International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 12 (2020) pp. 001-013

Received: 19 August, 2019; Accepted: 20 Jan, 2020; Published: 8 Feb, 2020

A Futuristic Hybrid Image Retrieval System based on an Effective Indexing Approach for Swift Image Retrieval

Shikha Bhardwaj[1], Gitanjali Pandove[2] and Pawan Kumar Dahiya[3]

[1] Department of Electronics and Communication,
DCRUST, Murthal and UIET, Kurukshetra University, Haryana 136119, India
shikpank@yahoo.com

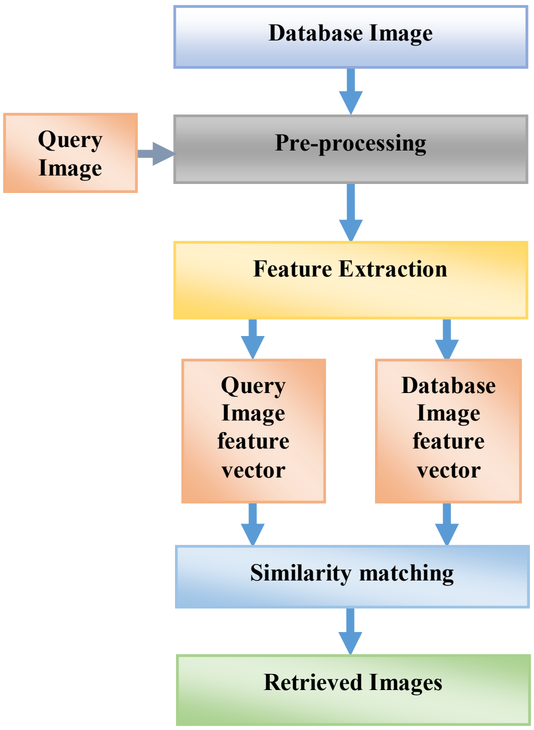
[2,3] Department of Electronics and Communication,
DCRUST, Murthal, Sonepat, Haryana, India
pandove_geet@yahoo.co.in
pawan.dahiya@gmail.com

Abstract*: As the multimedia content is increasing rapidly, there is an urgent need of an effective image retrieval system and the key to this peculiar problem is denoted by Content-based image retrieval (CBIR) system. But, for retrieving some particular images from a massive database, the retrieval process becomes time consuming. So, in order to reduce the retrieval time, image indexing is utilized. The present work highlights an effective image retrieval system with an indexing technique to reduce the retrieval time. This paper focuses on the formation of a hybrid image retrieval system in which texture, color and shape attributes of an image are withdrawn by using gray level co-occurrence matrix (GLCM), color moment and region props procedure respectively. Then, extracted fused features are optimally selected by using principal component analysis (PCA). Afterwards, two types of Indexing techniques namely, similarity-based indexing and cluster-based indexing have been tested on the developed hybrid system to find the best amongst them. The results of the hybrid color descriptor based on Cluster-based Indexing technique depict that the proposed system has enhanced results. Average precision of 93.8%, 79.6%, 70%, 98.7%, 93.5% and 79.5% has been obtained on Corel-1K, Corel-5K, Corel-10K, COIL-100, GHIM-10 and ZUBUD datasets respectively.

***Keywords*:** Similarity based indexing, cluster based indexing, Principal component analysis, Gray level co-occurrence matrix, color moment, region props

I. Introduction

In this current era, digital images and videos are prevalent in almost all the realms of technology. The increasing pace of digital images has in turn led to the development of different effective techniques for image retrieval, indexing, storage, etc.[1]. Among these techniques, an eminent technique is known by the name, content-based image retrieval (CBIR) system. This technology is based on retrieving the desired images from vast repositories based on basic image attributes like color, texture, shape, spatial information, etc. [2]. Basically, a basic image retrieval system can be disintegrated into two types: (1) Text-based image retrieval (TBIR) (2) Content-based image retrieval. In TBIR, images are retrieved on the basis of given annotations [3]. But, this technique is prone to many disadvantages like homonyms, spelling errors, etc. Therefore, the drawbacks associated with this technique can be resolved by using CBIR systems. In CBIR system, basic features like color, texture, shape, etc. are extracted from an image and a particular feature vector is formed. Then, the same procedure is applied to the complete set of database images and again feature vector is formed. Then, the two formed feature vectors are compared in similarity matching stage by using a specific distance metric and finally top N



similar images are retrieved [4] as shown in Fig. 1.

Figure 1. A basic architecture of a CBIR system

To develop a hybrid CBIR system, many types of features can be extorted depending on the requirements of the user. Texture defines the discernible patterns of an image. Different texture descriptors like discrete wavelet transform (DWT), curvelet transform [2], gabor transform, color co-occurance matrix (CCM), tamura features etc. have been used for texture extraction. Color is the basic attribute of an image and techniques like color histogram [5], color coherence vector,dominant color descriptor, color correlogram, etc, can be used for its extraction. Shape is also a prominent feature of an image which bears the semantic information and has been categorized into region based and boundary based. B-splines, curvature scale space (CSS), hough transform, fourier transform, zernike moments etc. are handful of shape descriptors used in CBIR system [3]. Different edge detection techniques like robert, prewitt, sobel, etc. can also be utilized for identification and location of keen disruption.

But, the retrieval of desired images from huge databases becomes too time consuming because each and every query image has to be matched with all the images in the database. Therefore, to reduce the time of retrieval, different techniques can be used and image indexing is one of these techniques. Indexing is a technique in which a database index of each image is maintained in a tabular format which can be used to access and find any image of that database [6]. Utilization of indexing techniques also leads to less storage requirements of the data and it can be used as a sorting technique.

The main benefactions of this paper are as follows:

- To generate a hybrid system by extracting color, texture and shape features of an image.
- To select the optimal fused features by using Principal component analysis (PCA).
- In order to reduce the retrieval time of the proposed hybrid system, two different and novel indexing techniques have been tested and analyzed.
- Finally, the execution of the proposed hybrid CBIR system based on an efficient indexing technique is evaluated by calculating precision, recall, retrieval time etc.
- Lastly, the proposed system has been evaluated on six benchmark CBIR datasets.

The remaining organization of the paper is as follows: Related work is described in Section II. Preliminaries in the form of utilized feature extraction with different indexing methods are given in section III. Section IV gives description of the proposed methodology and experimental setup and results are given in section V. Conclusion with future trends are given in Section VI.

II. Related Work

In this section, varied state-of-the-art techniques utilized for indexing have been discussed with their short-comings. Different hybrid CBIR systems have also been analyzed here. Color has been considered as an eminent feature of an image and is an obligatory source of human perception. Color histogram is the basic technique for extracting color features of an image. In addition to extract color attributes by using color histogram, an optimization technique based on K-means has been proposed by Rejito et al. [7]. But, this optimization based on K-means is dependent on the value of K. Therefore, many latest optimization techniques like monarch butterfly optimization (MBO) [8] and earthworm optimization algorithm (EWA) [9] have also been utilized.

Dimensionality reduction is the foremost advantage of indexing techniques. Basic intrinsic pattern (BIP) [10] is another indexing technique which is based on image intensity for indexing applications. But, this algorithm works well on only the texture characteristics of an image like smoothness, roughness etc. and the other features associated with color, region and shape have not been considered in this algorithm. Anti-pole tree algorithm [6] has also been used for image indexing purpose but this algorithm works only for sorted data and pre-sorting of data is not always feasible. Moreover, this technique is also less efficient in its performance.

Different types of human oriented techniques can also be utilized for the development of a user based interactive CBIR system. Interactive genetic algorithm (IGA) [11] has been deployed for fetching the most similar images based on the user's query image and to reduce the retrieval time of the developed system, K-means clustering has been used. But, in developing IGA based systems, there are many parameters which should be considered for an effective manufacturing. Small population size, interactive user-friendly system, small number of combinations and generations [12] are some of the attributes which are mandatory to be considered before the development of any IGA based image retrieval system. Therefore, selection of these parameters is a typical task.

Another indexing technique has been proposed by Megha et al. [13]. A R+-tree indexing approach has been utilized which is based on multi-dimensional analysis. This approach is generally used over the data which possesses multi-dimensional information. In this technique, generally the shape of a rectangle is chosen to represent the nearby grouped objects or images. Again, this R+-tree structure has several drawbacks. It involves recurrent insertion of objections. It is a very slow process and to satisfy the user's requirements, retrieval process becomes too complex and indeed is based on a large number of nodes.

In order to create a multi-dimensional feature matrix in a CBIR system, various descriptors like speeded up robust features (SURF), maximally stable extermal regions (MSER), color correlograms and improved color coherence vector (ICCV) have been used [14]. But, this multi-dimensionality increases the storage space for the algorithm and an efficient indexing technique is required to solve this issue. The foremost demand of any retrieval system is to capture the higher semantic information and for this purpose perceptual uniform descriptor (PUD) has been proposed by Shenglan Liu et al. [15]. This descriptor is based on a combination of gradient direction features with color attributes of an image.

Many hybrid CBIR systems have been developed so as to retrieve the required images from vast storehouse of images. A CBIR system based on three-level hierarchical system has been developed by Pradhan et al. [16]. It utilizes adaptive tetrolet transform to extract the textual information while edge joint histogram and color channel correlation histogram has been used respectively to analyze shape and color features.

But, this hybrid system does not utilize any indexing technique for the fast retrieval of images. Divergent statistics of histogram [17] has also been used for the extraction of texture features in an image retrieval system. There are many more techniques which can be used for the extraction of textural information. Curvelet transform, wavelet transform [18], local binary pattern (LBP) [19], discrete cosine transform (DCT), gabor filter are among the common texture descriptors.

Like texture, different types of global color descriptors have been used in the domain of CBIR. Color layout descriptor (CLD), edge histogram descriptor (EHD), color and edge directivity descriptor (CEDD), etc have been analyzed by lakovidou et al. [20]. In order to extract information from a specific part of an image, a region-based descriptor is required and edge-integrated minimum spanning tree (EI-MST) is an algorithm which can be used for region-based extraction [21]. Sometimes, to enhance the efficiency of the developed CBIR system, different partitions of the complete dataset has been done and clusters have been formed [22]. But, this clustering approach should be combined with an effective indexing technique for faster retrieval as well as for occupying less storage space.

In the current era, the focus of the researcher community has been shifted from machine learning to deep learning. Many CBIR systems have been developed where deep learning techniques have been utilized to extract image features. Like Arun et al. [23] describes a hybrid deep learning architecture (HDLA) which uses boltzmann machines in the upper layers and softmax model in the lower levels. Another technique based on deep convolutional neural network (DCNN) [24] has been developed which extracts the features of an image using convolution layers which subsequently employs max pooling layers for its execution. Again, the deployment of any indexing technique is missing here, which should be needed to save large number of features in proper index table with less storage requirement.

Different types of deep learning techniques like deep neural networks, restricted boltzmann machines, deep belief networks (DBN), auto encoders etc., can be used for the formation of a CBIR system [25].

In addition to images, indexing can be applied to images related to remote sensing [26], medical images [27], specifically for large medical databases namely MEDLINE [28], plant biology etc. Indexing has many advantages in the field of image processing. In order to remove redundant images from the server, again indexing technique based on locality sensitive hashing (LSH) [29] has been used.

Thus, the described hybrid CBIR systems and the indexing techniques in literature suffers from one or the other drawback. In order to overcome the above mentioned issues, a hybrid CBIR system is proposed in this paper which is based on an effective combination of gray level co-occurrence matrix (GLCM), color moment (CM) and region props procedure. To fasten the retrieval accuracy of the proposed system, an efficacious indexing technique is also utilized. This experimental analysis have been done on six benchmark CBIR systems.

III. Preliminaries

Various techniques used in the evolution of the proposed hybrid CBIR system are described in this section.

A. *Color moment*

In image retrieval systems, varied techniques have been used for the extraction of color features. Among these, color histogram [30] is the conventional method of color feature extraction. Though it is very simple and is invariant to scale and angle rotation but it does not convey any spatial information regarding an image. Color coherence vector (CCV) [31] has also been used as a color feature descriptor but the feature vector produced by this method is of high dimensionality which is against the basic requirements of any CBIR system. Dominant color descriptor (DCD) [2] also suffers from the lack of complete spatial information and moreover, if the obtained feature vector is compact, then it is utilized feasibly otherwise vague results can be produced.

Color auto-correlogram (CAC) [32] has a high computation time and cost and it is very sensitive to noise also. Therefore, based on these facts and conclusions, the color moment has been chosen as an effective color feature extraction technique. It is robust, fast, scalable, consumes less time and space. Color moments (CM) are metrics which signify color distributions in an image. According to probability theory, its distribution can be effectively characterized by its moments. Color moment can be computed for any color space model. Different moments specify diverse analytical and statistical measures. This color descriptor is also scale and rotation invariant but it includes the spatial information from images [28].

If I_{ij} specifies the i_{th} color channel and j_{th} image pixel, the number of pixels in an image are N, then index entries associated with the particular color channel and region r is given by first color moment, which signifies average color in an image denoted as:

$$Mean(E_{r,i}) = \frac{1}{N}\sum_{j=1}^{N} I_{ij} \tag{1}$$

The next moment is given by standard deviation which is obtained from a color distribution of a particular image and is defined as the square root of the variance. It is given as:

$$Standard\ deviation(\sigma_{r,i}) = \left(\frac{1}{N}\sum_{j=1}^{N}(I_{ij} - E_{r,i})^2\right)^{\frac{1}{2}} \tag{2}$$

The third moment is called as skewness and signifies that how asymmetric is the color distribution. It is given as:

$$Skewness\ (S_{r,i}) = \left(\frac{1}{N}\sum_{j=1}^{N}(I_{ij} - E_{r,i})^3\right)^{\frac{1}{3}} \tag{3}$$

Fourth moment is denoted by Kurtosis and it signifies the color distribution shape, emphasizing particularly on tall or flat shape of the color distribution.

B. *Grey level co-occurrence matrix*

To withdraw the texture features from an image, different second-order statistical methods have been utilized like: Gray level difference matrix (GLDM) [33], Gray level run length matrix (GLRLM) etc. but with the change in gray level variance, GLDM also becomes variant. On the other hand, GLRLM suffers from meagerness to represent the pattern of an image and also its computational involvement is high. For analyzing a signal in time-frequency zone, initially fourier based transforms like discrete cosine transform (DCT) and discrete fourier transform (DFT) were in trend. But, these traditional methods are deficient, as they do not convey the local information of an image. Also, the images regenerated by these techniques are of deprived standard, especially at the edges because of the presence of high- frequency bristly components.

In the wavelet domain, both gabor wavelet and discrete wavelet transform are highly prominent. But, due to a large dimension of feature vectors, Gabor wavelet takes more time in image analysis. Discrete wavelet transform (DWT) also suffers from many disadvantages like ringing near discontinuities, variance with shift, the lack of directionality of decomposition functions etc. Local binary pattern (LBP) suffers from the production of very long histograms which slows down the process of image recognition, sometimes sensitivity to noise and the effect of the center pixel is sometimes not included. Based on these conclusions, the best second-order statistical texture feature extraction technique is gray level co-occurence matrix (GLCM) [34] which has many dominant attributes like: (1) Rotation invariance (2) Diverse applications (3) Simple implementation and fast performance (4) Numerous resultant (Haralick) parameters (5) Precise inter-pixel relationship.

It is used to withdraw texture features for the retrieval and classification of different images. The spatial correlation enclosed by the pixel duplet in any given image is also computed by this geometrical method. A co-occurance matrix denoted by P_{dis} *(m,n)* specifying grey levels, contains information about two pixels: Gray level content ***m*** is denoted by the first pixel and content ***n*** is denoted precisely by the second pixel which further is separated by a distance denoted by *dis*. These specifications are chosen according to a specific angle. Matrices produced by this technique gives the gray level spatial frequencies which further gives association among pixels which are adjacent and have distinct distances amidst them [35].

Therefore, the description of GLCM is as follows:

$$P_{dis,\Theta}(m,n) = P_r(I(p_1) = m \wedge I\,(p_2) = n \wedge \|p_1 - p_2\| = d_{is} \quad (4)$$

In equation (4) *I* is the given image and the positional information in the same image *I* is given by p_1 and p_2, the probability is denoted by P and Θ denotes the range of different angle directions given by 0º, 45º, 90º and 135º. Therefore, a GLCM image is represented by d as its vector used for movement, δ as its radius and θ as orientation. A generalized GLCM matrix can be represented by Fig. 2.

Gray Tone	0	1	2	3
0	# (0,0)	# (0,1)	# (0,2)	# (0,3)
1	# (1,0)	# (1,1)	# (1,2)	# (1,3)
2	# (2,0)	# (2,1)	# (2,2)	# (2,3)
3	# (3,0)	# (3,1)	# (3,2)	# (3,3)

Figure 2. A generalized GLCM

Therefore, for a given specific test image with gray tone values, a GLCM matrix is formed which is the spatial co-occurance dependence matrix.

Prominent four types of GLCM feature parameters which are subjected to be used for the extraction of the textual content of an image and are denoted by:

$$Contrast = \textstyle\sum_m \sum_n [P(m,n) * (m-n)]^2 \quad (5)$$

The disparity between the topmost and the bottom most conterminous pixel sets is given by Contrast intensity.

$$Correlation = \frac{\sum_m \sum_n (m-u_x) * (n-u_y) * P(m,n)}{\sigma_i \sigma_j} \quad (6)$$

The correspondence between a reference pixel and its adjoining pixels in an image is diagnosed by making use of Correlation. It considers the mean and standard deviation of a matrix by encapsulating both the row and column of that particular matrix.

$$Homogeneity = \textstyle\sum_m \sum_n \frac{P(m,n)}{1+|m-n|} \quad (7)$$

In the spatial domain, the proximity among gray levels in an image is defined by the term homogeneity.

$$Energy = \sqrt{\textstyle\sum_m \sum_n P(m,n)^2} \quad (8)$$

Energy of a texture denotes the cyclic consistency of gray level allocation in an image.

C. Region props procedure

Shape features of an object provide valuable information about the identity of an object. These features can be categorized into two types: based on its region and based on its boundary. Information about an object's internal regions is depicted by region-based descriptors while boundary based shape descriptors are based on the usage of boundary information of an object. Fourier descriptors are among the popular technique of boundary-based shape descriptors. They are robust, contains perceptual characteristics but information about local features is not present because only magnitudes of the frequencies are present in Fourier transform and location information is missing.

Curvature scale space (CSS) is another shape descriptor which analyzes the boundary of an object as a 1D signal and finally represents the signal in scale space. The major issue with this technique is the superficial projections on the shape of an image. Angular radial transform (ART) produces a high dimensional feature vector which causes a hindrance in the performance of a CBIR system. Image moments, Zernike moments, Hu moments, Canny edge detector are also among the prominent shape and edge descriptors but suffers from one or the other functioning issue. Properties of image regions are efficiently measured by using region props which measures the number of connected components in an image. Many types of region props such as Center of gravity (Centroid) [36], mass, dispersion, eccentricity, axis of least inertia, hole area ratio, etc. can be used to find the shape related information in an image. In this paper, some of the region props are used to find the largest connected component. They are:

Mass: It is the total number of pixels present in one class. It is given as:

$$\textstyle\sum_{mn} h(m,n) \quad (9)$$

Where $h = \begin{Bmatrix} 1 \; if \; s(m,n) \in C \\ 0 \; if \; s(m,n) \; C \end{Bmatrix}$

Centroid: The center value of all the pixels is denoted by centroid and is also known as the center of mass. C denotes the cluster, h specifies mask over the same cluster C over image *S(m,n).* A Centroid is given by:

$$y_C = \frac{\sum_{mn} m * h(m,n)}{mass} \tag{10}$$

And $$z_C = \frac{\sum_{mn} n * h(m,n)}{mass} \tag{11}$$

In these equations (y_c, z_c) are the co-ordinates of the centroid.

Mean: It is defined as the average value of all pixels and is denoted by:

$$\mu_C = \frac{\sum_{mn} I_{mn} * h_C(m,n)}{mass} \tag{12}$$

Variance: It is the analysis which measures the distance of the spread from an average value of given random numbers.

$$\sigma_C^2 = \frac{\sum_{mn} (I_{mn} - \mu_C) * h_C(m,n)}{mass} \tag{13}$$

Dispersion: Dispersion is defined as the total distance from the centroid to every class present in an image. It is given as:

$$Dispersion = \sum_j dist(O_{D,} O_{j,D}) \tag{14}$$

Where $dist(\, O_{D,} O_{j,D})$ gives information related to the distance metric

Centroid of class *D* is O_D

Centroid of region *j* of class *D* is $O_{j,D}$

D. Principal component analysis

Use of extraneous, redundant and higher number of features can lead to the slow convergence and low performance of the designed CBIR system. Therefore, to remove this issue, the PCA technique has been utilized in this paper for dimensionality reduction. It works by zeroing the weakest principal components which leads to a projection of the data in the lower dimension. Orthogonal linear projection [37] is used to obtain these features in the lower dimension. In general PCA can be defined as:

$$Z = YC \tag{15}$$

Where $\boldsymbol{Z} \in \boldsymbol{R^{TXP}}$ is the projected data matrix. Here, P depicts the principal components of Y and P ≤ M. The ultimate goal is to find the projection matrix $\boldsymbol{C} \in \boldsymbol{R^{MXP}}$ or the singular value decomposition (SVD) technique for Y should be applied.

$$Y = U\Delta \, V^T \tag{16}$$

In equation (16) matrix of left singular values of $\boldsymbol{m \; x \; l}$ size is given by U and matrix of $\boldsymbol{n \; x \; l}$ right singular values is given by V. Diagonal matrix of singular values is given by Δ and l is the rank of matrix Y, where $\boldsymbol{l \leq min \; (m, \; n)}$. If the $\boldsymbol{m \; x \; n}$ matrix of factor scores is denoted by Fs then, it becomes: $\boldsymbol{F_s} = \boldsymbol{U}\Delta$. Finally, the product of *Y* and *V* gives the projection [35] values on the principal components by using equation (16) and is given by:

$$\boldsymbol{YV} = \boldsymbol{U}\Delta \, \boldsymbol{V^T V} = \boldsymbol{U}\Delta = \boldsymbol{F_s} \tag{17}$$

Thus, PCA can be utilized for reducing the number of features.

E. Image indexing (Cluster-based vs Similarity-based)

In order to apply image indexing on the PCA based hybrid image retrieval system, two types of indexing techniques have been analyzed and experimented namely: Cluster based and Similarity based.

Cluster based indexing approach is generally utilizes K-means clustering algorithm. K-means is a partitioning algorithm which uses a two-step iterative process to find the user specified K clusters. If the set of observations $(\boldsymbol{y_1}, \boldsymbol{y_2}, \boldsymbol{y_3} \dots \boldsymbol{y_n})$ are given then, the K-means clustering algorithm disintegrates the n observations into K parts (K≤ n) denoted by S = $(\boldsymbol{S_1}, \boldsymbol{S_2}, \boldsymbol{S_3} \dots \boldsymbol{S_k})$ in order to minimize the variance. Thus, the main objective of this algorithm [38] can be given as:

$$min_{S_k} \sum_k \sum_{y \in \boldsymbol{S_k}} \pi_y dist(y, m_k) \tag{18}$$

Here, in equation (18), S_k denotes k clusters, m_k denotes the centroid (centroid is the mean of the various cluster members) of S_k, $\pi_y > 0$ denotes the weight of y and $dist$ is the distance function used.

Similarity based indexing is a technique which is based on the usage of different similarity measures like Euclidean, manhattan, cosine, minkowski, [39] etc. In this algorithm, the fusion of three attributes i.e, color, texture and shape results in the formation of a fused feature matrix. The complete procedure of this indexing can be explained by the means of an algorithm.

1. Row-wise sorting of fused feature matrix.
2. Disintegration of obtained feature matrix into K parts.
3. Calculation of mean for each of the K obtained parts.
4. Then, mean selection for each obtained fragment.
5. For each image, nearest mean value from the fragment is chosen.
6. Finally, image index equal to the nearest fragment/part index is assigned.
7. End

To find out the nearest fragment index, divergent similarity measures have been utilized in this algorithm. Therefore, this algorithm is called as similarity based indexing.

IV. Proposed Methodology

The proposed system consists of an optimal feature selection based hybrid image retrieval system where CM, GLCM and region props procedure is used respectively to extract color, texture and shape features. This hybrid system contains fused features of the above three mentioned parameters. For the optimal feature selection, PCA is used which selects the best and precise features from the fused features. Lastly, two indexing techniques have been analyzed and tested on the proposed system in order to fasten the retrieval time of the system. Cluster-based indexing and similarity-based indexing are the two utilized approaches and amongst these cluster-based indexing achieves superior results. Therefore, the final proposed indexing-based hybrid image retrieval system has been designed by using K-means cluster-based indexing. Formally, the proposed technique can be divided into two parts: (1) Index creation stage (2) Testing/Index Prediction stage.

A. *Index creation stage*

This stage is related to the creation of indexes using cluster based indexing and further it can be divided into four sub-parts.

1) Feature extraction stage

It is the first stage of this phase and is based on the extraction of features from the utilized dataset using CM, GLCM and region props procedure to extract color, texture and shape features respectively. After this stage, three independent feature vectors are formed for each of the attribute.

2) Fusion stage

The three feature vectors are fused together in this stage for the formation of a hybrid feature vector (HFV). These three feature vectors are normalized by using minimum-maximum normalization for the creation of HFV. Normalization brings the feature dimensions into a common range.

3) Feature selection stage

The hybrid feature vector contains the fused features of all the three attributes of an image. But, to select the best features among the fused features and for dimensionality reduction, principal component analysis (PCA) has been used, which selects only 50 features for the formation of a indexing-based image retrieval system.

4) Clustering and index creation stage

The selected features obtained by using PCA have been utilized as an input to the clustering algorithm. Suitable centroid has been obtained by using K-means clustering algorithm, which is an image, based on which matching with the query is performed. K-means is a two-phase iterative process and the steps can be given by means of an algorithm.

1. Initialize the initial value of K randomly.
2. K clusters are formed by connecting every given input with the nearest mean.
3. Then, the arithmetic mean (Centroid) of each of the K clusters becomes the new mean and step 2 and 3 are repeated unless a convergence is reached and finally desired clusters are obtained.
4. End

Then, finally desired clusters are obtained and these clusters serve as indexes to different images of the dataset and index is formed for each cluster. The architecture of this part of the proposed system is shown in Fig. 3(a).

B. Index allotment/Testing stage

In this stage, the indexes which are created by using clustering approach in the previous section are utilized to retrieve an image by using a specific distance metric in spite of matching each and every image of the database. The main subsections of this stage are as follows:

1) Query formulation stage

It is the first stage of this section. The user enters the query image to test the working of the proposed system. Here, each and every image is utilized as a query image.

2) Feature extraction and Fusion stage

Again, the features of the query image are extracted using CM, GLCM and region props. Finally, a hybrid feature vector (HFV) is created which is formed by fusing the independent feature vectors of three used techniques.

3) Feature selection stage

The optimal features of the HFV are selected by using principal component analysis (PCA) to reduce the dimensions of the HFV and to select the best and optimal features from the fused feature vector.

4) Index prediction and similarity matching stage

In this stage, the matching of the query image is performed in two levels. In the index prediction phase, the matching of the query image is done with the already created cluster index and the best cluster is chosen. In the second level, by using a specific distance metric, matching of the query image is performed with images in that selected cluster and finally top N images are retrieved based on the minimal distance. The basic architecture depicting this part is given in Fig. 3(b).

V. Experimental set-up and Simulation results

To analyze the retrieval efficiency of the implemented technique, six benchmark datasets namely Corel-1K, Corel-5K, Corel-10K [40], COIL-100 [41], GHIM-10 [42] and ZUBUD [43] have been utilized. There are a diversity of images present in these databases which have been used successfully in many retrieval techniques. All the experiments are performed in MATLAB R2018a, core i3 processor, 4 GB memory, 64-bit windows. A brief description of these datasets is given in Table 1 and sample images from each dataset is shown in Fig. 4(a-f).

Image Database	Total images	Total classes	Size of images
Corel-1K	1000	10	256 × 384 or 384 × 256
Corel-5K	5000	50	256 × 384 or 384 × 256
Corel-10K	10000	100	256 × 384 or 384 × 256
COIL-100	7200	100	128 x 128
GHIM-10	10000	20	300 × 400 or 400 × 300
[illegible]	[illegible]	[illegible]	[illegible]

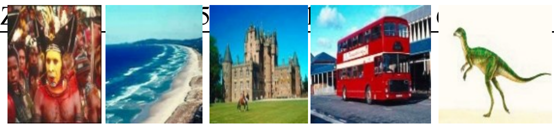


Table 1. Brief description of utilized datasets

Figure 4(a). Sample images from Corel-1K dataset

Figure 4(b). Sample images from Corel-5K dataset

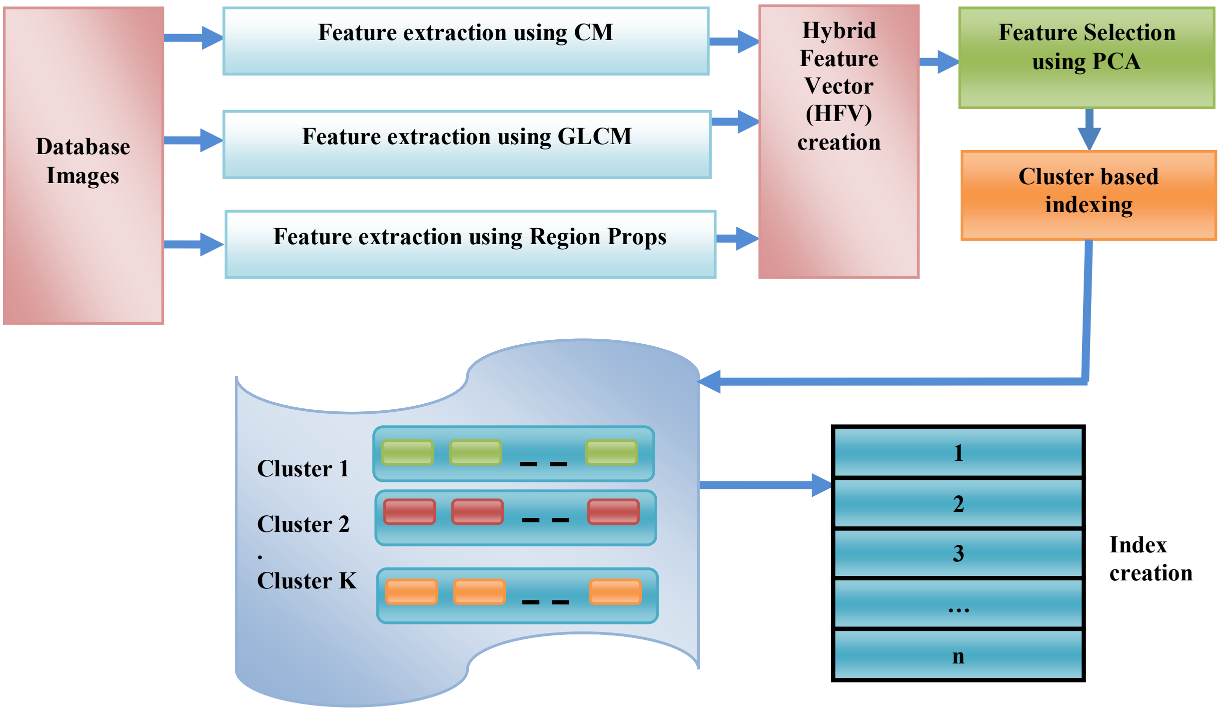


Figure 3(a). A framework of an index creation stage

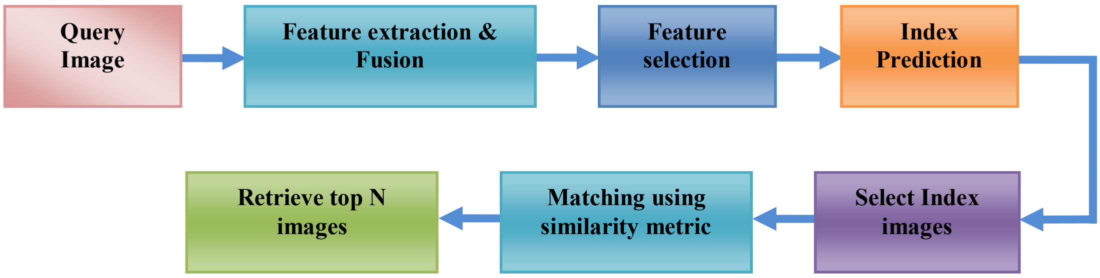


Figure 3(b). A framework of an index prediction stage

Figure 4(c). Sample images from Corel-10K dataset

Figure 4(d). Sample images from GHIM-10 dataset

A. Evaluation metrics

There are different evaluation parameters which can be utilized for calculating the capability of a particular CBIR system. But, the prominent metrics are precision and recall [44-45].

$$Precision(P_i) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \quad (19)$$

Figure 4(e). Sample images from COIL-100 dataset

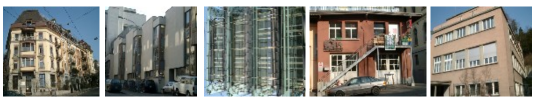


Figure 4(f). Sample images from ZUBUD dataset

$$Recall(R_i) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ in\ the\ database} \quad (20)$$

Here, the total retrieved images are 10 for every dataset except ZUBUD dataset where 5 images are retrieved because it has only 5 images per category. In case of recall, the number of relevant images of a particular dataset depends on the total number of images per category of a particular dataset. Corel-

1K, Corel-5K and Corel-10K has 100 images, GHIM-10 has 500, COIL-100 has 72 and ZUBUD has 5 images as relevant images.

B. Retrieval performance analysis

The proposed system has been obtained by extracting color, texture and shape features of an image by using CM, GLCM and Region props procedure respectively. Then, for the feature selection, PCA has been used. In order to lessen the retrieval time of the proposed technique, two divergent indexing techniques namely, similarity-based indexing and cluster-based indexing have been analyzed and tested over the proposed technique. The results in terms of average precision obtained over all the six datasets by using similarity-based indexing are given in Table 2.

Datasets	Average Precision (%) based on Similarity-based Indexing
Corel-1K	92
Corel-5K	78.8
Corel-10K	69
COIL-100	97.5
GHIM-10	92.3
ZUBUD	77.5

Table 2. Average precision (%) by using similarity-based indexing

Similarly, average precision obtained by using cluster-based indexing technique is given in Table 3.

Datasets	Average Precision (%) based on Cluster-based Indexing
Corel-1K	93.8
Corel-5K	79.6
Corel-10K	70
Coil-100	98.7
GHIM-10	93.5
ZUBUD	79.5

Table 3. Average precision (%) by using cluster-based indexing

From Table 2 and Table 3, it can be seen that the results of cluster-based indexing techniques are higher than the results of similarity-based indexing. It is due to the fact that in case of cluster-based indexing the similarity of a query image is performed only with the selected and the best cluster in place of all images in the database while in similarity-based indexing, query matching is performed with all the images in the database which is time consuming and average precision of all images in the database while in similarity-based indexing, query matching is performed with all the images in the database which is time consuming and average precision obtained is also low. Finally, indexing based on K-means clustering has been selected as the technique to be utilized over the hybrid CBIR system.

There are different distance metrics which have been utilized for calculating the similarity between a query image and complete database images. Both the indexing techniques have been evaluated with varied distance metrics. The results obtained by using varied distance metrics on both the indexing techniques are given in Table 4 and Table 5. Some of the prominent distance metrics are:

$$Distance_{Euclidean} = \sqrt{\sum_{j=1}^{n} \left|(I_j - D_j)^2\right|} \quad (21)$$

$$Distance_{City\ block} = \sum_{j=1}^{n} \left|I_j - D_j\right| \quad (22)$$

$$Distance_{Minkowski} = \left[\sum_{j=1}^{n} \left|(I_j - D_j)^{1/P}\right|\right] \quad (23$$

Datasets	Distance Metrics on Similarity-based indexing			
	Euclidean	City block	Minkowski	Cosine
Corel-1K	91.6	92.0	91.2	90.3
Corel-5K	77.4	78.8	76.5	76.0
Corel-10K	69.0	69.0	68.6	67.1
Coil-100	96.2	97.5	95.7	95.2
GHIM-10	91.3	92.3	90.6	88.4
ZUBUD	76.2	77.5	76.12	75.0

Table 4. Average precision (%) by using various distance metrics on similarity-based indexing

Datasets	Distance Metrics on Cluster-based indexing			
	Seuclidean	City block	Correlation	Cosine
Corel-1K	92.0	93.8	90.1	91.5
Corel-5K	78.5	79.6	76.8	77.2
Corel-10K	68.7	70	64.2	66.4
Coil-100	96.2	98.7	93.7	95.6
GHIM-10	92.1	93.5	89.5	91
ZUBUD	77.68	79.5	74.23	76.86

Table 5. Average precision (%) by using various distance metrics on cluster-based indexing

From Table 4, it can be concluded that in case of similarity-based indexing, city-block distance metric has superior results as compared to other similarity measures. Similarly, in cluster-based indexing again city-block distance obtains the highest results. Generally, in case of indexing based on K-means clustering, squared euclidean distance metric has been used but the proposed system has been realized in such a way, that other distance metrics can also be tested.

Among, these distance metrics again, City block obtains augmented results because it is based on absolute value results in comparison to squared value results. Moreover, it is fast in computation.

After applying the K-means cluster based indexing the whole dataset has been divided into K clusters and mean of each cluster is calculated which forms the centroid of each cluster. Then, each query image is matched with the total images of the best cluster based on the minimum value of distance metric and thus indexing is performed. The number of images for each cluster for Corel-1K dataset is given in Table 6.

Cluster (K=10)	Record count (No. of images/cluster)
1	132
2	131
3	97
4	144
5	26
6	103
7	42
8	145
9	102
10	78

Table 6. Number of clusters present in Corel-1K dataset

From Table 6, it can be seen that for Corel-1K dataset, the whole dataset has been divided into 10 clusters (K=10) and the record count represents the number of images present in each cluster. Then, if a query image is given by the user, its best matching cluster is obtained based on the minimum distance and finally, only the images present in that particular cluster are matched with the given query image to retrieve the top N images.

Recall is also considered as an eminent measure to represent the retrieval efficiency of any CBIR system. Together, precision vs recall curves proves to be very effective in calculating and analyzing the performance of any image retrieval system. Precision vs recall curves by varying the number of images from 10 to 50 for the utilized datasets are shown in Fig. 5(a) and 5(b).

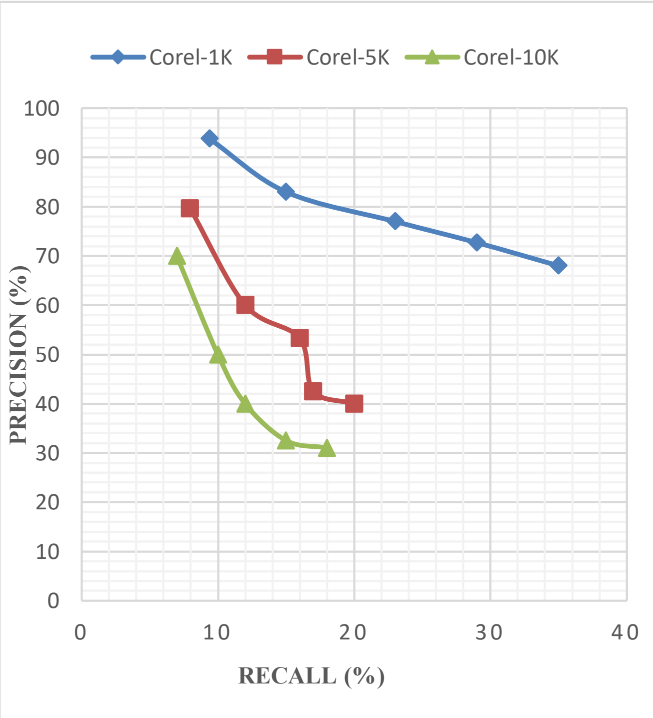


Figure 5(a). P vs R curves on Corel-1K, Corel-5K and Corel-10K

Thus, from the curves given in Fig. 5(a) and 5(b), it can be seen that as there is a decrease in the value of precision, there is an increase in the value of recall.

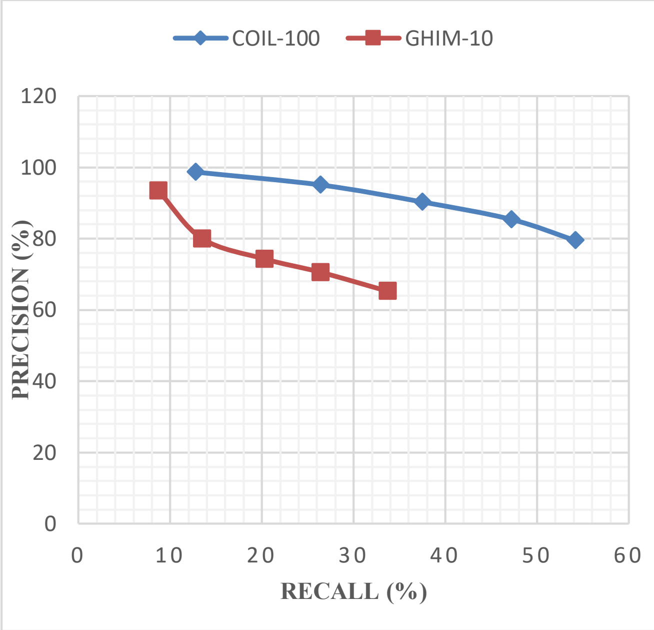


Figure 5(b). P vs R curves on COIL-100 and GHIM-10

The comparative analysis based on COIL, GHIM and ZUBUD datasets is given in Table 7, 8 and 9.

Dataset	Average Precision (%)				
COIL-100	Hybrid only	Ref. [31]	Ref. [12]	Ref. [13]	Proposed (Hybrid + Indexing)
	94.4	95	82.5	96.5	98.7

Table 7. Comparative analysis based on COIL-100 dataset

Dataset	Average Precision (%)				
GHIM-10	Hybrid only	Ref. [17]	Ref. [21]	Ref. [43]	Proposed (Hybrid + Indexing)
	85.7	76.99	73.7	57.51	93.5

Table 8. Comparative analysis based on GHIM-10 dataset

Dataset	Average Precision (%)				
ZUBUD	Hybrid only	Ref. [19]	Ref. [18]	Ref. [20]	Proposed (Hybrid + Indexing)
	76.2	79.12	75.85	70	79.5

Table 9. Comparative analysis based on ZUBUD dataset

The proposed PCA based hybrid color descriptor with cluster-based indexing technique has been compared with many state-of-the-art techniques in order to validate the performance of the proposed system. The comparative analysis based on Corel-1K, Corel-5K and Corel-10K is given in Table 10.

Dataset	Semantic name	Average Precision (%)						
		Hybrid only	Ref. [7]	Ref. [14]	Ref. [33]	Ref. [2]	Ref. [37]	Proposed (Hybrid + Indexing)
Corel-1K	Africa	82	92	95	68	72.4	81	86.5
	Beach	94	42	60	65	51.5	92	99
	Building	82	38	55	80	59.55	79	80.5
	Bus	95	52	100	90	92.35	93	88
	Dinosaur	100	100	100	100	99	99	100
	Elephant	83	68	90	80	72.7	79	100
	Flower	100	92	100	85	92.25	99	98
	Food	90	78	100	85	72.35	88	100
	Horse	82	97	100	95	96.6	80	99
	Mountain	87	25	75	75	55.75	85	87
	Average	88.8	68	87.5	82.3	76.5	87.5	93.8
Corel-5K	**Average**	77.3	55.75	72.56	64.5	55.25	75.4	79.6
Corel-10K	**Average**	68.4	48.92	65.45	57.54	49.58	68.9	70

Table 10. Comparative analysis based on Corel-1K, Corel-5K and Corel-10K dataset

The related techniques given in Table 10 lack in performance due to several issues. Firstly, the described systems did not include any indexing technique in their working to reduce the retrieval time. Moreover, the presented proposed technique in this paper is based on a hybrid CBIR system which can be used to retrieve complex and irregular images as well.
From the comparative analysis given in Tables 7, 8, 9 and 10, it can be clearly seen that the proposed technique has superior results as compared to the latest and famous techniques based on different datasets. Individual Graphical user interface (GUI) has been designed for each dataset based on the top N retrieval. The GUI's for each dataset have been given in Fig. 6(a-f).

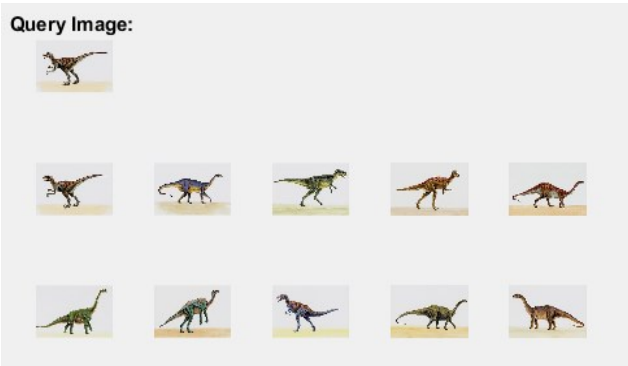


Figure 6(a). Retrieval results from Corel-1K dataset

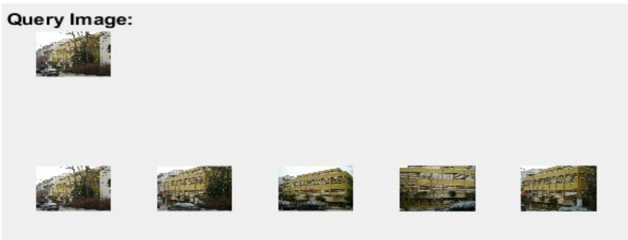


Figure 6(b). Retrieval results from ZUBUD dataset

Figure 6(c). Retrieval results from Corel-5K dataset

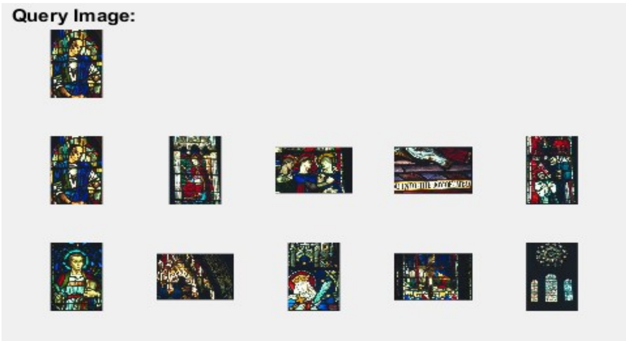


Figure 6(d). Retrieval results from Corel-10K dataset

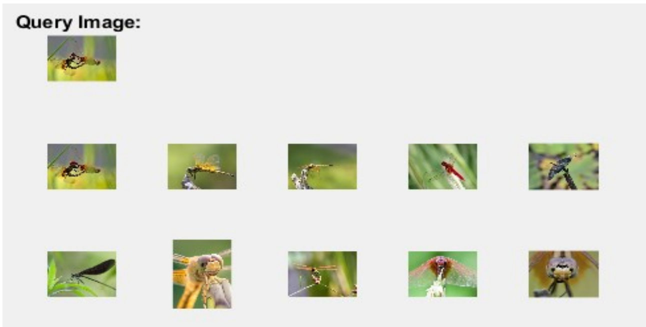


Figure 6(e). Retrieval results from GHIM-10 dataset

Figure 6(f). Retrieval results from COIL-100 dataset

From Fig 6(a-f), it can be concluded that the top ten images are retrieved from the desired category of the dataset which in-turn belongs to the native category of the query image for all the datasets except ZUBUD dataset where only 5 images are retrieved. This is so because it has only 5 images per category.

C. Time performance analysis

Since, the proposed hybrid descriptor is based on indexing, therefore the retrieval time is indeed an important aspect to be considered. This section describes the retrieval time of the proposed technique. The comparative analysis of the retrieval time based on hybrid system only (without indexing) with indexing based hybrid system has been given in Table 11.

Datasets	Avg. Precision on Hybrid system only	Retrieval time in (sec)	Avg. Precision on Hybrid system with indexing	Retrieval time in (sec)
Corel-1K	88.8	1.25	93.8	0.112
Corel-5K	77.3	1.59	79.6	0.65
Corel-10K	68.4	2.620	70	1.54
COIL-100	94.4	2.320	98.7	1.010
GHIM-10	85.7	2.453	93.5	1.51
ZUBUD	76.2	1.435	79.5	0.27

Table 11. Retrieval time comparison between two levels of the proposed system

From Table 11, it can be easily seen that the proposed system based on indexing has lesser retrieval time as compared to the retrieval time of hybrid system only. In Table 11, the average precision is given in terms of percentage (%). The comparison of the retrieval time of the proposed system with the related techniques is given in Fig. 7.
From Fig. 7, it can be deduced that the proposed system is effective in reducing the time required for the retrieval of images, thus has a fast execution.

VI. Conclusion with future trends

The proposed system has been designed by using a combination of color, texture and shape descriptors. Then,

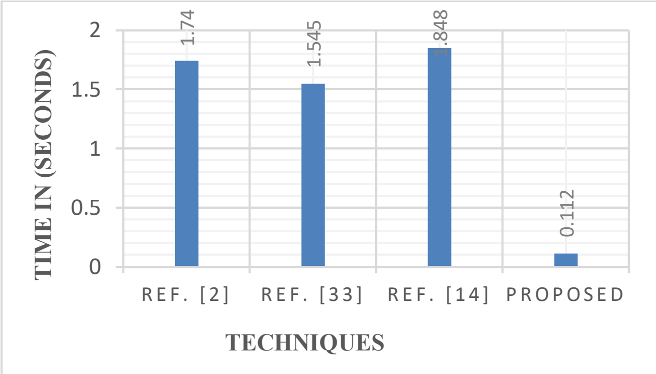


Figure 7. Comparison of the retrieval time with the related techniques on Corel-1K dataset

PCA has been used for best feature selection. Finally, cluster-based indexing technique has been used for the faster retrieval of images. To validate the results of the proposed system, the obtained results have been compared to many state-of-the-art techniques in terms of average precision. The retrieval time has also been compared to many related techniques in order to satisfy the obtained results. In future, our work will be focused on developing a deep learning based hybrid CBIR system. Also, for the purpose of optimization, different optimization algorithms like monarch butterfly optimization (MBO), earthworm optimization algorithm (EWA), elephant herding optimization (EHO), moth search (MS) algorithm, etc. could be utilized. In order to further enhance the indexing approach, some preference learning or structured query language (SQL) based indexing technique could be used.

References

[1] C. Reta, J. A. Cantoral-ceballos, I. Solis-moreno, and J. A. Gonzalez, “Journal of Visual Communication and Color uniformity descriptor : An e fficient contextual color representation for image indexing and retrieval,” *J. Vis. Commun. Image Represent.*, vol. 54, no. March, pp. 39–50, 2018.

[2] S. Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, “New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features,” *IET Image Process.*, Vol. 11, pp. 89–98, 2017.

[3] A. Alzu’bi, A. Amira, and N. Ramzan, “Semantic content-based image retrieval: A comprehensive study,” *J. Vis. Commun. Image Represent.*, Vol. 32, pp. 20-54, 2015.

[4] S. Bhardwaj, G. pandove and P. K. Dahiya, “A Journey from basic Image Features to Lofty Human Intelligence in Content Based Image Retrieval : Motivation and Applications.” *Proc. in 5th Int. Conf. on Computing for Global Sustainalble development,* Delhi, 14-16 March, 2018, pp. 1-6.

[5] G. H. Liu and J. Y. Yang, “Content-based image retrieval using color difference histogram,” *Pattern Recognit.*, vol. 46, no. 1, pp. 188–198, 2013.

[6] A. N. Bhute and B. B. Meshram, "Content Based

Image Indexing and Retrieval," *Int. J. of Graphics & Image Process,* Vol. 3, no.4, pp. 235-246, 2013.

[7] J. Rejito, A. S. Abdullah, and D. Setiana, "Image indexing using color histogram and k- means clustering for optimization CBIR in image database," *J. of physics*, series 893, pp. 1-8, 2017.

[8] Y. Feng, X. Yu and G.G. Wang, "A novel monarch butterfly optimization with global position updating operator for large-scale 0-1 knapsack problems", *Mathematics,* Vol. 7, no. 11, pp.1056, 2019.

[9] G. G. Wang, S. Deb and L.D.S. Coelho, "Earthworm optimization algorithm: a bio-inspired metaheuristic algorithm for global optimization problems", *International Journal of Bio-Inspired Computation*. 2018, 12(1):1-22. doi: 10.1504/I J B IC.2015.1 00 04283

[10] R. I. Minu, G. Nagarajan, P. J. T, and A. Pravin, "BIP : A dimensionality reduction for image indexing," *ICT Express*, no. xxxx, pp. 1–5, 2018.

[11] V. N. Pahune, R. Pusdekar, and N. Umare, "Human Oriented Content Based Image Retrieval Using Clustering and Interactive Genetic Algorithm-A Survey," *Int. J. of Science Tech. & Engg*., Vol. 3, no. 5, pp. 63–68, 2016.

[12] H. Farooq and M. Tariq, "A Comparative Study on User Interfaces of Interactive Genetic Algorithm," *Procedia - Procedia Comput. Sci.*, vol. 32, pp. 45–52, 2014.

[13] M. M. Shah, "Multi-Dimensional Image Indexing with R + -Tree," *Int. J. of Innovations & Advancement in Computer Science*, vol. 3, no. 1, pp. 38–42, 2014.

[14] H. A. Elnemr, "Combining SURF and MSER along with Color Features for Image Retrieval System Based on Bag of Visual Words," *J. of Comp. sciences*, pp. 213-222, 2016.

[15] S. Liu *et al.*, "Perceptual uniform descriptor and Ranking on manifold : A bridge between image representation and ranking for image retrieval," *J. of Latex*, arXiv:1609.07615v1 [cs.CV] 24 Sep, pp. 1-14, 2016.

[16] J. Pradhan, S. Kumar, A. K. Pal, and H. Banka, "A hierarchical CBIR framework using adaptive tetrolet transform and novel histograms from color and shape features," *Digit. Signal Process. A Rev. J.*, vol. 82, pp. 258–281, 2018.

[17] I. Shakir, "A New Approach for Content Based Image Retrieval Using Statistical Metrics," *J. of Adv. Research in Dynamical & Control Systems*, Vol. 10, 13-Special Issue, pp. 2349-2359, 2018.

[18] S. Anand, S. Jaiswal, and P. K. Ghosh, "A Machine Learning Prediction Model for Automated Brain Abnormalities Detection," *Recent Patents on Comp. Sci*., Vol. 11, pp. 17-24, 2018.

[19] P. Srivastava and A. Khare, "Utilizing multiscale local binary pattern for content-based image retrieval," *Mutimedia Tools and App*., DOI 10.1007/s11042-017-4894-4, pp. 1-27, 2017.

[20] C. Iakovidou, N. Anagnostopoulos, A. Kapoutsis, Y. Boutalis, M. Lux, and S. A. Chatzichristofis, "Localizing global descriptors for content-based image retrieval," *EURASIP J. on Advances in Signal Proces.,* 10.1186/s13634-015-0262-6, pp.1-20, 2015.

[21] Y. Liu, L. Huang, S. Wang, X. Liu, and B. Lang, "Efficient Segmentation for Region-based Image Retrieval Using Edge Integrated Minimum Spanning Tree." *Proc. in 23rd International Conference on Pattern Recognition (ICPR)*, 4-8 Dec. ancun, Mexico, 2016, pp. 1-6.

[22] S. Pandey, P. Khanna, and H. Yokota, "Clustering of hierarchical image database to reduce inter-and intra-semantic gaps in visual space for finding specific image semantics," *J. Vis. Commun. Image Represent.*, vol. 38, pp. 704–720, 2016.

[23] K. S. Arun and V. K. Govindan, "A Hybrid Deep Learning Architecture for Latent Topic-based Image Retrieval," *Data Sci. Eng*., vol. 3, no. 2, pp. 166–195, 2018.

[24] K. Lin, H.F.Yang, and C.S. Chen, "Flower Classification with Few Training Examples via Recalling Visual Patterns from Deep CNN", Proc. In *International Conference on Computer vision, graphics and image process*. (CVGIP), 2015, pp. 1-9.

[25] L. Weibo, W. Zidong, L. Xiaohui, Z. Nianyin, L. Yurong, and E.A. Fuad, "A survey of deep neural network architectures and their applications", *Neurocomputing*, vol. 234, pp. 11-26, 2017.

[26] P. Sadeghi-tehran, P. Angelov, N. Virlet, and M. J. Hawkesford, "Scalable Database Indexing and Fast Image Retrieval Based on Deep Learning and Hierarchically Nested Structure Applied to Remote Sensing and Plant Biology," *J. of Imaging*, vol. 5, no. 33, pp. 1-21, 2019.

[27] Y. Kumar, A. Aggarwal, S. Tiwari, and K. Singh, "Biomedical Signal Processing and Control An efficient and robust approach for biomedical image retrieval using Zernike moments," *Biomed. Signal Process. Control*, vol. 39, pp. 459–473, 2018.

[28] Y. Wang, L. Huang, S. Guo, L. Gong, and T. Bai, "A novel MEDLINE topic indexing method using image presentation q," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 130–137, 2019.

[29] S. Thaiyalnayaki, J. Sasikala, and R. Ponraj, "Indexing near-duplicate images in web search," *Mater. Today Proc.*, vol. 5, no. 1, pp. 1943–1949, 2018.

[30] S. M. Singh and K. Hemachandran, "Content - Based Image Retrieval using Color Moment and Gabor Texture Feature," *Int. J. of Comp. Science Issues*, vol. 9, no. 5, pp. 299–309, 2012.

[31] R. Chaudhari and A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 1, no. 5, pp. 386-392, 2012.

[32] V. Vinayak, "CBIR System using Color Moment and Color Auto-Correlogram with Block Truncation Coding," Int. J. of Comp. App., vol. 161, no. 9, pp. 1-7, 2017.

[33] L. Nanni, S. Brahnam, S. Ghidoni, E. Menegatti, and T. Barrier, "Different Approaches for Extracting Information from the Co-Occurrence Matrix," *PLOS one*, vol. 8, no. 12, pp. 1-9, 2013.

[34] S. Bhardwaj, G. Pandove, and P. K. Dahiya, "An Intelligent Multi-resolution and Co-occuring local

pattern generator for Image Retrieval," *EAI Endorsed Transactions on Scalable Information Systems,* sis: 18, pp. 1–12, 2019.

[35] E. Mehdi, E. Aroussi, and N. El Houssif, "Content-Based Image Retrieval Approach Using Color and Texture Applied to Two Databases (Coil-100 and Wang)", *Springer International Publishing, https:// doi.org/10.1007/ 978-3-319-76357-6_5*, pp. 49-59, 2018.

[36] F. Song, Z. Guo, and D. Mei, "Feature selection using principal component analysis," *Proc. in Int. Conf. on System Science, Engineering Design and Manufacturing Informatization Feature,* pp. 27-30, 2010.

[37] M. Morchid, R. Dufour, P. Bousquet, and G. Linarès, "Feature selection using Principal Component Analysis for massive retweet detection," *Pattern Recognit. Lett.*, vol. 49, pp. 33–39, 2014.

[38] J. Cao, Z. Wu, J. Wu, and W. Liu, "Towards information-theoretic K-means clustering for image indexing," *Signal Processing*, vol. 93, no. 7, pp. 2026–2037, 2013.

[39] Y. Mistry, D. T. Ingole, and M. D. Ingole, "Content based image retrieval using hybrid features and various distance metric," *J. Electr. Syst. Inf. Technol.*, no. 2016, pp. 1–15, 2017.

[40] http://wang.ist.psu.edu/docs/related/

[41] https://www.kaggle.com/jessicali9530/coil100

[42] http://www.ci.gxnu.edu.cn/cbir/

[43] http: //www .vision. ee. ethz. ch /showroom /zubud/

[44] M. Ennekhli, R. El Ouazzani and M. El Haziti "A New Content-Based Image Retrieval System Using Parzen Relevance Feedback and Kullback-Leibler Divergence" *Int. J. of Computer Information Systems and Industrial Management Applications,* Vol. 10, pp. 18-27, 2018.

[45] G. Liu, "Content-Based Image Retrieval Based On Visual Attention And The Conditional Probability," *In proc. of Int. Conf. on Chemical, Material and Food Engg.*, pp. 838–842, 2015.

Author Biographies

Shikha Bhardwaj received her B.Tech degree in ECE from Kurukshetra University, Haryana, India in 2006, Master's degree in ECE from M.M University, Ambala, Haryana, India in 2009. Now, pursuing Ph.D in Electronics and Communication Engineering Department from DCRUST University, Murthal, Haryana, India. Since 2011, she is currently working as an Assistant Professor in University Institute of Engineering & Technology, Kurukshetra University, Haryana, India. Her major research areas are Image processing, Artificial intelligence, Machine learning

Gitanjali Pandove received her Ph.D degree in ECE. She is working as Professor in ECE Deptt., in DCRUST, Murthal, Haryana, India. Her research areas are optical communication, digital image processing, power systems, etc. She has more than 70 publications in various international reputed journals.

Pawan Kumar Dahiya received his Ph.D degree in ECE. He is working as a Professor in ECE Deptt., in DCRUST, Murthal, Haryana, India. His research areas are Digital and Embedded System Design, ANPR, VANETs, Soft Computing Techniques, Image Processing, Internet of Things (IoT), etc. He has more than 75 publications in various national and international reputed journals.