ do
   a. Measure the similarity metric between Q_{IMAGE} and $NewDB_{COUNTIMAGES}$ using normalized texture and edge feature
   b. Combine the calculated similarity metric of the texture and edge feature of the images $\in NewDB_{COUNTIMAGES}$

   end
9. Bubble sorting is performed on $NewDB_{COUNTIMAGES}$ according to the value of each image which is obtained from step 8(b)
10. return K number of similar images $IM_1, IM_2, \ldots, IM_K$

3.4. Similarity measure

Similarity measure is mandatory in all kinds of the retrieval system as it gives the measure of distance between the low-level visual contents of two images. Distance information is the deciding factor of the similarity measure [1]. The very low resultant value indicates that corresponding database image is very close to the given query image. LBP and edge features similarity are estimated through the Manhattan distance measure [17] which is given by the Eqs. (9) and (10):

$$LBP\_SM(QLBP_{IMAGE}, NewDBLBP_{COUNTIMAGES}) = \sum_{i=1}^{N} |f_{QLBP_{IMAGE}}(i) - f_{NewDBLBP_{COUNTIMAGES}}(i)| \tag{9}$$

where $f_{QLBP_{IMAGE}}(i)$ is the ith LBP feature of the query image, $f_{NewDBLBP_{COUNTIMAGES}}(i)$ denotes ith LBP feature of an image in the new database and N holds the total count of LBP feature in an image.

$$\begin{aligned} EDGE\_SM(QEDGE_{IMAGE}, NewDBEDGE_{COUNTIMAGES}) = &\sum_{i=1}^{N} |f_{QEDGE\_R_{IMAGE}}(i) - f_{NewDBEDGE\_R_{COUNTIMAGES}}(i)| \\ &+ \sum_{i=1}^{N} |f_{QEDGE\_G_{IMAGE}}(i) - f_{NewDBEDGE\_G_{COUNTIMAGES}}(i)| \\ &+ \sum_{i=1}^{N} |f_{QEDGE\_B_{IMAGE}}(i) - f_{NewDBEDGE\_B_{COUNTIMAGES}}(i)| \end{aligned} \tag{10}$$

where, $f_{QEDGE\_R_{IMAGE}}(i)$, $f_{QEDGE\_G_{IMAGE}}(i)$ and $f_{QEDGE\_B_{IMAGE}}(i)$ denotes ith edge feature of the query image on the R, G and B color channels, $f_{NewDBEDGE\_R_{COUNTIMAGES}}(i)$, $f_{NewDBEDGE\_G_{COUNTIMAGES}}(i)$ and $f_{NewDBEDGE\_B_{COUNTIMAGES}}(i)$ gives the i^{th} edge feature of the new database image on the R, G and B color channels. NR, NG and NB have the total number of edge features in R, G and B color channels of an image respectively.

The obtained LBP and edge feature distance measure values randomly vary from one another in an unbounded way. Hence, normalization is essential to limit the large and small variations in the feature value to the range of [0, 1]. Min-Max normalization [21] is implemented through the Eqs. (11) and (12) on the texture and edge feature distance measure information.

$$NORMAL\_LBP\_SM(i) = \frac{LBP\_SM(i) - \min(LBP\_SM)}{\max(LBP\_SM) - \min(LBP\_SM)}, i = 1, 2, \ldots, K \tag{11}$$

where K is the total number of images in the new database which is not a fixed one. It will change according to the selection rule in the first stage. $LBP\_SM(i)$ denotes LBP based similarity measure value of the ith image in the new database. $\min(LBP\_SM)$ and $\max(LBP\_SM)$ denote minimum and maximum texture feature similarity value for the whole set of images in new database.

$$NORMAL\_EDGE\_SM(i) = \frac{EDGE\_SM(i) - \min(EDGE\_SM)}{\max(EDGE\_SM) - \min(EDGE\_SM)}, i = 1, 2, \ldots, K \tag{12}$$

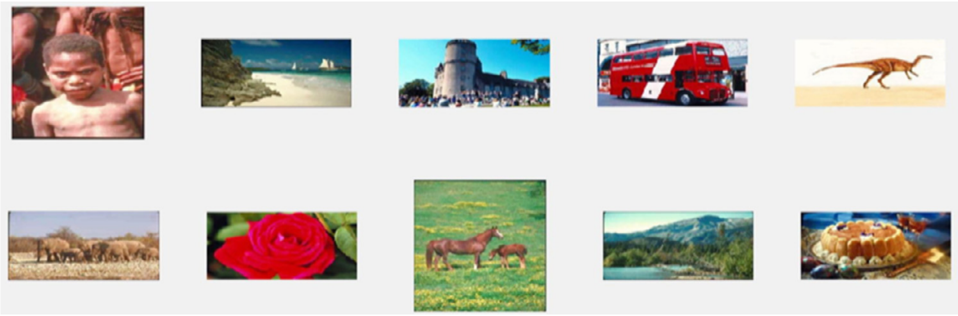

Fig. 3. Sample images from Wang's database. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

where K holds the value of total number of images in new database which clearly depends on the selection rule of the first stage. $EDGE\_SM(i)$ denotes edge distance measure of ith image in the new database, $\min(EDGE\_SM)$ and $\max(EDGE\_SM)$ is the low and high edge feature distance value of the new database images.

In the proposed framework, the estimated normalized feature values are combined using equal weights and are organized in ascending order using bubble sort to achieve homogeneous images at the top level of the retrieval. Furthermore, the proposed CBIR system is probed with the standardized performance measures over the different number of retrieval results.

4. Experimental results and discussion

For the purpose of experimentation and verification, experiments are conducted over the Wang's [22], Corel-5K [23] and Corel-10K [23] databases. Each database contains 1000, 5000 and 10,000 images respectively of size either 256×384 or 384×256(Wang's) and 126×187or 187×126(Corel-5K and Corel-10K). Corel-5K is the subset of the Corel-10K database. In Wang's database, 1000 images are divided into 10 groups and a sample image from each class (African tribes, Food, Sea, Buildings, Bus, Dinosaurs, Elephants, Flowers, Horse and Mountains) is shown in Fig. 3. Like Wang's database Corel-5K and Corel-10K database images are also divided into 50 and 100 classes and each class has 100 images into it. Here, the experiments are executed in the MATLAB R2013a, an environment along with the dual core processor, 2 GB memory and 64 bit windows operating system and the experiment is performed over the three databases.

4.1. Stage 1: Image selection

Retrieval accuracy and similar image search space of this hybrid system is highly dependent on the result of global visual content descriptor. Color is the easily assessable, more powerful and widely used visual content in image and video based retrieval systems. Beyond that, here it plays an additional role as a filter to restrict the set of images which do not fall within the limits defined in Eqs. (3) and (4). Images in the new subset database will randomly differ with respect to a query image because threshold values obtained in the selection step is fully based on the first two color moments of that image. At the end of this stage, search space for the similar image is reduced which is comparatively less than the large databases used in the initial stage. Further, the size of the databases is not predictable. Efficiency of the proposed system's image selection process is tabulated in Table 1 which is evaluated by measuring the average number of images involved in second level feature extraction process and their retrieval time. The experimental results of this work show that it approximately takes 490, 1850 and 3500 number of images from the Wang's, Corel-5K and Corel-10K database for the subsequent process. Size of the image is also responsible for the feature extraction time and retrieval efficiency. Here, feature calculation time of the Corel-5K and Corel-10K database image is less than the Wang's database image since image size of the Corel database is less. Moreover, Corel-5K is a subset of the Corel-10K database so that the average time taken for feature extraction is same in both databases but the retrieval time of these databases is different as it depends on the size of the database.

4.2. Stage 2: Texture and edge feature extraction

The feature extraction process is carried out with images from the new database and query image. For texture feature extraction, each image in the subset is converted to it corresponding gray scale form and then LBP extraction is applied over those images which serve texture information in 256 values. LBP creation of an image is shown in Fig. 4.

Meanwhile, edge features of selected images are evaluated on V channel of the HSV color space image. From the inference of edge detection [11], the Canny edge detection algorithm is employed to determine edges in the image. Then, H, S and V channels are converted to R, G and B channels and a histogram is required to give the edge features of the image. Feature vector length of the each channel is 256. Hence, the collective edge feature count is 768 for three channels. The edge features extraction outcome is shown in Fig. 5.

Please cite this article as: L.K. Pavithra, T.S. Sharmila, An efficient framework for image retrieval using color, texture and edge features, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.08.030

Table 1
Efficiency of the proposed image selection process.

CP		Image DB								
		Wang's			Corel-5K			Corel-10K		
Feature descriptors		Color feature	Texture feature	Edge feature	Color feature	Texture feature	Edge feature	Color feature	Texture feature	Edge feature
Average number of images involved in feature extraction	TM	1000	1000	1000	5000	5000	5000	10,000	10,000	10,000
	PM	1000	490	490	5000	1850	1850	10,000	3500	3500
Feature extraction time (in s) on single image		0.039	0.503	0.129	0.029	0.478	0.099	0.029	0.478	0.099
Image retrieval time (in s) with image selection (PM)	First stage		0.0550			0.1490			0.2710	
	Second stage		1.0937			3.2424			7.342	
	Total		1.1087			3.3914			7.613	
Image retrieval time (in s) without image selection			3.9213			18.328			43.654	

TM- Traditional Method, PM- Proposed Method, CP- Comparison parameters

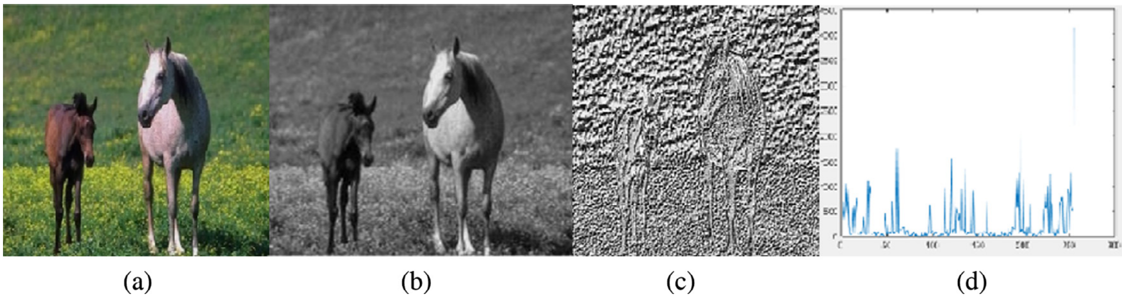

Fig. 4. LBP feature creation (a) Input image (b) Gray image (c) LBP of the gray image (d) Histogram of (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

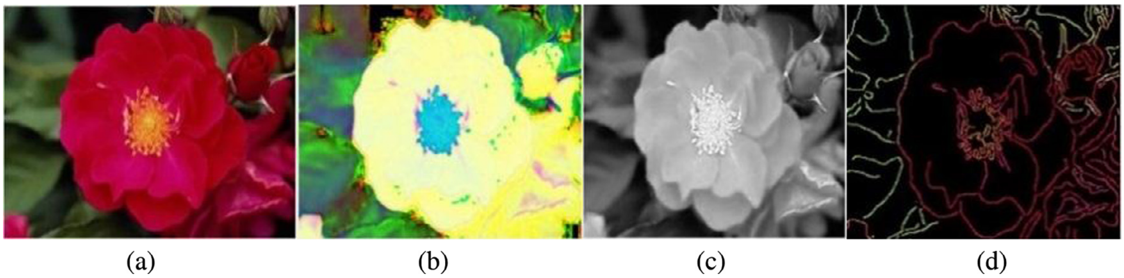

Fig. 5. Edge feature extraction (a) Input image (b) HSV converted image (c) V channel of HSV image (d) Edge extracted RGB color space image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

4.3. Performance assessment

In order to assess the efficiency of the proposed work, two performance measures precision and recall [15] are utilized to estimate the correctness and robustness of the given retrieval system. They are given a below:

$$\text{Precision} = \frac{\text{RI}}{\text{TNI}} \tag{13}$$

$$\text{Recall} = \frac{\text{RI}}{\text{TDB}} \tag{14}$$

where RI denotes the total number of relevant images retrieved, TNI is the total number of images retrieved,TDB is the total number of relevant images in the database.

In the first experiment, Wang's database is taken to execute the proposed work and its efficiency is evaluated over the same database using precision and recall. The name of each class in the Wang's database and its label value is given as follows: African tribes-1, Sea-2, Building-3, Bus-4, Dinosaurs-5, Elephants-6, Flowers-7, Horse-8, Mountain-9 and Food-10. From each class, twenty images are randomly selected to measure the performance of the proposed hybrid system. Then,

Please cite this article as: L.K. Pavithra, T.S. Sharmila, An efficient framework for image retrieval using color, texture and edge features, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.08.030

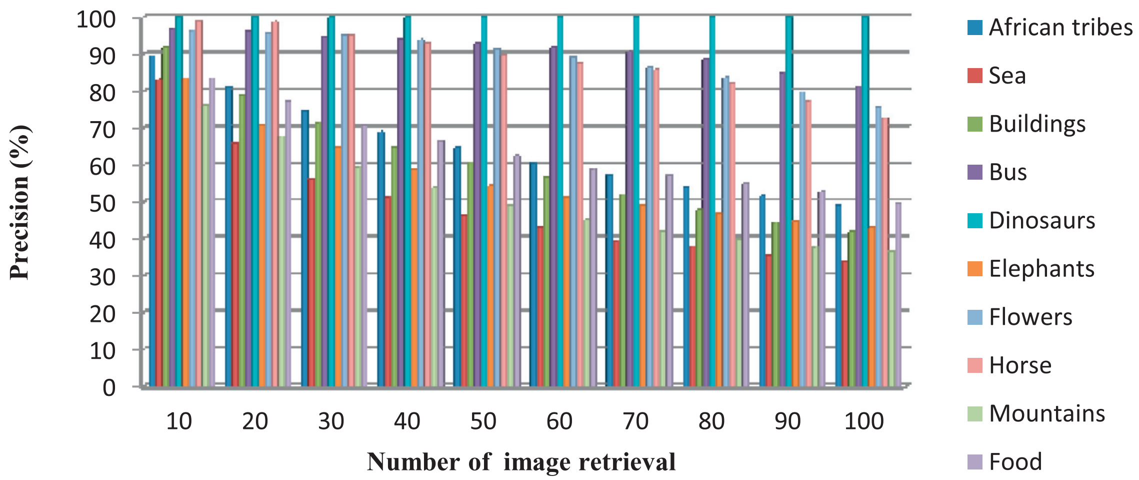

Fig. 6. Performance of the proposed system on the Wang's database. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

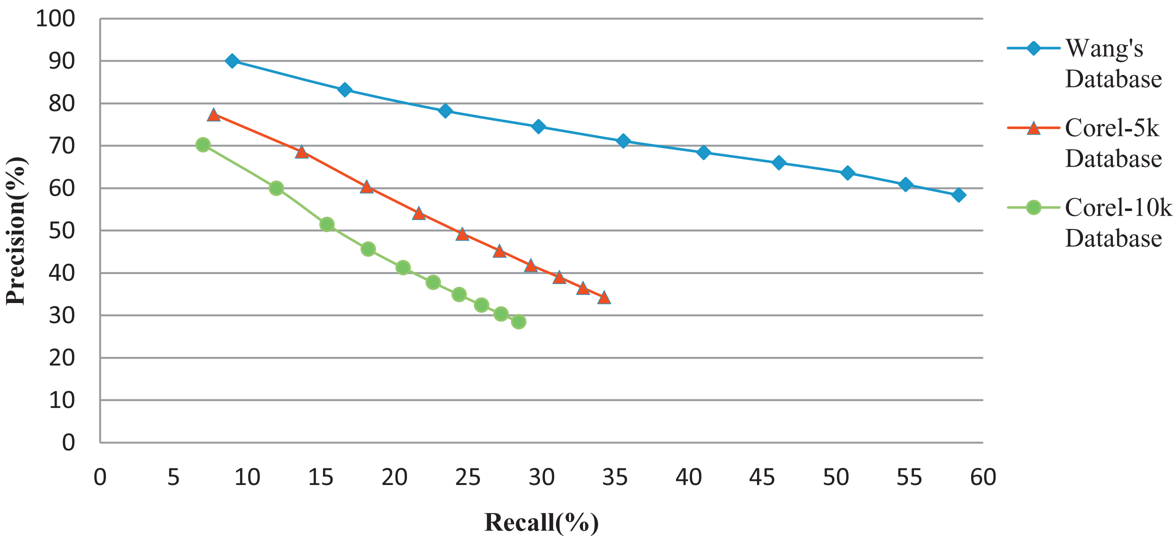

Fig. 7. Precision versus recall graph of the proposed system. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

the top 10 to 100 retrieval results are considered for overall performance evaluation of the proposed hybrid retrieval system which is shown in Fig. 6.

In experiments 2 and 3 the same procedure is carried out over Corel-5K and Corel-10K databases. The average Precision (AP) and Average Recall (AR) measure are used to estimate the performance of the proposed work by varying the retrieval ranges from 10 to 100. Then, the precision versus recall curve in the Fig. 7 is plotted using this information.

A combination of three different visual contents of the image will give large amount of highly matched images than a single feature for similar retrieval. This is to ensure the combined feature based system performance is better than an individual feature based performance. The top 20 retrieval results of the independent and combined visual content-based system is examined on the three standard databases and the accuracy results are given in Table 2. Moreover, a comparison of these two methods (individual and combined feature based methods) over the Wang's database is plotted in Fig. 9.

From the observation of average precision on 3 individuals and combined features, it is clearly visible that the efficiency of the hybrid framework averagely gives 83.225%, 68.605% and 59.98% retrieval accuracy on Wang's, Corel-5K and Corel-10K databases which is higher than the individual descriptor based systems. Individual color descriptor based retrieval is estimated by color moments with the help of Eqs. (1) and (2) and weighted distance is measured [18] using Eq. (15) and is

Fig. 8. Retrieval results for the query images (a–b) Wang's database (c) Corel-5K (d) Corel-10K. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Performance measures of the individual and the proposed hybrid feature using average precision (%).

Image Databases	Feature Extraction Techniques			
	Individual feature			Combined feature
	Color Moments AP (%)	Local Binary Pattern AP (%)	Edge AP (%)	Proposed Method AP (%)
Wang's	60.75	61.275	53.625	83.225
Corel-5K	30.035	35.3	19.085	68.605
Corel-10K	25.395	29.55	18.865	59.98

AP – Average Precision

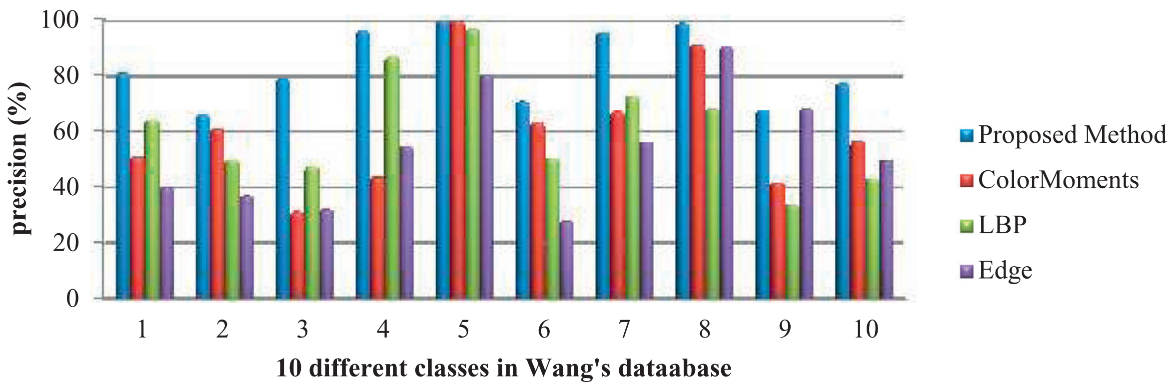

Fig. 9. Individual and combined feature comparison in Wang's database. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: L.K. Pavithra, T.S. Sharmila, An efficient framework for image retrieval using color, texture and edge features, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.08.030

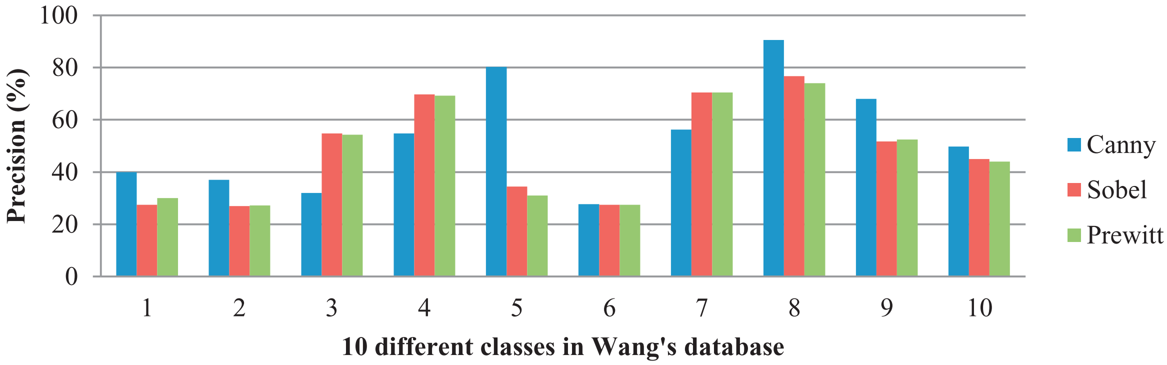

Fig. 10. Top 20 retrieval results of Canny, Sobel and Prewitt edge detection based CBIR. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Performance comparison of the proposed hybrid system with different retrieval techniques using average precision (%).

Wang's Database	Lin et al. [9]	Irtaza et al. [15]	Wang et al. [16]	Walia et al. [17]	Walia et al. [18]	Proposed Work
African Tribes	55.5	53	80.5	41.25	73	81*
Sea	66	46	56	71*	39.25	66
Building	53.5	59	48	46.75	46.25	78.75*
Bus	84	73	70.5	59.25	82.5	96.25*
Dinosaurs	98.25	99.75	100*	99.5	98	100*
Elephants	63.75	51	53.75	62	59.25	70.75*
Flowers	88.5	76.75	93	80.5	86	95.75*
Horse	87.25	70.25	89	68.75	89.75	98.75*
Mountain	48.75	62.5	52	69*	41.75	67.75
Food	68.75	70.75	62.25	29.25	53.45	77.25*
Average Precision	71.425	66.2	67.2	60.91	66.92	83.225*
Corel-5K (AP)	58.753	53.238	60.58	49.67	56.72	68.605*
Corel-10K (AP)	44.775	32..157	49.225	27.759	40.786	59.98*

Note: * indicates the best results, AP-Average Precision.

sorted in ascending order.

$$d(Q_{IMAGE}, DB_{IMAGES}) = \sum_{c=1}^{3} w_1 |Q_{IMAGE} m_c - DB_{IMAGES} m_c| + w_2 |Q_{IMAGE} std_c - DB_{IMAGES} std_c| \tag{15}$$

$w1$ and $w2$ are the user-defined weight vectors for the colorful moments here $w1 = 0.6$ and $w2 = 0.4$ which give higher performance than other combinations ((0.9, 0.1), (0.8, 0.2), etc.). Q_{IMAGE} and DB_{IMAGES} define the query and new database images. C is the index of the color channel which takes values 1, 2 and 3 (1-Red, 2-Green and 3-Blue), m and std define the mean and standard deviation values of color moments. Our proposed system will not consider the color moments distance between query and database images, since this information is used as a threshold in the first stage of this work. For the observation of individual color descriptor based image retrieval, Eq. (15) is taken to estimate the distance between the query and database. Moreover, Canny edge detector performance is investigated and compared over the other traditional edge detectors such as Sobel and Prewitt which is depicted in Fig. 10. Among 10 classes, Canny edge detector based retrieval efficiency is higher in 6 classes than the Sobel and Prewitt edge detectors. Additionally, Canny edge detection performance is better than Angular Radial Transform (ART) [18] shape descriptor. This is also a reason for selecting Canny edge detection for image retrieval.

In order to show the efficiency of the proposed work, it is compared with the state-of-the-art methods in the CBIR which is tabulated in Table 3–4. The top 20 retrieval results of each method are taken for performance evaluation.

Every user of the retrieval system wants to get comparable images at the top of the displayed search result. Therefore, this work considers the first 20 retrieval result to measure the accuracy of the hybrid system. Generally, the precision of a CBIR system gives the relevant image ratio for the resultant images. Hence, the proposed work uses this metric to quantify the retrieved results. Table 3 illustrates the retrieval result of each class image in Wang's database and it exemplifies the retrieval result of Corel database (5K & 10K). It is observed that the proposed work averagely gives 83.225% similar images for the given query image on the Wang's database. Among 10 classes, the proposed work attained higher accuracy in 6 class images in the Wang's database, namely, "African Tribes, Bus, Elephants, Flowers, Horse and Food". On account of the same background and single object present in the "Dinosaur" class images they are extensively retrieved by the proposed framework and the hybrid system in [16]. At the same time, steerable filter and Pseudo Zernike Moments [17] are used to differentiate "Sea and Mountain" class images effectively from other class images and gives 5%, 1.25% improved performance respectively than the proposed work. It is identified that proposed (Sea-66%, Building-78.75%, Elephants-70.75%, Mountain-67.75%) and

Table 4
Performance comparison of the proposed hybrid system with different retrieval techniques using average recall (%).

Wang's Database	Lin et al. [9]	Irtaza et al. [15]	Wang et al. [16]	Walia et al. [17]	Walia et al. [18]	Proposed Work
African Tribes	11.1	10.6	16.1	8.25	14.6	16.2*
Sea	13.2	9.2	11.2	14.2*	7.85	13.2
Buildings	10.7	11.8	9.6	9.35	9.25	15.75*
Bus	16.8	14.6	14.1	11.85	16.5	19.25*
Dinosaurs	19.65	19.95	20*	19.9	19.6	20*
Elephants	12.75	10.2	10.75	12.4	11.85	14.15*
Flowers	17.7	15.35	18.6	16.1	17.2	19.15*
Horse	17.45	14.05	17.8	13.75	17.95	19.75*
Mountains	9.75	12.5	10.4	13.8*	8.35	13.55
Food	13.75	14.15	12.45	5.85	10.69	15.45*
Average Recall	14.285	13.24	14.1	12.545	13.384	16.645*
Corel-5K (AR)	11.7506	10.6476	12.116	9.934	11.344	13.721*
Corel-10K (AR)	8.955	6.4314	9.845	5.5518	8.1572	11.996*

Note: * indicates the best results AR-Average Recall.

Table 5
Time complexity in feature extraction and retrieval (in seconds).

Image Databases		Lin et al. [9]	Irtaza et al. [15]	Wang et al. [16]	Walia et al. [17]	Walia et al. [18]	Proposed Work
Wang's	FET	75,632.685	0.7924	7.2224	41.288	2.7935	0.671
	RT	4.832	4.758	4.186	4.9746	4.096	1.1087
Corel-5K	FET	151,208.31	0.722	6.8252	39.6222	2.7464	0.606
	RT	9.485	8.431	10.574	8.4635	7.486	3.3914
Corel-10K	FET	378,458.322	0.722	6.8252	39.6222	2.7464	0.606
	RT	120.483	103.478	135.219	73.486	38.475	7.613

FET- Feature Extraction Time, RT- Retrieval Time

other retrieval techniques [9] (Sea-66%, Building-53.5%, Elephants-63.75%, Mountain-48.75%), [15] (Sea-46%, Building-59%, Elephants-51%, Mountain-62.5%), [16] (Sea-56%, Building-48%, Elephants-53.75%, Mountain-52%), [18] (Sea-39.25%, Building-46.25%, Elephants-59.25%, Mountain-41.75%) have less pertinent retrieval results in " Sea, Building, Dinosaurs and Mountain" class images compared to other class images (African Tribes, Bus, Elephants, Flowers, Horse and Food). The "Sea, Building, Elephants and Mountain" class images have more number of pixels related to blue, white and sand color because most of the image regions of these class images are occupied by sky and land objects. On account of similar objects in multiple classes, it gives distinct class images on the top of the search for the specific query image which is evident from the result shown in Table 3. The "Building and Elephants" class images are significantly retrieved (approximately 78.75% and 70.75%) by the proposed work with the help of the min-max normalized texture and edge features from the reduced set images. Then, experiments 2 and 3 give approximately 68.605% and 59.98% average precision over the databases, Corel-5K and Corel-10K.

From the retrieved result, recall estimates the amount of ground truth images in a particular class. Here, the first 20 retrieval results are considered for performance evaluation so that the recall value of the retrieved results can vary from 0 to 20. It is observed that in the 20 retrieved results, the proposed work takes approximately 16.645% images from the same class which is comparatively higher than the other methods used in Wang's database. Moreover, the proposed work gives 1.605%–3.787% improved average recall in Corel-5K and 2.151%–6.4442% improved average recall in Corel-10K database. Additionally, feature extraction time and retrieval time of the CBIR system is used to evaluate the rapidness of the proposed framework and other techniques [9,15,16,17 and 18] that is depicted in Table 5.

Feature extraction and response time of the CBIR system are also key parameters to measure the efficiency of this system. In CBIR, the amount of images taken for feature extraction and distance calculation is responsible for the retrieval time. Even though Lin et al. [9] gives higher similar retrieval (on Wang's database) than the other techniques [15–18], it consumes more amount of time to extract the color feature named as CHKM (CCM- 0.1953s, DBPSP-0.4897s, CHKM-7563s) because the 16-bin values are obtained by running K-Means over all images in the database. Hence, the convergence time of this CHKM feature is high [24]. Moreover, if an image is randomly added to the database, this feature is recalculated for all images. This is the drawback of the texture (CCM and DBPSP) and color (CHKM) based image retrieval technique used in [9]. Walia et al. [18] approach uses approximately 4.096s to retrieve applicable images from the database since it takes the least amount of retrieval time than the methods in comparison [9,15,16,17]. The feature extraction time of the ART based shape features only increase time for feature extraction in [18]. Further, it takes the first 30 images from the color based retrieved results for texture and shape details extraction. Therefore, it is not widely to retrieve images from the large database. The proposed framework scans the whole database for color feature extraction and it takes about 0.039s and 0.029s to extract the color features respectively from the Wang's and Corel database images (Ref Table 1). After the color feature extraction, it picks the comparable images from the database with respect to the query image. This step takes averagely 0.0550s on the Wang's database. As the database size increases, processing time of this step also increases such that Corel-5K and

Corel-10K databases takes 0.1490s and 0.2710s to perform this selection. Due to the small amount of images utilized in the second stage, the proposed work uses approximately 1.0937s, 3.2434s and 7.342s to complete the texture and edge feature extraction, and distance measure calculation over the three databases. Therefore, the total time taken to execute this work is 1.1087s, 3.3914s and 7.613s with respect to the databases used in the three experiments. Thus, the retrieval time of this work is less than the time taken by the state-of-the art techniques used in comparison.

5. Conclusion

This paper proposed a new hybrid framework for efficient retrieval of similar images using color, texture and edge features. The search space selection and feature extraction plays a major role in this proposed research. Approximately, one-third of the search space is reduced by suitable image selection, which is achieved by employing color moments for fast response over similar images. From the selected images, the texture and edge features are extracted using LBP and Canny edge detection algorithms, which lift the correctness and robustness of the proposed retrieval system with less time complexity. The proposed system has been experimented on the Wang's, Corel-5K and Corel-10K databases. The evaluation results showed that the performance of this proposed hybrid strategy is better than the individual feature based retrieval system. The proposed framework is also compared with the state-of-the-art methods, which has been revealed the efficiency of proposed system in the form of accuracy and retrieval time. This work can be applied on multiple databases and can use machine learning techniques to design the low-level feature integrated CBIR system with low-feature dimension.

References

[1] Alzu'bi A, Amira A, Ramzan N. Semantic content-based image retrieval: a comprehensive study. J Vis Commun Image Represent 2015;32:20–54. http://dx.doi.org/10.1016/j.jvcir.2015.07.012.
[2] Wang XY, Wu JF, Yang HY. Robust image retrieval based on color histogram of local feature regions. Multimedia Tools Appl 2010;49:323–45. doi:10.1007/s11042-009-0362-0.
[3] Pass G, Zabih R. Histogram refinement for content-based image retrieval. In: IEEE Workshop on Appl of Comput Vis; 1996. p. 96–102. doi:10.1109/ACV.1996.572008.
[4] Subrahmanyam M, Jonathan Wu QM, Maheshwari RP, Balasubramanian R. Modified color motif co-occurrence matrix for image indexing and retrieval. Comput Electr Eng 2013;39(3):762–74. http://dx.doi.org/10.1016/j.compeleceng.2012.11.023.
[5] Sree Sharmila T, Ramar K, Sree Renga Raja T. Developing an efficient technique for satellite image denoising and resolution enhancement for improving classification accuracy. J Electron. Imaging 2013;22(1). doi:10.1117/1. JEI.22.1.013013.
[6] Sree Sharmila T, Ramar K, Sree Renga Raja T. Impact of applying pre-processing techniques for improving classification accuracy. Signal Image Video Process 2013;8(1):149–57. doi:10.1007/s11760-013-0505-7.
[7] Brandt S, Laaksonen J, Oja E. Statistical shape features for content-based image retrieval. J Math Imaging Vis 2002;17:187–98.
[8] Yue J, Li Z, Liu L, Fu Z. Content-based image retrieval using color and texture fused features. Math Comput Model 2011;54:1121–7. doi:10.1016/j.mcm.2010.11.044.
[9] Lin CH, Chen RT, Chan YK. A smart content-based image retrieval system based on color and texture feature. Image Vis Comput 2009;27:658–65. doi:10.1016/j.imavis.2008.07.004.
[10] Kiranyaz S, Ferreira M, Gabbouj M. Automatic object extraction over multiscale edge field for multimedia retrieval. IEEE Trans Image Process 2006;15(12):3759–72. doi:10.1109/TIP.2006.881966.
[11] Agarwal S, Verma AK, Dixit N. Content based image retrieval using color edge detection and discrete wavelet transform. In: IEEE Int Conf on Issues and Chall in Intell Comput Tech (ICICT); 2014. p. 368–72. doi:10.1109/ICICICT.2014.6781310.
[12] Liu GH, Yang JY. Content-based image retrieval using color difference histogram. Pattern Recognit 2013;4(1):188–98. http://dx.doi.org/10.1016/j.patcog.2012.06.001.
[13] Dey M, Raman B, Verma M. A novel colour- and texture-based image retrieval technique using multi-resolution local extrema peak valley pattern and RGB colour histogram. Pattern Anal Appl 2015. doi:10.1007/s10044-015-0522-y.
[14] Srivastava P, Binh NT, Khare A. Content-based image retrieval using moments of local ternary pattern. Mob Netw Appl 2014;19(5):618–25. doi:10.1007/s11036-014-0526-7.
[15] Irtaza A, Arfan Jaffar M, Aleisa E, Choi TS. Embedding neural networks for semantic association in content based image retrieval. Multimedia Tools Appl 2014;72:1911–31. doi:10.1007/s11042-013-1489-6.
[16] Wang XY, Yu YJ, Yang HY. An effective image retrieval scheme using color, texture and shape feature. Comput Stand Interfaces 2011;33:59–68. https://doi.org/10.1016/j.csi.2010.03.004.
[17] Walia E, Pal A. Fusion framework for effective color image retrieval. J Vis Commun Image Represent 2014;25:1335–48. http://dx.doi.org/10.1016/j.jvcir.2014.05.005.
[18] Walia E, Vesal S, Pal A. An effective and fast hybrid framework for color image retrieval. Sens Imaging 2014;15(93):1–23. doi:10.1007/s11220-014-0093-9.
[19] Ojala T, Pietikainen M, Maenpaa T. Multi resolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002;24(7):971–87. doi:10.1109/TPAMI.2002.1017623.
[20] Mojsilovic A, Hu H, Soljanin E. Extraction of perceptually important colors and similarity measurement for image matching, retrieval and analysis IEEE Trans Image Process 2002; 11(11): 1238–1248. doi:10.1109/TIP.2002.804260.
[21] Jain A, Nandakumar K, Ross A. Score normalization in multimodal biometric systems. Pattern Recognit 2005;38:2270–85. doi:10.1016/j.patcog.2005.01.012.
[22] Wang's Database. http://wang.ist.psu.edu [Accessed 01.03.2016].
[23] Corel-5K and Corel-10K databases. http://www.ci.gxnu.edu.cn/cbi/Dataset.aspx. [Accessed 01.03.2016].
[24] Lin CH, Chen CC, Lee HL, Liao JR. Fast K-means algorithm based on level histogram for image retrieval. Experts Syst Appl 2014;41:3276–83. http://dx.doi.org/10.1016/j.eswa.2013.11.017.

L. K. Pavithra is a research scholar of SSN College of Engineering, Chennai, India. She received the B.Tech. degree in Information technology from Anna University, Tirunelveli, in 2011 and the M.E Degree in Computer Communication Engineering from Anna University, Chennai, in 2013. Her areas of interest are image and video retrieval, texture analysis and pattern recognition.

T. Sree Sharmila is working as an Associate Professor in the Department of Information Technology, SSNCE, Chennai, India. She received her Ph.D. from Anna University, in 2013, M.E in Computer Science and Engineering from Annamalai University, in 2005 and B.Tech. in Information Technology from Manonmaniam Sundaranar University, Tirunelveli, in 2003. Her areas of interest are image and video processing.

Please cite this article as: L.K. Pavithra, T.S. Sharmila, An efficient framework for image retrieval using color, texture and edge features, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.08.030