

# Content based image retrieval using hybrid features and various distance metric

Yogita Mistry\*, D.T. Ingole, M.D. Ingole

*Department of Electronics Engineering, PRM Institute of Technology & Research, Badnera, Maharashtra 444701, India*

Received 17 June 2016; received in revised form 28 September 2016; accepted 10 December 2016

Available online 5 January 2017

## Abstract

In last decade, large database of images have grown rapidly and will continue in future. Retrieval and querying of these image in efficient way is needed in order to access the visual content from large database. Content based image retrieval (CBIR) provides the solution for efficient retrieval of image from these huge image database. In this article a hybrid feature based efficient CBIR system is proposed using various distance measure. Spatial domain features including color auto-correlogram, color moments, HSV histogram features, and frequency domain features like moments using SWT, features using Gabor wavelet transform are used. Further, to enhance precision binarized statistical image features, color and edge directivity descriptor features are employed for developing efficient CBIR system. Various distance metrics are used for retrieval.

The experiments are performed using WANG database which consists of 1000 images from 10 different classes. Experimental result shows that the proposed approach performs better in terms of precision compared to other existing systems.

© 2017 Electronics Research Institute (ERI). Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** CBIR; DTCWT; SWT moments; Minkowski distance; Mahalanobis distance

## 1. Introduction

Recently collection of images have rapidly grown and continues to increase in future due to widespread use of internet. Information is collected from books, newspapers is digitized and available in digital image format. Every user is able to access this information using internet. To access these large databases, efficient method is required

\* Corresponding author.

E-mail addresses: [yogitadhumal@rediffmail.com](mailto:yogitadhumal@rediffmail.com) (Y. Mistry), [dtingole@gmail.com](mailto:dtingole@gmail.com) (D.T. Ingole), [mdingole@gmail.com](mailto:mdingole@gmail.com) (M.D. Ingole).

Peer review under the responsibility of Electronics Research Institute (ERI).



Production and hosting by Elsevier

<https://doi.org/10.1016/j.jesit.2016.12.009>

2314-7172/© 2017 Electronics Research Institute (ERI). Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

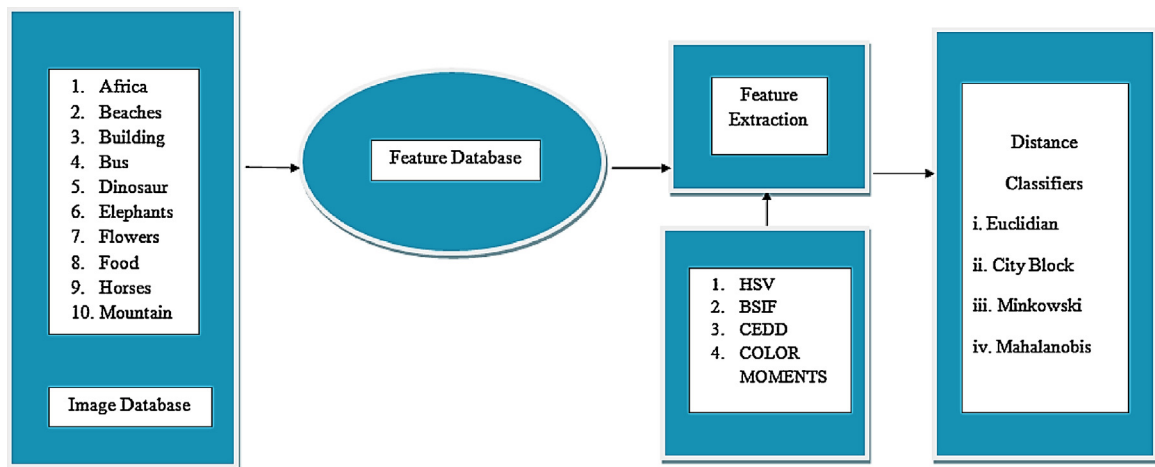


Fig. 1. Block diagram of CBIR system.

for querying the indexed image database. It is difficult to locate the relevant information from such a huge database. Content based image retrieval is the method for efficient retrieval of relevant images.

The image content can be classified into three types (a) spatial, (b) semantic and (c) low-level contents. In the spatial content of an image is the positioning of the objects from the given image. The meaning of the image is represented by the semantic content. Color, shape and texture features are considered to be the low-level contents. Most of the CBIR methods are based on low-level features. Content-based algorithm uses visual content of the image for retrieval removing the disadvantages of text-based retrieval systems.

Content Based Image Retrieval (CBIR) is the process of searching and retrieving images from a database on the basis of features that are extracted from the image themselves. In this paper image classes are used such as Africa, beaches, building, bus, dinosaur, elephants, flowers, food, horses and mountains. Features are extracted from the entire image database and the feature vectors have been stored.

Features are extracted using HSV histogram, Binarized Statistical Image Features (BSIF), Color and Edge Directivity Descriptors (CEDD) and Color moments. Feature extraction using HSV histogram includes color space conversion, color quantization and histogram computation.

Feature extraction using BSIF includes conversion of input RGB to gray image, patch selection from gray image, subtract mean value from all components, patch whitening process followed by estimation of ICA components. CEDD feature extraction process consists of HSV color two stage fuzzy linking system.

The color moments feature extraction process includes conversion of RGB into individual components and involves computation of mean and standard deviation of each component. These stored feature vectors are then compared with the feature vectors of query image. The minimum distance using distance classifiers such as Euclidean distance, City block distance, Minkowski and Mahalanobis, results in the comparison are then retrieved. The process of CBIR is shown in Fig. 1. The details of each and every feature extraction process is explained in subsequent section. Comparing to existing algorithm which consists of spatial and frequency domain features, in this article, two additional features BSIF and CEDD which improves the precision of the proposed algorithm.

The precision of image retrieval is dependent on the (1) feature extraction process, (2) feature similarity method. Some of the CBIR algorithms uses shape features extracted from the shape of an object and objects are classified with higher accuracy compared conventional features like texture and color. The color content of an image plays very important role in content-based image retrieval. Global histogram is used to represent popularly for color contents. Three RGB color channels are used to represent global color histogram. These three individual color histogram provides similarity between different images as it is scale and rotation invariant features.

In this paper hybrid features which combines three types of feature descriptors, including spatial, frequency, CEDD and BSIF features are used to develop efficient CBIR algorithm. Individual analysis of descriptors is also

studied and results are presented. The rest of the paper is organized as follows. Section 2 briefly reviews important algorithms of CBIR techniques. Various spatial, frequency and hybrid domain feature extraction methods are explained in Section 3. Section 4 presents simulation results and discussions. Finally, Section 5 concludes the paper.

## 2. Related work

Wavelet transform is introduced with the advancement in multiresolution transform research. Discrete wavelet transform is a multiresolution approach and used in CBIR. Wavelet based features were first proposed in [Mandal et al. \(1996\)](#) which selects 64 largest Haar wavelet coefficients in each of the 3 color band. Algorithm was proved ineffective as low frequency coefficients tend to be more dominant than those of the high frequency coefficients for images with sharp color changes. Daubecheis wavelet for generation of feature vector is used in [Jun-Lee et al. \(2008\)](#). Wavelet coefficients in the lowest frequency band and their variances are used as feature vector of 768 dimension.

Each sub-band local features including the high frequency bands and total energy of each sub-band are computed as feature vector leading to good frequency resolution, but no space resolution ([Kekre et al., 2010](#)). A new edge based shape feature representation method with multiresolution enhanced orthogonal polynomials model and morphological operations for effective image retrieval is presented in [Krishnamoorthy and Devi \(2013\)](#). The Pseudo Zernike moments based global shape features, which are invariant to basic geometric transformations, are extracted and are used for retrieving similar images with Canberra distance metric. CBIR system is presented using shape feature descriptor and the modified Zernike moments based on the Zernike moments with minimum geometric error and numerical integration error ([Ma et al., 2011](#)).

In [Goyal and Walia \(2014\)](#), experimental analysis of pixel-based dense descriptors such as local binary pattern (LBP), local directional pattern (LDP) and their variants are done. These descriptors are used as local features along with ZMs global features in achieving higher and accurate retrieval rate in CBIR system. A new feature scheme called enhanced Gabor wavelet correlogram (EGWC) is proposed for image indexing and retrieval. EGWC uses Gabor wavelets to decompose the image into different scales and orientations, refer [Moghaddam and Dehaji \(2013\)](#).

In [Daisy et al. \(2012\)](#), features such as shapes and texture are extracted from the query and reference images and are compared by means of Euclidean distance. The morphological operation with spatiallyvariant structuring element is used for feature extraction. Based on HSV color model, a method of object-based spatial-color feature (OSCF) for color image retrieval is proposed in [Huang et al. \(2012\)](#). Firstly, objects are extracted from color, then image features are represented by objects in it. Color and spatial-color feature are adopted for description of objects. The new method only pays attention to main central objects.

A method is proposed [Lee et al. \(2012\)](#) for efficient image retrieval that applies a weighted combination of color and texture to the wavelet transform, based on spatial-color and second order statistics, respectively. Smith and Chang used statistical characteristics (average and variance) calculated for each subband as texture features [Smith and Chang \(1994\)](#). They compared effectiveness of texture classification for the features constructed by means of the wavelet approach, homogeneous decomposition into subbands (without scaling, each subband contains a part of a signal of certain frequency), discrete cosine transform, and spatial decomposition.

In [Do and Vetterli \(2000\)](#), the mean and standard deviation of the distribution of the wavelet transform coefficients are used to construct the feature vector. In the case of transformation with  $N$  filters, the image is represented by a feature vector of dimension  $2N$ . In [Sumana et al. \(2012\)](#), authors computed a new texture feature by applying the generalized Gaussian density to the distribution of curvelet coefficients which is called curvelet GGD texture feature. The purpose was to investigate curvelet GGD texture feature and compare its retrieval performance with that of curvelet, wavelet and wavelet GGD texture features. The work consists of minimizing low-level features describing an image by using a reduced descriptor that combines color and texture information which is wavelet transformation is explored in [Gallas et al. \(2012\)](#).

A retrieval method which combines color and texture feature is based on Dual-Tree Complex Wavelet (DT-CWT) transform and Rotated Wavelet Filter (RWF) is proposed in [Wu et al. \(2010\)](#). Multiple feature fusion and matching to retrieve images from the image database is proposed in [Priya and David \(2011\)](#). Three features namely color, texture and shape are used. A modified color histogram is used to extract color features, the standard DWT method was combined with RWF features and DT-CWT are combined to select texture features and active contour model is used

to select the shape features. A CBIR method is described that represents each image in the database by a vector of feature values row mean and column mean of images for CBIR using 2D DWT [Sai and Patil \(2011\)](#).

### 3. Feature extraction

Feature extraction is the most crucial step in the content based image retrieval system as it directly affects the retrieval rate. This process involves extracting image features to a distinguishable extent for better classification. There are various techniques of feature extraction used in CBIR. We use combination of spatial and frequency domain features extracted from input images.

#### 3.1. Color auto-correlogram

CBIR algorithms use color histograms as one of the feature and it is easy to compute. As the histogram is simply the probability of any pixels of color  $c_i$  in an image, this feature does not take into account the space information of color across different areas of the image. The “Color auto-correlogram” is one of the technique which is used to integrate spatial information with color histograms.

Consider a pixel  $P_1$  of color  $C_i$  in the image  $I$ , at distance  $k$  away from  $p_1$  pick another pixel  $p_2$ , what is the probability that  $p_2$  is also of color  $C_i$ . The auto-correlogram of image  $I$  for color  $C_i$ , with distance  $k$  is defined as:

$$\gamma_{c_i}^k(I) = Pr[p_1 - p_2 = k, p_2 \in I_{c_i} | p_1 \in I_{c_i}] \quad (1)$$

From the above equation it is clear that color auto-correlogram shows the spatial autocorrelation of color changes variation with distance [Huang \(1998\)](#). The auto-correlogram integrates the color information and the space information. For each pixel in the image, the auto-correlogram method needs to go through all the neighbors of that pixel. The computation complexity is  $O(k * n^2)$ , where  $k$  is the number of neighbor pixels, which is dependent on the selection of distance. The computation complexity is higher as the distance  $k$  is large. In this paper, 64-D feature vector is extracted using this method.

#### 3.2. Color moments

Color moments is the second type of feature extracted in this experimentation. Two steps are used for this purpose: (1) In first step, input RGB image is separated into R, G and B component images (2) second step involves computation of mean (which describes the center of probability distribution of the signal) and standard deviation of each of the component. Color moments generates 6-D feature vector.

$$Mean = \frac{1}{N} \sum_{j=1}^N I_{i,j} \quad (2)$$

$$Standard\ deviation = \sqrt{\frac{1}{N} \sum_{j=1}^N (I_{i,j} - E_{r,i})^2} \quad (3)$$

#### 3.3. HSV histogram features

Feature extraction using HSV histogram includes three steps: (1) color space conversion, (2) color quantization and (3) histogram computation. In the first step input RGB image is converted into HSV color space. In the HSV color space, colors are distinguished by hue, saturation is the percentage of white light added to a pure color and value refers to the perceived light intensity [LiHua et al., 1999](#). RGB to HSV color space conversion is done using the equation:

$$H = \cos^{-1} \frac{1/2[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \quad (4)$$

$$S = 1 - \frac{3[\min(R, G, B)]}{R + G + B} \quad (5)$$

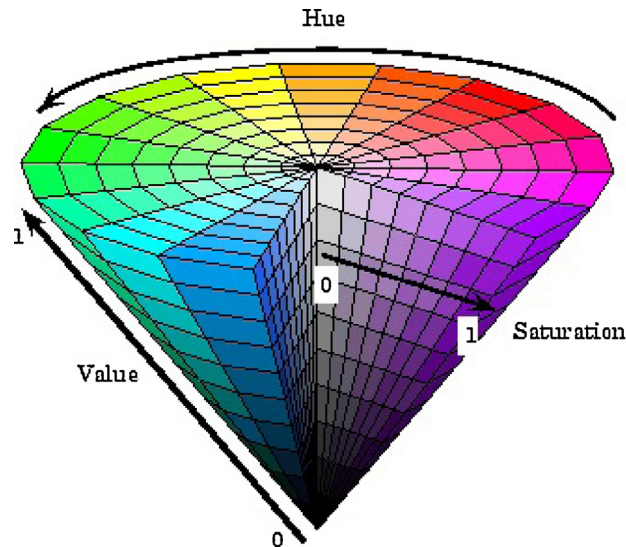


Fig. 2. Planes of HSV color space.

$$V = \left\lfloor \frac{R + G + B}{3} \right\rfloor \quad (6)$$

HSV is cylindrical geometries, with hue, their angular dimension, starting at the red primary at  $0^0$ , passing through the green primary at  $120^0$  and the blue primary at  $240^0$ , and then back to red at  $360^0$ . The HSV planes are shown as in Fig. 2.

To minimize the computational complexity and to reduce the feature size, HSV color model is quantized. Color quantization lowers the number of colors and its levels used in an image. In last step, for each quantized image a histogram is obtained which shows the frequency distribution of quantized HSV values of each pixel in an given image. In this method various intervals that are used: H: 8 bins; S: 2 bins and V: 2 bins. In the feature extraction step, we concatenate  $8 \times 2 \times 2$  histogram resulting in 32-D feature vector. We select less number of segments as the power of human eye to distinguish colors is limited (KONG, 2009) (Fig. 2).

### 3.4. Stationary Wavelet transform (SWT) moments

SWT decomposes the input image into various frequency bands which permits the isolation of the frequency into various sub-bands. 1D-SWT separates the high and the low-frequency portions of a signal by using filters. SWT decomposition results in four sub-band images called as low-low (low resolution approximation component), low-high (vertical), high-low (horizontal), and high-high (diagonal). There are various types of wavelets introduced for example Haar, Daubechies Wavelet, Morlet Wavelet and Mexican Hat Wavelet.

In this simulation, 4-level decomposition is applied to the input image that means wavelet transformation is applied iteratively 4 times. Figure below shows 4-level decomposition of input image.

Next step is to compute moments from these decomposed sub-bands. Here, mean and standard deviation are the two features extracted from each sub-band at each level. Standard deviation is used to measure variability. By using the standard deviation parameter, one can measure how much variation exists from the mean. The resulting feature vector length is 20-D (Fig. 3).

### 3.5. Gabor wavelet transform

Gabor functions provide the optimal resolution in both the time (spatial) and frequency domains. Gabor wavelet transform is the optimal basis to extract local features due to following (Shen and Bai, 2006):

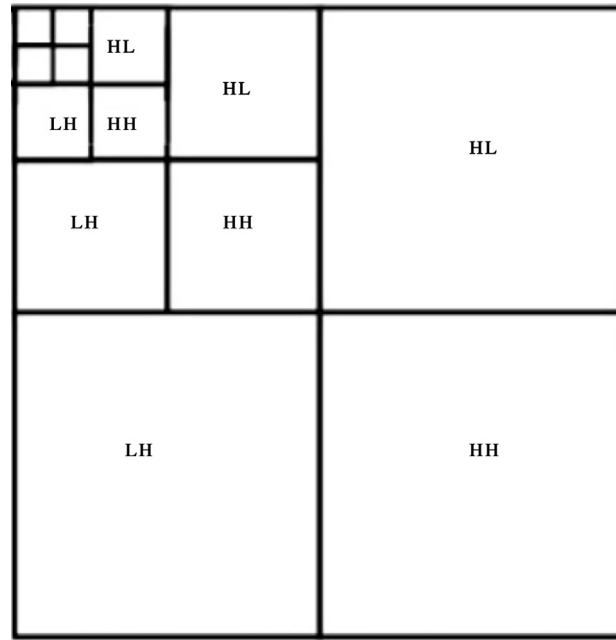


Fig. 3. 4-Level wavelet transformation of an image.

**Biological motivation:** The modeling of the higher vision visual cortex of mammalian brains is best suited by a family of self-similar 2D Gabor wavelets. **Mathematical motivation:** The multi-resolution and multi-orientation properties of the Gabor wavelet transform can be employed for measuring local spatial frequencies.

Gabor wavelet transform is used in many image analysis applications including face recognition, texture classification and facial expression classification. For a given image  $I(x, y)$  with size  $M$  times  $N$  its discrete Gabor wavelet transform is given by,

$$G_{mn}(x, y) = \sum_s \sum_t I(x - s, y - t) \psi_{mn}^*(s, t) \quad (7)$$

where  $s$  and  $t$  are the mask size variables.  $\psi_{mn}^*$  is the complex conjugate of  $\psi_{mn}$  which is a self-similar function generated from dilation and rotation of the mother wavelet. The self-similar Gabor wavelets are computed using the generating function:

$$\psi_{mn}(x, y) = a^{-m} \psi(\bar{x}, \bar{y}) \quad (8)$$

### 3.6. Binarized Statistical Image Features (BSIF)

Binarized Statistical Image Feature (BSIF) is a statistical learning based image descriptor generating a binary code string, which is generated by thresholding (Kannala and Rahtu, 2012). The independent vectors are extracted from a random samples of training image data set using independent component analysis (ICA). BSIF image features are inspired from LBP features.

In this process, the statistical filters are designed manually. Predefined texture filters are generated using natural image database. To create BSIF features, input image is convolved with these predefined textural filters. The image patches are used to learn filters using ICA. The process can be summarized as (1) Conversion of input RGB to gray image (2) Patch selection from the gray image (3) Subtract mean value from all the components (4) Patch whitening process (5) Estimation of ICA components (Hyvriinen et al., 2009).

For the input image  $f(x, y)$ , and  $l \times l$  size filter  $w_i$  the filter response is computed by convolution,

$$B_i = w_i * f(x, y) \quad (9)$$

where  $S_i$  is the response of  $i$ -th filter. Binary output is generated using the rule,  $b_i = 1$ , if  $B_i > 0$  otherwise,  $b_i = 0$ .



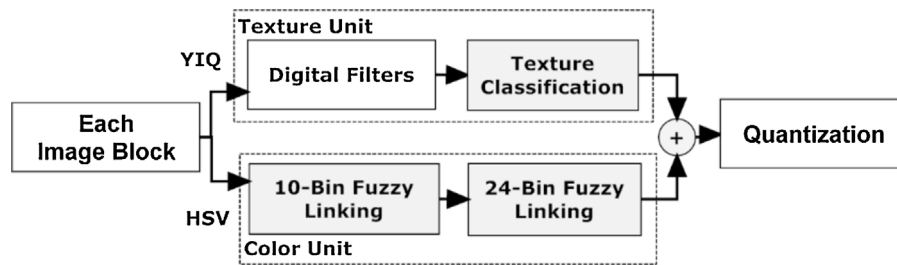


Fig. 4. CEDD descriptor extraction process.

Filter learning process involves collection of 500,000 natural image patches collection randomly. The patch size are of the sizes  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$ ,  $15 \times 15$  and  $17 \times 17$ . Higher patch results in better image retrieval. These filter patches are of different bit length ranging from 5, 6, 7, 8, 9, 10 and 11.

### 3.7. Color and Edge Directivity Descriptor (CEDD)

Color and Edge Directivity Descriptor (CEDD) uses color and texture information into single histogram bin reducing the feature size. CEDD feature extraction process consists of a HSV color two-stage fuzzy-linking system (AS, 2008). The first stage produces 10-bin histogram which represents various colors such as white, red, orange, magenta.

Second stage of the algorithm divides each of the color into one of three hues resulting in 24-bin histogram as output. Each bin corresponds different colors like dark green, light yellow, dark yellow, dark magenta etc. A set of four Takagi-Sugeno-Kang rules are used. To compute textural information, various features are extracted from high frequency bands of wavelet transform.

Final 6-bin histogram uses the five digital filters proposed by the MPEG-7, so finally generating 144 ( $24 \times 6$ ) regions. The generalized structure of the texture descriptor generation is shown in Fig. 4.

## 4. Simulation results

For evaluation of the proposed method a general purpose WANG database containing 1000 Corel images of 10 different subjects in JPEG format of size  $384 \times 256$  or  $256 \times 384$  is used. The image set comprises 100 images in each of 10 different classes as elephants, buses, flowers, horses and beaches. This database is used extensively to test many CBIR systems because the size of the database and the availability of class information.

The first step in this algorithm consists feature extraction using color auto-correlogram and color moments. Next, conversion of input RGB color space into HSV color space. After converting into HSV color space, HSV histogram features are computed. Later stage involves frequency domain feature extraction. After decomposing the input image into 4-level using DWT, wavelet transform moments are computed (mean and standard deviation) from each orientation and scale.

Gabor wavelet features are extracted at different orientation and scale of Gabor filter. In addition to this, BSIF textural descriptor features are computed using the input. We found that, filter size  $13 \times 13$  with 8 bit length resulting in higher accuracy compared to all other filter sizes. Also, statistical filter size higher than 13 was not resulting in improvement of the accuracy (retrieval results). Hence, we choose  $13 \times 13$  filter size.

In addition to this, 144 dimension CEDD features are also extracted from the input image. Final feature vector consists of all the above descriptors fusion and which is applied to similarity measurement.

### 4.1. Similarity measurements

Once the features of the input database images are created then the user can give an image as a query to retrieve similar images from the database. The feature vector of the query image is computed again by using the procedure explained above. Similarity measurement is another important issue in CBIR in which the query image is compared with other database images for the similarity. To compute the similarity between the input query image and the database

Table 1  
Precision using spatial domain features.

Class	Precision			
	Euclidean	City block	Minkowski	Mahalanobis
Africa	0.62	0.59	0.54	0.53
Beaches	0.71	0.68	0.66	0.61
Building	0.53	0.55	0.44	0.49
Bus	0.77	0.78	0.70	0.72
Dinosaur	0.89	0.88	0.82	0.79
Elephants	0.55	0.60	0.55	0.51
Flowers	0.89	0.87	0.80	0.79
Food	0.55	0.52	0.49	0.50
Horses	0.52	0.51	0.44	0.49
Mountains	0.55	0.58	0.51	0.50

images, the difference between the query image feature vector and the database image feature vector is calculated by using various distance metrics. Smaller the difference more similar the two images are. The distance metrics which are evaluated in this work are Euclidean distance, City block distance, Minkowski distance and Mahalanobis distance. In addition to this we have used  $L_1$ ,  $L_2$ , cosine and Chebyshev distance metric for experimentation.

*Euclidean distance:* Euclidean distance metric is most widely used for similarity measurement in image retrieval because of its efficiency and effectiveness (Malik and Baharudin, 2013). It measures the distance between two vectors by computing the square root of the sum of the squared absolute differences and is given as:

$$D_E = \sqrt{\sum_{i=1}^n (|I_i - D_i|)^2} \quad (10)$$

*City block distance:* City block distance metric is also called the Manhattan distance. The city block distance metric has robustness to outliers. It is computed as,

$$D_C = \sum_{i=1}^n |I_i - D_i| \quad (11)$$

*Minkowski distance:* The generalized form of the Minkowski distance can be defined as:

$$D_M = \left[ \sum_{i=1}^n (|I_i - D_i|)^p \right]^{\frac{1}{p}} \quad (12)$$

*Mahalanobis distance:* This distance metric is computed as,

$$D_{Mh} = \sqrt{(\vec{x} - \vec{\mu})^T S^{-1} (\vec{x} - \vec{\mu})} \quad (13)$$

The effectiveness of CBIR system is measured by using parameters like Precision and is defined as,

$$P(I) = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (14)$$

Extensive experiments are performed using spatial domain, transform domain and hybrid (combination of spatial and frequency domain) features and results are presented. In first case we evaluated effect of spatial domain features only on the precision of the developed CBIR system. Spatial domain features are combination of (1) auto correlogram (2) color moments and (3) HSV histogram features. Results are shown in Table 1 using various distance metrics. As it can be seen from Table 1, precision is lowest in this case. Results are better in case of euclidean and city block distance metric compared to two other distance metric.

Second experiment involves computation of precision using frequency domain features. Here, we extracted SWT moments and Gabor wavelet transform features for efficient CBIR system. Table 2 shows precision using all four



Table 2

Precision using frequency domain features.

Class	Precision			
	Euclidean	City block	Minkowski	Mahalanobis
Africa	0.72	0.69	0.62	0.68
Beaches	0.81	0.78	0.69	0.67
Building	0.63	0.62	0.54	0.59
Bus	0.87	0.88	0.79	0.78
Dinosaur	0.93	0.95	0.92	0.89
Elephants	0.65	0.69	0.61	0.59
Flowers	0.93	0.91	0.88	0.89
Food	0.65	0.62	0.59	0.58
Horses	0.62	0.61	0.54	0.59
Mountains	0.63	0.64	0.55	0.57

distance metric for frequency domain features. Compared to Table 1, precision is higher in this case. Similarly, precision is better in case of euclidean and city block distance metric compared to Minkowski and Mahalanobis distance metric.

Third experiment involves CBIR using CEDD and BSIF feature fusion. In the process, 144-D CEDD feature vector and 256-D BSIF feature fusion is employed. Table 3 shows the precision values using the same database. The retrieval performance better than spatial as well as frequency domain methods.

Table 3

Precision using CEDD and BSIF feature fusion.

Class	Precision			
	Euclidean	City block	Minkowski	Mahalanobis
Africa	0.75	0.76	0.71	0.72
Beaches	0.83	0.78	0.72	0.69
Building	0.71	0.68	0.66	0.68
Bus	0.84	0.81	0.81	0.79
Dinosaur	0.95	0.95	0.91	0.92
Elephants	0.71	0.69	0.63	0.61
Flowers	0.93	0.87	0.87	0.91
Food	0.72	0.71	0.68	0.61
Horses	0.75	0.73	0.65	0.71
Mountains	0.81	0.80	0.73	0.87

Table 4

Precision using hybrid features.

Class	Precision			
	Euclidean	City block	Minkowski	Mahalanobis
Africa	0.81	0.79	0.72	0.78
Beaches	0.92	0.88	0.82	0.77
Building	0.79	0.78	0.74	0.76
Bus	0.93	0.89	0.88	0.87
Dinosaur	0.99	0.98	0.94	0.95
Elephants	0.79	0.76	0.71	0.69
Flowers	0.99	0.97	0.93	0.95
Food	0.80	0.78	0.74	0.77
Horses	0.85	0.81	0.74	0.79
Mountains	0.88	0.87	0.83	0.85

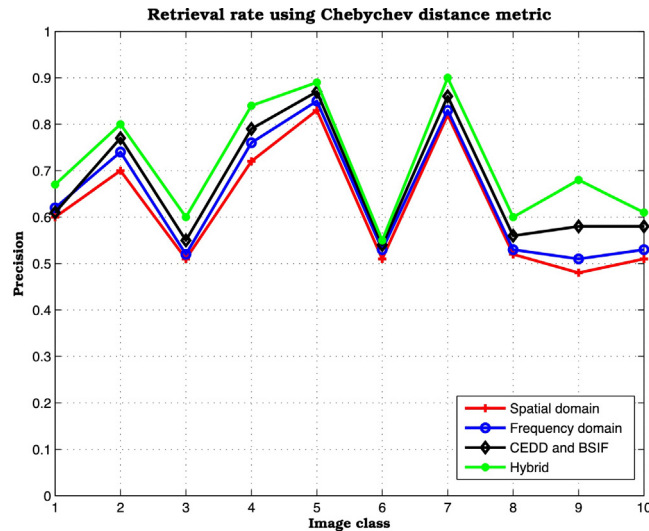


Fig. 5. Retrieval rate using Chebyshev distance metric for four types of feature descriptors.

Final set of experiment consists of precision measurement using hybrid features. In this, both spatial, frequency domain and CEDD and BSIF features are extracted and precision evaluation is done using the same database. Table 4 shows precision values using hybrid feature method. Compared to other two methods, this approach results in better precision values. Also, Euclidean distance metric performs well compared to rest of the distance metric. Additional experiments are performed in order to evaluate the performance of various distance metrics and effect of CEDD and BSIF feature extraction process. Firstly, we evaluated precision performance of the proposed CBIR algorithm using various distance metric for different feature descriptors.

Fig. 5 shows the precision obtained using Chebyshev distance metric for different class of images. We performed experiments for four feature descriptors: (1) Spatial domain features (2) Frequency domain features (3) CEDD and BSIF descriptors and (4) Hybrid features combining all the above features. In this figure, precision is plotted for the various types of image classes and feature types. We used four types of features as shown in the figure. In the Fig. 5, class 1 refers to Africa image database, class 2 shows Beaches image database, class 3 depicts Building image database, class 4 refers to Bus images, class 5 refers to Dinosaur database, class 6 means Elephants image database, class 7 refers

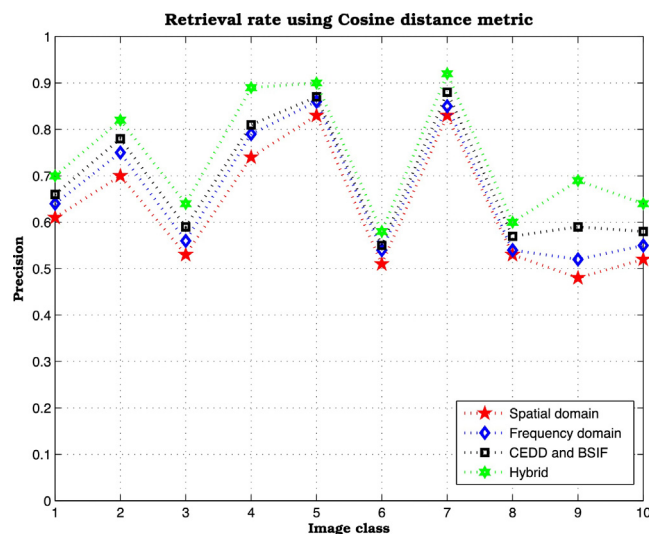


Fig. 6. Retrieval rate using Cosine distance metric for four types of feature descriptors.

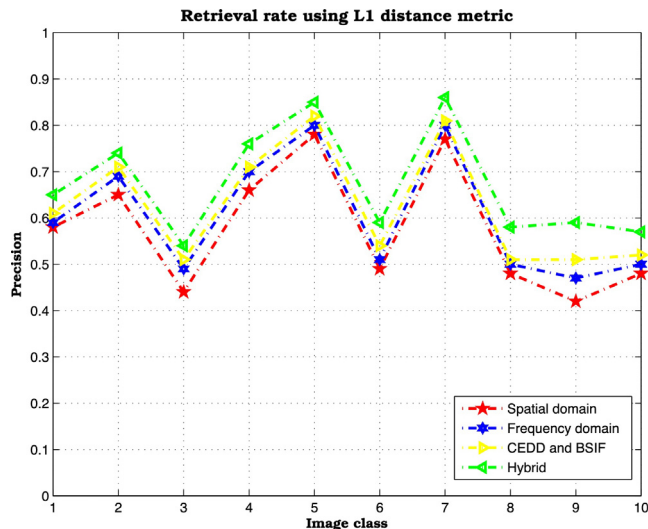


Fig. 7. Retrieval rate using  $L_1$  distance metric for four types of feature descriptors.

Flowers image database, class 8 shows Food images, class 9 refers Horses database and class 10 refers to Mountains image database. From the figure, we conclude that hybrid features performs better compared to all other types of features.

Fig. 6 shows the precision obtained using Cosine distance metric for different class of images and for four types of feature descriptors. It is evident from the figure that, the combination of CEDD and BSIF features performs better compared to rest of the feature descriptors. Also, the hybrid features precision is highest in this distance metric too.

Figs. 7 and 8 shows the precision obtained using  $L_1$  and  $L_2$  distance metric for different class of images and for four types of feature descriptors. It is observed from the figure that, compared to cosine and Chebyshev features the precision is low in the  $L_1$  distance metric. As expected, hybrid features precision is highest in both the distance metric.

CEDD and BSIF features are used in this experiment. In BSIF various predefined texture filters are generated using natural image database. Filter design in ICA involves natural random image patch collection. Different patch sizes are used to evaluate the performance of the algorithm. We have employed sizes including  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$ ,  $15 \times 15$  and  $17 \times 17$ . Effect of patch size on precision is demonstrated in Fig. 9. Figure shows

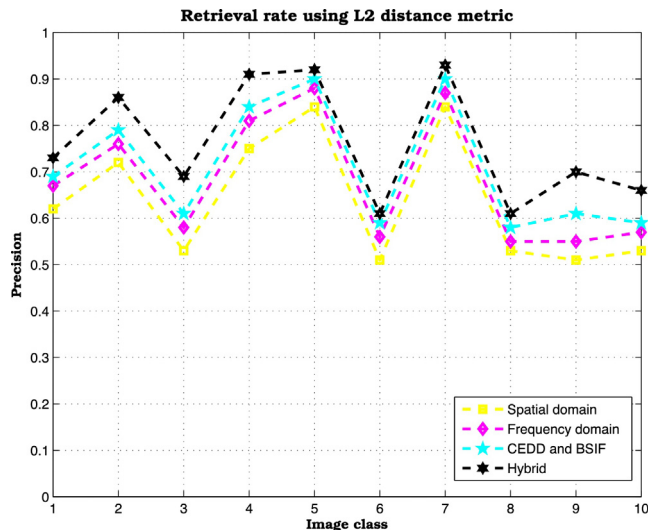


Fig. 8. Retrieval rate using  $L_2$  distance metric for four types of feature descriptors.

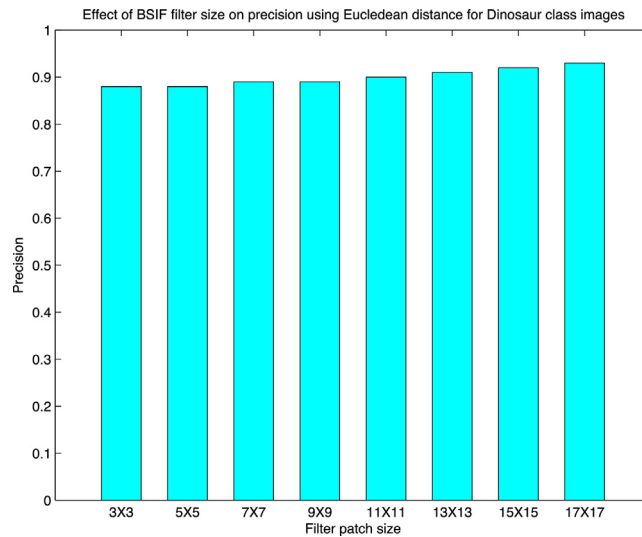


Fig. 9. Retrieval rate for Dinosaur image database using Euclidean distance metric with various patch size.

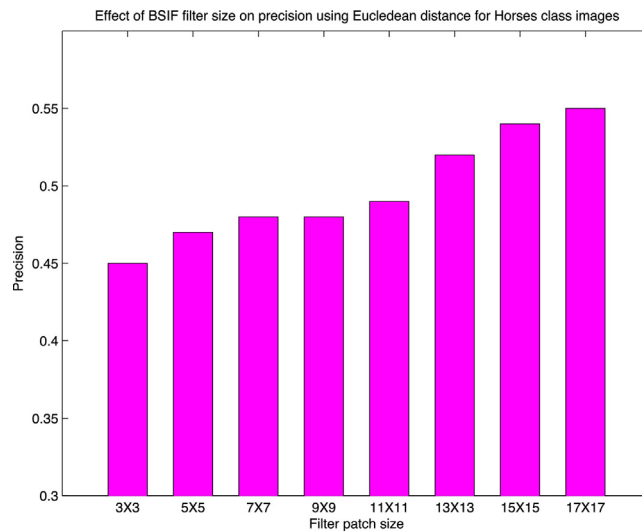


Fig. 10. Retrieval rate for Horse image database using Euclidean distance metric with various patch size.

precision values obtained using various patch size as mentioned. Evaluation is done using Euclidean distance metric for Dinosaur image database.

Similar experiments are also performed for various distance metrics. Fig. 10 depicts another such example for Horses image database. From both the figures, we conclude that as patch size increases the precision also increases. Also, there is no significant change in the result for patch size  $15 \times 15$ , and  $17 \times 17$ .

Frequency domain features includes moment based features after 4-level SWT decomposition of the image. We analyzed effect of number of decomposition levels on the precision using various distance metric and for different image classes. Fig. 11 depicts precision obtained for Dinosaur image database using euclidean distance metric. It is observed that, higher levels of decomposition results in better precision in all the image classes apart from the types of distance metric used. 4-level decomposition resulted in highest precision hence we have used 4-level SWT.

Finally, we evaluated CEDD feature histogram bins and its effect on precision. Fig. 12 shows the effect of number of histogram bins on precision using euclidean distance metric for Dinosaur class image. It is clear from the figure that, precision increases as the number of histogram bins increases.

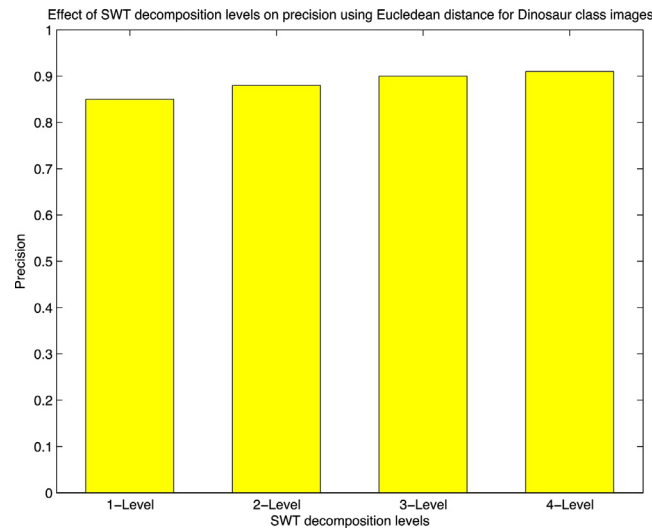


Fig. 11. Effect of number of decomposition levels on retrieval rate for Dinosaur image database using Euclidean distance metric.

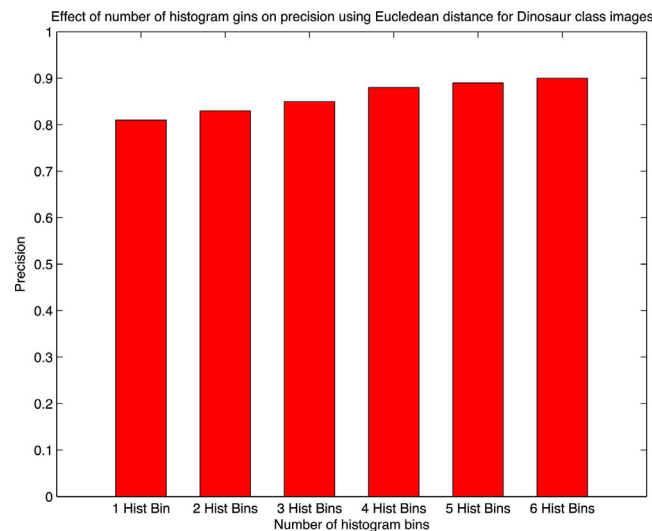


Fig. 12. Effect of number of histogram bins on retrieval rate for Dinosaur image database using Euclidean distance metric.

## 5. Conclusion

A new hybrid feature scheme is proposed for efficient CBIR in this paper based on spatial, frequency, CEDD and BSIF feature descriptors. In addition to spatial features such as color auto-correlogram, moments, and HSV histogram features and frequency domain features like SWT moments, Gabor wavelet transform features, CEDD and BSIF features are used to increase precision of the presented approach. Using spatial domain features lowest precision results were obtained whereas frequency domain features resulted in better precision as compared to spatial domain features. Finally, in order to improve the precision we extracted CEDD and BSIF descriptors, which resulted in highest precision shown in Table 5. Both global and local features are combined to obtain higher retrieval rate. Experimental results obtained using the proposed approach are better compared to existing methods. Currently we are working on techniques for reducing the feature vector size which will further reduce the execution time.

Table 5

Comparison of various CBIR methods.

	Algorithm 1 as in Youssef et al. (2012)	Algorithm 2 as in Baharudin et al. (2012)	Algorithm 3 as in Jadhav and Patil (2012)	Proposed Algorithm
Average Precision	0.703	0.81	0.525	0.875

## References

- Chatzichristofis, S.A., Boutalis, Y.S., 2008. CEDD: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In: *Proceedings of the 6th International Conference on Computer Vision Systems. ICVS'08*. Springer-Verlag, Berlin, Heidelberg, pp. 312–322. <http://dl.acm.org/citation.cfm?id=1788524.1788559>.
- Baharudin, B., et al., 2012. Effective content-based image retrieval: combination of quantized histogram texture features in the dct domain. In: *Computer & Information Science (ICCIS), 2012 International Conference on*. Vol. 1. IEEE, pp. 425–430.
- Daisy, M.M.H., TamilSelvi, S., Prinza, L., 2012. Gray scale morphological operations for image retrieval. In: *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*. IEEE, pp. 571–575.
- Do, M.N., Vetterli, M., 2000. Texture similarity measurement using kullback-leibler distance on wavelet subbands. In: *Image Processing, 2000. Proceedings. 2000 International Conference on*. Vol. 3. IEEE, pp. 730–733.
- Gallas, A., Barhoumi, W., Zagrouba, E., 2012. Image retrieval based on wavelet sub-bands and fuzzy weighted regions. In: *Communications and Information Technology (ICCIT), 2012 International Conference on*. IEEE, pp. 33–37.
- Goyal, A., Walia, E., 2014. Variants of dense descriptors and zernike moments as features for accurate shape-based image retrieval. *Signal Image Video Process.* 8 (7), 1273–1289.
- Huang, C., Han, Y., Zhang, Y., 2012. A method for object-based color image retrieval. In: *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*. IEEE, pp. 1659–1663.
- Huang, J., 1998. *Color-spatial image indexing and applications*. Cornell University, Ph.D. thesis.
- Hyvriinen, A., Hurri, J., Hoyer, P.O., 2009. *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision.*, 1st Edition. Springer Publishing Company, Incorporated.
- Jadhav, S.M., Patil, V., 2012. An effective content based image retrieval (cbir) system based on evolutionary programming (ep). In: *Advanced Communication Control and Computing Technologies (ICACCCT), 2012 IEEE International Conference on*. IEEE, pp. 310–315.
- Jun-Lee, S., Lee, Y.-H., Ahn, H., Rhee, S.B., 2008. Color image descriptor using wavelet correlogram. In: *The 23rd international conference on Circuits/systems, computers and communication*.
- Kannala, J., Rahtu, E., 2012. BSIF: binarized statistical image features. In: *Proc. 21st International Conference on Pattern Recognition (ICPR 2012)*, Tsukuba, Japan, pp. 1363–1366.
- Kekre, H., Thepade, S.D., Athawale, A., Shah, A., Verlekar, P., Shirke, S., 2010. Energy compaction and image splitting for image retrieval using kekre transform over row and column feature vectors. *Int. J. Comput. Sci. Netw. Security* 10 (1), 289–298.
- KONG, F.-H., 2009 July. Image retrieval using both color and texture features. In: *8th international conference on Machine learning and Cybernetics*.
- Krishnamoorthy, R., Devi, S.S., 2013. Image retrieval using edge based shape similarity with multiresolution enhanced orthogonal polynomials model. *Digital Signal Process.* 23 (2), 555–568.
- Lee, Y.-H., Rhee, S.-B., Kim, B., 2012. Content-based image retrieval using wavelet spatial-color and gabor normalized texture in multi-resolution database. In: *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*. IEEE, pp. 371–377.
- LiHua, C., Wei, L., GuoHui, L., 1999. Research and implementation of an image retrieval algorithm based on multiple dominant colors. *J. Comput. Res. Develop.* 36 (1), 96–100.
- Ma, Z., Zhang, G., Yan, L., 2011. Shape feature descriptor using modified zernike moments. *Pattern Anal. Appl.* 14 (1), 9–22.
- Malik, F., Baharudin, B., 2013. Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the dct domain. *J. King Saud Univ. Comput. Inform. Sci.* 25 (4), 207–218.
- Mandal, M.K., Aboulnasr, T., Panchanathan, S., 1996. Image indexing using moments and wavelets. *Consumer Electronics, IEEE Transactions on* 42 (3), 557–565.
- Moghaddam, H.A., Dehaji, M.N., 2013. Enhanced gabor wavelet correlogram feature for image indexing and retrieval. *Pattern Anal. Appl.* 16 (2), 163–177.
- Priya, R., David, V.K., 2011. Optimized content based image retrieval system based on multiple feature fusion algorithm. *Int. J. Comput. Appl.* 31 (8).
- Sai, N., Patil, R., 2011. Image retrieval using 2d dual-tree discrete wavelet transform. *Int. J. Comput. Appl.* (0975-8887) 14 (6).
- Shen, L., Bai, L., 2006. A review of gabor wavelets for face recognition. *Pattern Anal. Appl.* 9 (3), 273–292.
- Smith, J.R., Chang, S.-F., 1994. Transform features for texture classification and discrimination in large image databases. In: *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*. Vol. 3. IEEE, pp. 407–411.
- Sumana, I.J., Lu, G., Zhang, D., 2012. Comparison of curvelet and wavelet texture features for content based image retrieval. In: *Multimedia and Expo (ICME), 2012 IEEE International Conference on*. IEEE, pp. 290–295.

- Wu, J., Wei, Z., Chang, Y., 2010. Color and texture feature for content based image retrieval.
- Youssef, S.M., Mesbah, S., Mahmoud, Y.M., 2012. An efficient content-based image retrieval system integrating wavelet-based image sub-blocks with dominant colors and texture analysis. In: Information Science and Digital Content Technology (ICIDT), 2012 8th International Conference on. Vol. 3. IEEE, pp. 518–523.