CrossMark

Categorical image retrieval through genetically optimized support vector machines (GOSVM) and hybrid texture features

Aun Irtaza · M. Arfan Jaffar

Abstract Content-based image retrieval (CBIR) systems provide a potential solutions of retrieving semantically similar images from large image repositories against any query image. The research community is competing for more efficient and effective methods of content-based image retrieval, so they can be employed in serving time critical applications in scientific and industrial domains. In this paper, we have combined genetic algorithm and support vector machines to reduce the existing gap between high-level semantic content of the images and the information provided by their low-level descriptors. To maximize the performance of proposed technique, an efficient feature extraction method is introduced, which is based on the concept of in-depth texture analysis. To further enhance the capabilities of proposed method, we employed a way through which the risk of mis-associations can be avoided. To justify the effectiveness of the proposed method, we compared it against several popular CBIR techniques and show a significant improvement in terms of accuracy and stability based on Corel image gallery.

Keywords CBIR · Genetic algorithm · Relevance feedback · SVM-based semantic association

A. Irtaza · M. A. Jaffar
National University of Computer and Emerging Sciences,
Islamabad, Pakistan

A. Irtaza
Department of Computer Science, University of Engineering
and Technology, Taxila, Pakistan
e-mail: aun.irtaza@gmail.com

M. A. Jaffar (✉)
College of Computer and Information Sciences,
Al Imam Mohammad Ibn Saud Islamic University (IMSIU),
Riyadh, Saudi Arabia
e-mail: arfan.jaffar@gmail.com; arfan.jaffar@ccis.imamu.edu.sa

1 Introduction

Due to the exponential growth of digital image records, many industrial, medical, educational, and scientific applications [1] demand efficient retrieval of images based on the actual image contents. This is the reason that content-based image retrieval has gained much attention in last decade [2]. Content-based image retrieval (CBIR) systems provide a potential solution of retrieving semantically similar images from large image repositories as that of the query image. A lot of research has been taken place in the area of CBIR, but still CBIR is not a mature research area as it is suffered from many inherent problems such as semantics evaluation, results formulation, semantic gap between low-level features and high-level visual concepts, and the way to deal with enormously large image repositories.

Content-based image retrieval (CBIR) systems represent images through low-level visual features, and then use any dissimilarity metric to return the results, which appear close to that representation as their response. There are three main flaws in this approach due to which the retrieval output is not satisfactory: (1) CBIR systems rely only on image features; they rank the retrieved images on the basis of the feature distance with query image and generate the output; in this regard, they do not verify their output. The problem with this approach is that many images may appear as the response images while they are not relevant at all. (2) Secondly, they do not consider similarity among neighbors of the query image for output finalization; therefore, the generated output is very inconsistent, as the probability of right semantic association on the basis of single image (query image) is far less than the right semantic association probability of multiple images (top neighbors) [3]. (3) Wrongly associated outputs are usually compromised in case of complex images; for example, when human perception subjectivity is not incor-

porated in the system in the form of relevance feedback, user is not able to evaluate the results and guide the CBIR system, so system output contains mostly wrong results. If somehow we become able to overcome these shortcomings, we can increase the image retrieval performance of CBIR systems.

Keeping the aforementioned points in mind, in this paper, the focus of research was on finding the ways through which such kind of shortcomings can be avoided and performance of CBIR systems can be improved. For this, we emphasized on following points in this paper:

– For content-based image retrieval and semantic association purposes, a multiple support-vector-machine-based architecture is introduced, which is empowered by genetic algorithm (GA).
– To raise the right semantic association probability and bring the output verification, a semantic association scheme is introduced, which also utilizes the neighborhood of query images and guarantees much consistent image retrieval output.
– To achieve high semantic retrieval performance, an effective feature extraction measure is introduced, which performs an in-depth texture analysis of images through best nodes of the wavelet packets tree, Eigen values of Gabor filter, and curvelet transform of the images.
– To avoid the risk of mis-associations, relevance feedback is incorporated in proposed method, which gives more freedom to the user to guide CBIR system in case of complex images.

Using GA as the component of training for support vector machines is due to three facts: (1) Image retrieval problem is considered as the optimization problem, in which we are concerned with the maximization of the relevant results against the query images. This can be achieved in a best way when the retrieval abilities of an efficient image classifier such as SVM can be combined with a fast convergence algorithm such as GA. (2) When feature placement in feature vectors is independent of position, like in proposed features, then GA can be used to explore the most optimal arrangements of features, so the distance can be normalized for right output generation. (3) Support vector machines are not able to perform well when training sets contain far more negative examples as compared to the positive training examples. So to handle this problem, we introduced a scheme through which we extend the positive training set for support vector machines through GA. For this, we generate new feature vectors in the positive training set from already existing feature vectors, which are present in the positive training set. Due to these mentioned reasons, we used GA with support vector machines, and presented a system, which ensures high retrieval results against any query image.

The rest of paper is organized as follows. Section 2 provides the details of related work done in the area of content-based image retrieval. Section 3 is focusing on the proposed method introduced in this paper. Experimentation and results are covered in Sect. 4. Finally, we conclude in Sect. 5.

2 Related work

IBM's QBIC system [4] is considered as the first CBIR system [5], which open the horizons for the research in the area of content-based image retrieval. Then, many CBIR systems appeared, which aimed to address the image searching problem in more effective way by addressing new signature types and image similarity detection measures. For signature development, the researchers, focused on either texture [6–8], color [9], or shape [10] features or any of these combinations [11,12]. Color features are widely used in CBIR, which may be ascribed to the superior potentiality of three-dimensional domain over the single-dimension domain of a gray-level images [9,13]. Texture features are also very powerful visual features and have the ability to capture repetitive patterns of a surface. Texture features are widely used in domain-specific applications of CBIR such as Aeriel and medical imaging [14,15]. In [16], an image retrieval system CTDCIRS (color texture and dominant color-based image retrieval system) is provided for image retrieval using three features called dynamic dominant color (DDC), Motif co-occurrence matrix (MCM), and difference between pixels of scan pattern (DBPSP). Initially, the image is divided into eight coarse partitions using color quantization algorithm, and the eight dominant colors are obtained from eight partitions. Next, the texture of the image is represented by the MCM and DBPSP. The three features (dominant color, MCM, and DBPSP) are integrated to facilitate the image retrieval system. In [17], a texture feature based on curvelet transform is proposed. The technique makes use of curvelet transform and is combined with a region-based vector codebook sub-band clustering (RBSC) for dominant color extraction with efficient curvelet-based sub-band texture extraction. Shape features are also very important type of visual features in the domain of CBIR, but as they suffer from the issues like inaccuracy of segmentation, they are not widely used for content-based image retrieval [18,19]. In [20], an image is uniformly divided into eight coarse partitions as a first step. After the above coarse partition, the centroid of each partition is selected as its dominant color. Texture of an image is obtained by using gray-level co-occurrence matrix (GLCM). Color and texture features are normalized. Shape information is captured in terms of edge images computed using gradient vector flow fields. Invariant moments are then used to record the shape. Some other visual features are also proposed for content-based image retrieval, such as salient points

and spatial features. SIFT [21] and SURF [22] are the well-known visual features based on salient points. These visual features have shown promising results for image retrieval, but due to the high dimensionality of SIFT- and SURF-based feature vectors, these features can cause over-fitting problem when employed with support vector machines for image retrieval purposes. In [21], a bag of features based model is presented by integrating SIFT and LBP features and is combined with weighted k means clustering algorithm for image retrieval. In [5], Wang et al., introduced a semantics classification method, which uses a wavelet-based approach for feature extraction, and then for the image comparison, they used image-segmentation-based region matching approach. IRM proposed in this work is not efficient for texture classification due to uncertain modeling. So to address this issue, their idea was further processed by Chen et al. [3]. They introduced an unsupervised clustering-based technique, which generate multiple clusters of retrieved results and give more accurate results as compared to the previous work, but their method suffers from issues such as numbers of clusters identification and segmentation uncertainty, due to which the results of this technique are not reliable.

An important focus of research in content-based image retrieval is on relevance feedback (RF) [2,23,24]. The main theme of relevance feedback is to keep the user in a loop through feedback requests to improve the performance of CBIR. But it exhibits some limitations like over sensitivity, and inability to accumulate knowledge thats why these systems are still not able to give the robust solutions. Another important reason due to which RF is not always popular is the unwillingness of users to provide feedback during interactive retrieval tasks. Therefore, a mechanism is required through which users should be able to retrieve images in an automatic way or in case of complex images; minimum feedbacks should be required to achieve the desired output. Historically, RF systems use machine learning techniques such as EM and KNN to bring semantically similar results in response of any query image [23]. New relevance feedback learning methods have recently been proposed among which SVM [25–27], neural networks [24,28], and Bayesian inference are most popular ones [29]. SVM and neural network-based methods consider the retrieval process as classification problem, in which relevant and irrelevant images are considered as two separate training sets; these approaches usually fail to produce good results in case of imbalanced feedback samples. So a lot of research has been carried out to handle this problem. SVM active learning [26], which plays an important role in CBIR relevance feedback research, selects the samples near the SVM boundary and queries the user for labels. Then, it selects the points near the SVM boundary and regards them as most informative images. Constrained similarity measure-based support vector machines (CSVM) [30], which are employed with Adaboost consider reposi-

tory images belonging to two clusters and learns a boundary to separate them and return the results after sorting them; the major flaw of this technique is its inability to handle the imbalanced training sets. Asymmetric bagging and random subspace for support vector machines (ABRSSVM) [23] came with classifier committee learning paradigm to handle three main problems due to which SVM classification-based RF techniques usually fail. These include instability of SVM to perform for small sets, SVM hyperplane bias, and overfitting in case of high-dimensional feature vectors and noisy images in the training set. This technique works by bagging only the negative examples to make equal sets for training, therefore can easily be misguided in case of less positive feedbacks by the user.

To improve the retrieval process in CBIR, an important focus of research is on evolutionary computation. GAs are used to improve indexing [1,31,32], combining retrieval descriptors [33], and RF-based retrieval [34,35]. The main emphasis of GA-based techniques is on the selection of parameters, i.e., parameters of SVM or feature subset selection [25,36], to see which features are effective, its importance increases when the number of features are large. But none of these techniques are able to handle classifier association problems and are unable to deal with positive asymmetric bagging to increase the correct association rate. In [35], a RF mechanism for CBIR method based on an interactive genetic algorithm (IGA) is proposed. Color attributes such as the mean value, the standard deviation, entropy based on the gray-level co-occurrence matrix, and the image bitmap of a color image are used as the features for retrieval. To reduce the gap between the retrieval results and the users' expectation, the IGA is employed to help the users identify the images that are most satisfied to the user's need.

The technique we presented in the current paper is also based on the evolutionary computation. In the current paper, we focused on finding the ways through which GA can be utilized for SVM retrieval performance enhancement. For this, we resolved two issues: (1) when there are multiple semantic classes and training sets are imbalanced; (2) when the CBIR involves the system user in the image retrieval process and the output feedback contains far more negative responses as compared to the positive ones [23]. In the proposed method, GA is also utilized for true hyperplane generation for SVM classifier, even when small positive training sets are available. The proposed architecture is useful for both automatic and interactive CBIR systems which are utilizing the SVM classifier.
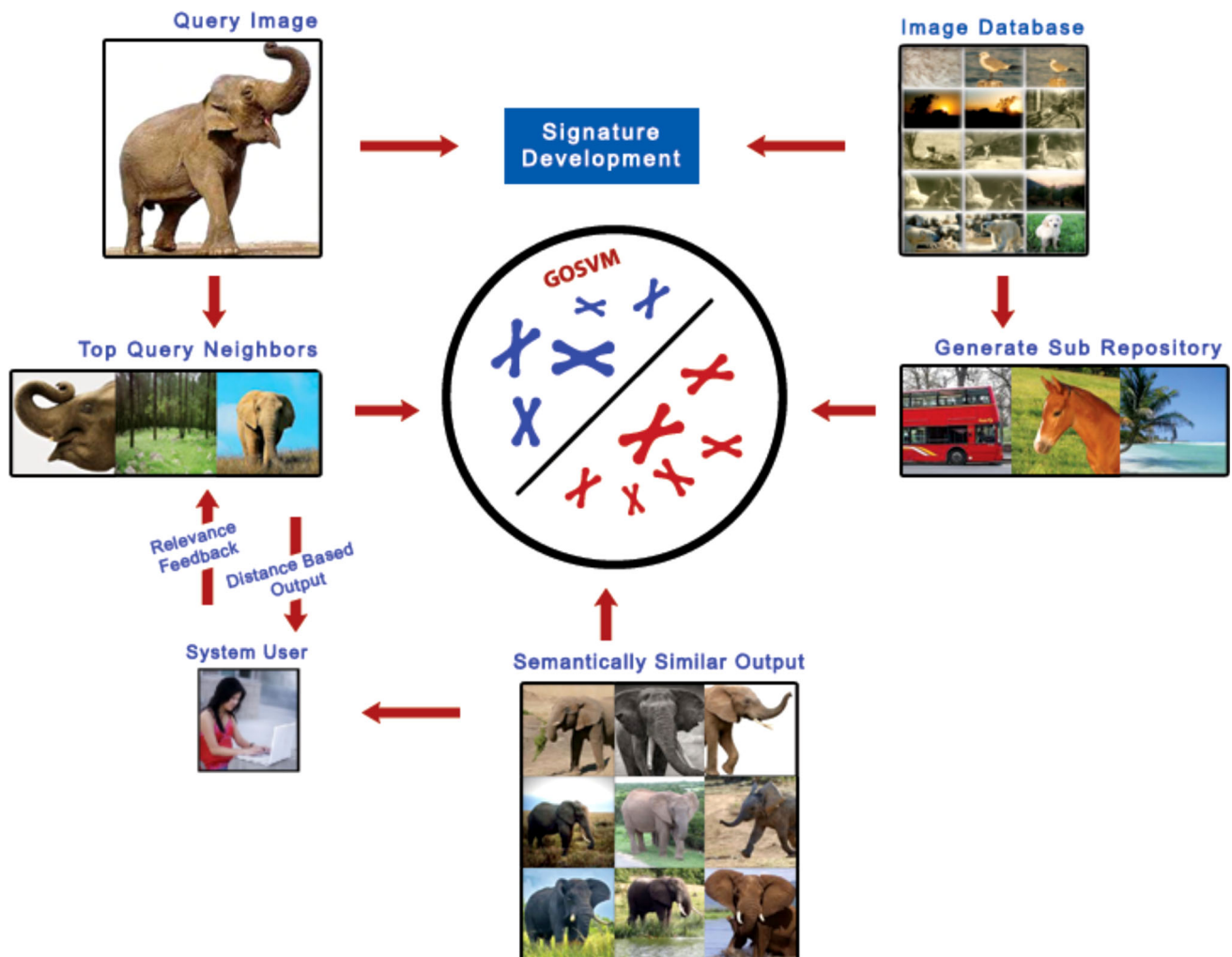
3 Proposed method

The proposed CBIR system considers categorical data for image retrieval purposes. As a first step for the implemen-

tation of CBIR process, CBIR system generates a feature repository. For this, it analyzes the repository images through curvelet transform and the best nodes of wavelet packets tree; the algorithm also uses smallest approximation image of wavelet packets tree to perform the detailed Gabor analysis and selects a set of Eigen values to completely represent them. Then, by fusing these features, it generates the corresponding feature vector for image representation purposes. After feature repository generation, 'n' bags of images (BOI) [13,31,37] are obtained by placing $R \geq 2$ example images (feature vector of images) in every image bag to represent all semantic classes. On these BOIs, a support vector machine (SVM)-based architecture is defined for every semantic class.

As SVM classifier is binary classifier, therefore training set for every SVM consists of positive samples (i.e., images belonging to a particular BOI) and negative samples (i.e., images belonging to all other BOIs). For every SVM, positive samples are passed to the GA, which returns genetically diverse chromosomes against them, and consider negative

samples as the chromosomes that are not diverse. Finally, all SVMs are trained on these two sets of chromosomes. These trained SVMs are used together with top neighbors of every image in the image repository to determine their semantic class.

Semantic association process is applied only once on all repository images. When semantic class is determined for all images, then CBIR system needs to process only the query images through the aforementioned process. For this, CBIR system is following query by example (QBE) scheme. Once semantic class of the query image is determined through top neighbors and trained SVMs, CBIR system returns images having same semantic class to the user after sorting them against query image. To further enhance the capabilities of proposed system and to avoid the risk of mis-association, system also returns 'M' top neighbors to the user and enables relevance feedback upon them. Architecture of the proposed method is presented in Fig. 1. Detail of the system is provided in following subsections.

Fig. 1 Architecture of the proposed method

3.1 Bags of images

There are $| D |$ images in image database belonging to $(A = n)$ categories. The CBIR divides the imagebase into $L = \{l_1, l_2, \ldots, l_n\}$ subsets as per the number of categories present in the imagebase. These subsets are known as bags of images (BOI). BOI is a semi-supervised image clustering technique, in which background knowledge for image clustering purposes is incorporated in the form of example images [13,31,37]. Each BOI contains $| D |_L = | D | /A * 0.3$ images from each category. For every BOI, system considers the images (feature vectors) present in image bag as positive examples $| d |^+$ and all other images present in other image bags as negatives examples $| d |^-$. This is represented as follows:

$$| d |^+ = \{i = 1, 2, \ldots, | l_i || I_i \in | D_{l_i} |\} \tag{1}$$

and

$$| d |^- = \{i = 1, 2, \ldots, | L - l_i || I_i \in | D_{L-l_i} |\} \tag{2}$$

On these sets, system defines category-specific support vector machines induced by the GA.

3.2 Imagebase representation

For imagebase representation, we analyze images through curvelet transform, wavelet packets, and Gabor features. The detail of the process is as follows.

3.2.1 Curvelet transform

Curvelet transforms [38] are the extension of the ridgelet transform to multiple scale analysis and cover the complete spectrum of frequency. This leads to why curvelet has better retrieval performance than normal wavelet transform [17]. Two main aspects should be highlighted. First, curvelets capture more accurate edge information or texture information than wavelets. Second, as curvelets are tuned to different orientations, curvelets capture more directional features and more accurate directional features than wavelets [17].

By tuning ridgelet into different orientations and scale, we can create curvelets, i.e., given an image function $f(x, y)$, the continuous ridgelet transform is given as [39]:

$$\Re_f(a, b, \theta) = \int \int \psi_{a,b,\theta}(x, y) f(x, y) \mathrm{d}x \mathrm{d}y, \tag{3}$$

where $a > 0$ is scale, $b \in R$ is the translation, and $\theta \in [0, 2\pi]$ is the orientation. The ridgelet is defined as [39]:

$$\psi_{a,b,\theta}(x, y) = a^{-\frac{1}{2}} \psi\left(\frac{x \cos \theta + y \sin \theta - b}{a}\right) \tag{4}$$

Ridgelets are oriented along angle θ and are constant along lines $x \cos \theta + y \sin \theta = \text{const}$. A ridgelet is linear in edge

direction and is much sharper than a conventional sinusoid wavelet [40]. Now if we compare it with 2D wavelets transform, then according to the wavelet domain:

$$\psi_{a1,a2,b1,b2}(x, y) = a1^{-\frac{1}{2}} a2^{-\frac{1}{2}} \psi\left(\frac{x - b1}{a1}\right) \psi\left(\frac{y - b2}{a2}\right) \tag{5}$$

By comparison, we can see that the ridgelet is similar to the 2D wavelet except that the point parameters $(b1, b2)$ are replaced by the line parameters (b, θ). In other words, the two transforms are related by:

Wavelet: ψ scale, point-position

Ridgelet: ψ scale, line-position

So this means that ridgelet can be tuned to different orientations and different scales to create the curvelets, the benefit of this scheme is to have a complete cover of the spectrum in frequency domain.

As a first step of feature extraction for any image, we take its representation in curvelet transform via wrapping. It gives the curvelet representation in the form of multiple bands and sub-bands. For every sub-band, we calculate its variance, and for every band, we take the mean of the variances of sub-bands of that band. We collect the mean values in one vector, which represents the curvelet transform of image.

3.2.2 Wavelet packets

Wavelet transform provides a suitable framework for analysis and characterization of images at different scales [41]. As significant texture information requires over complete decomposition, wavelet packet frames, which comprise of all possible combinations of sub-band tree decomposition, can serve better representation of textural analysis than standard dyadic wavelets [41].

The main difference between the discrete wavelet transformation and wavelet packets transformation decomposition is that despite just decomposing the approximation components, wavelet packets transformation decomposes the detailed components as well in order to create the complete binary tree. In this way, wavelet packets transformation assures the richest analysis of signals.

Wavelet packets procedure results in a large number of decompositions and its explicit enumerations are unmanageable. So it is necessary to find the optimal decompositions with respect to some reasonable criterion. One convenient criterion can be the selection of tree nodes on the basis of best entropy values. Shanon entropy is commonly used for this purpose, which can be calculated as [6]:

$$H(s) = -\sum_i s_i^2 \log(s_i^2) \tag{6}$$

where 's' is the signal, and $(s_i)^2$ is the coefficients of s in an orthonormal basis. This returns us the best node tree from complete binary tree of wavelet packets.

Now as a second step of feature extraction, we computed the complete Shannon entropy-based wavelet packets tree of the image up to the third level. This results in the form of 64 nodes of wavelet packets tree. But we are concerned with only those nodes, which have the best entropy values. So for this purpose, we generated the best entropy valued tree as corresponding best tree. Nodes of the best tree are used for the Haar-based feature generation using the following formula [8]:

$$f_r = \sqrt{\frac{\sum c_{ij}^2}{i \times j}}$$

where f_r is the computed Wavelet packets signature of the sub-image appeared at the node of best tree, c_{ij} represents the intensity value of all elements of sub-image. The $i \times j$ (where i represents the rows, and j represents columns) is the size of the sub-image [8]. We collect these values in a second vector.

3.2.3 Gabor features

Gabor filters are widely used in the area of pattern recognition and computer vision. Some successful applications of Gabor filter includes texture segmentation, feature extraction, iris recognition, face recognition, fingerprints identification, edge and contours detection, image compression, directional image enhancement, hierarchical image representation, and image recognition [7].

The motivation to use Gabor filters in computer vision application is that receptive fields of simple cells present in the primary visual cortex of mammals are oriented and they have characteristic spatial frequencies, which could be modeled as complex 2D Gabor filters. Gabor filters are robust to noise and can easily reduce image redundancy. Gabor filters can either be convolved on the whole image or it can be applied to different image portions. In such a case, an image region is described by the different Gabor responses generated through different angles, frequencies, and orientations. Gabor in 1946 proved that a signal's specificity in time and frequency is limited by the lower bound on the product of signal's bandwidth and duration; from this, the principal of uncertainty for information was derived [7].

Gabor's theory leads to the idea that a visual system should analyze visual information most economically by considering pairs of perceptive fields of symmetrical response and asymmetrical response profiles to achieve minimum uncertainty in both spatial localization and frequency [7].

Now as a third step of signature development, we take the smallest approximation image of wavelet packet decompo-

tion for Gabor analysis. In our implementation, we are using only the odd components of Gabor filter, so imaginary values can be avoided. Filters we used are defined by the following equation [15,42,43]:

$$G_o(x, y) = \exp\left(\frac{-x_\theta^2 - \gamma^2 y_\theta^2}{\sigma^2}\right) \sin\left(\frac{2\pi x_\theta}{\lambda}\right) \quad (7)$$

where

$$x_\theta = x \cos_\theta + y \sin_\theta \quad (8)$$
$$y_\theta = -x \sin_\theta + y \sin_\theta \quad (9)$$

and σ is the standard deviation of the Gaussian function, λ is the wavelength of the harmonic function, θ is the orientation, and γ is the spatial aspect ratio, which is left constant at 0.5. The spatial frequency bandwidth is the ratio σ/λ and is held constant and equal to .56. Thus, there are two parameters, which changes when forming a Gabor filter θ and λ. For obtaining the Gabor response, images are divided into 9×9 non-overlapping regions and are convolved by Gabor filter with aforementioned parameters. For generating the Gabor response, we are convolving the image with 12 Gabor filters, which are tuned to four orientations (θ) and three frequencies (γ). Orientation varies from 0 to $3\pi/4$ (stepping by $\pi/4$), and frequency varied from .3 to .5. After generating the response images, following scheme is used for feature extraction:

1. As there are four orientations and three frequencies, so on the basis of this, we obtain twelve Gabor-based response images after applying the aforementioned parameters.
2. On these response images, we obtain the eigenvector [44] corresponding to every Gabor response image. This results in the form of twelve eigenvectors. Eigenvectors are the linear transformations of the 2D squared matrices [44] and are widely used in computer vision applications like face and voice recognition and feature vector dimensionality reduction [45]. An eigenvector of a square matrix is a nonzero vector that is when multiplied by the matrix will yield a vector that will differ from the original at most by a multiplicative scalar [44]. The benefit of the scheme is that the multiple 2D Gabor response images will be represented by 1D vector.
3. The size of the Gabor response features in the form of eigenvectors is still very large; therefore, we take the mean of every eigenvector and merge it in one vector. So we have the representation of twelve Gabor response images in one vector. This vector serves for us the corresponding Gabor feature vector.

3.2.4 Hybrid features

Application of aforementioned procedure returns three feature vectors, representing texture features obtained from

curvelet transform, wavelet packets, and Gabor filters, respectively. Aggregation of these feature vectors in a single vector represents hybrid texture features against any image.

3.3 GA for CBIR

Let the query image and database images be represented by feature vectors as $X = (x_1, x_2, \ldots, x_n)$ and $Y_j = (y_{j1}, y_{j2}, \ldots, y_{jn})$, respectively, where '$n$' is the number of features in feature vectors, and x_i and y_i are the feature values of the ith feature. Simplest way of determining similarity between these two feature vectors is by computing the distance under the given dissimilarity metric. We can use the normalized Manhattan distance for this purpose, which can be defined as:

$$\text{Dist}(X, Y_j) = \sum_{i=1}^{n} | x_i - y_{ji} | \qquad (10)$$

After taking the distance by aforementioned dissimilarity metric, we can return the images, which appear similar and close to the query image.

But for the feature vectors, where feature placement is not restricted by the position, and they can be placed independently (e.g., as in our proposed features, those obtained by the curvelet transform are not bound to appear last in feature vectors, they can appear anywhere), we can optimize the distance between two feature vectors by simply arranging features in different order. So the process can be expressed as:

$$\hat{y}_m^j = \text{CBIR}\left(I_m\right) = \text{Retrieval}\left(F_m^j\right) \qquad (11)$$

$$\text{Retrieval}\left(F_m^j\right) = \text{Rank}\left(\left\|\text{argmin}\left(F_p^j - F_m^j\right)\right\|^{\alpha}\right) \qquad (12)$$

where

$$\alpha = \{1, 2, \ldots, l\}; p = \{1, 2, \ldots, | D || I_p \in D\}$$

where F_m^j, $\hat{y}_m^j$ represents the feature vector of the query image I_m and query results obtained by the retrieval represented as Retrieval(F_m^j) of CBIR system represented as CBIR(I_m). argmin$(F_p^j - F_m^j)$ represents that we should select the arrangement, which appears most similar to the query image in terms of distance $\|.\|$. α represents 'l' (a subset of all possible arrangements) possible arrangements of the feature vector. Results are returned after ranking on the basis of similarity with respect to query image. Ranking is represented as Rank(.) in Eq. (12).

The afore mentioned phenomena supports the use of GA [46,47]. As feature vectors can be arranged in multiple ways, in which some arrangements will be better than the original arrangement, therefore we can explore arrangements of the feature vectors by GA to find out, which arrangements

appear better than the original ones. Another important factor, which motivates us to use the GA is that SVM consider many images belonging to same semantic class as positive images for training purposes. Therefore, if we combine features from relevant training images, we can produce an effect like merging different image portions of different images to virtually generate new relevant images. So the GA is the best possible option for this purpose.

3.4 GA architecture

We want to optimize the retrieval process, by obtaining maximum positive images against any query image. For this positive training, images (feature vectors) are passed to the GA, which return chromosomes on which we train the support vector machines. In this regard, all positive images, which are passed to the GA, are considered as the elite parents. For population generation algorithm randomly selects two parents; and two cut point positions in each parent, than with the help of genetic operators it generates two new offspring. Offspring, which pass the evaluation test, also becomes the part of elite set. Algorithm generates multiple populations and returns a population, which gives the best fitness value, appeared against a fitness function for population selection. Further details of the process are provided in following subsections.

3.4.1 Structure of the chromosomes

Chromosomes are defined as follows:

$$\Theta = [\theta_{j,1}, \theta_{j,2}, \ldots, \theta_{j,N}], j = \{1, 2, \ldots, M\} \qquad (13)$$

where N are the number of genes in one chromosome, and M is the population size. We generate the population of chromosomes to perform the genetic operations. Initial population of chromosomes are the original feature vectors of positive images present in any BOI. So the parent chromosomes are represented as:

$$\Theta = [F_{\text{wtp}}^k, F_{\text{eig}}^q, F_{\text{cur}}^w] \qquad (14)$$

where $k = \{1, 2, \ldots, h\}$ are the features obtained from best nodes of wavelet packet tree represented as F_{wtp}^k, $q = \{1, 2, \ldots, 12\}$ are the features obtained from mean of the Eigen vectors of Gabor filter represented as F_{eig}^q and, $w = \{1, 2, \ldots, h\}$ are the features obtained from curvelet transform of the image represented as F_{cur}^w.

3.4.2 Population generation

We generate population of the chromosomes with the help of genetic operators crossover and mutation. Population size is kept constant, which is equal to 100. We are using crossover

and mutation for generating new offspring. The technique we are using for offspring generation is as follows:

$$\text{Parents} = \text{rand}(2); \tag{15}$$

$$\text{Cutpoints} = \text{rand}(2); \tag{16}$$

$$\hat{\Theta}_1 = [\theta_{P1,1} \to \theta_{P1,C1}; \theta_{P2,C1} \to \theta_{P2,C2};$$
$$\times \theta_{P1,C2} \to \theta_{P1,N};] \tag{17}$$

$$\hat{\Theta}_2 = [\theta_{P2,1} \to \theta_{P2,C1}; \theta_{P1,C1} \to \theta_{P1,C2};$$
$$\times \theta_{P2,C2} \to \theta_{P2,N};] \tag{18}$$

where $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are two new offspring. P and C are the parents and cutpoints, respectively. For generating new offspring, we are randomly selecting two parents and two random cut point positions. Then, by merging the genes of both parents as per Eqs. (17) and (18), we are generating two new offspring. In case both of the random numbers are the same for the parents, then we are applying mutation operator by randomly interchanging the positions of genes as per the cutpoint values in that particular parent. For elitism, we are selecting original chromosomes representing parents and the newly generated chromosomes, which passed the evaluation test for next iterations. Our chromosome evaluation test is as follows:

$$\hat{\Theta}_n = || \hat{\Theta} < \Theta ||, \quad n = \{P + 1 \to M\} \tag{19}$$

According to the evaluation test, we are selecting only those offspring chromosomes, which are giving less distance from any of the elite parents. 'n' is representing the remaining offspring, which can become elite other than the original parents. We kept the mutation rate constant as 0.05.

3.4.3 Fitness function

We used following fitness measuring criteria:

$$J(\Theta_j) = \left\| \text{argmin}\left(F^j_{\text{avg}|d|^+} - F^j_{\text{avg}\hat{\Theta}} \right) \right\|^{\alpha} \tag{20}$$

As per fitness criteria, we are generating multiple chromosome populations and selecting the population whose average distance is smallest as compared to the average distance of original positive training set. Here, α defines the number of populations, which we generate for selection purposes. For our implementation, we used $\alpha = 10$. An important thing to note is that this population is different from chromosome population. Chromosome population represents the number of chromosomes, which we would like to generate through genetic operators in one solution. While this population includes those complete solution sets, which are appearing after elitism-based GA.

3.4.4 Fitness function for RF

For relevance feedback, we used following fitness measuring criteria for GA [48]:

$$J(\Theta_j) = \frac{1}{2C \mid D \mid} \sum_{m=1}^{|D|} \sum_{p=1}^{C} \sum_{q=1}^{C} \delta\left(y^j_{m,p}, \hat{y}^j_{m,p}\right) \tag{21}$$

$$0 \leq J(\Theta_j) \leq 1 \tag{22}$$

where C represents the maximum possible positive images, or the images that belongs to same category as that of the query image. As in our image database, quantity of the category images is known so the value of C can be determined. $y^j_{m,p}$ indicates the pth member from y^j_m, and $\delta(.)$ is the Kronecker delta function [23].

$$\delta(a,b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{23}$$

If CBIR solution completely match the user solution, then the fitness function results in 1, and if they are totally mismatched it results 0. We are computing the retrieval precision against any query image, and measure the precision in every iteration.

3.5 Semantic association using support vector machines (SVM)

SVM separates two classes of points by a hyperplane. Suppose our input set belongs to two classes as [49]:

$$\{(x_i, y_i)\}_{i=1}^N y_i = \{+1, -1\} \tag{24}$$

where x_i and y_i are input sets and corresponding labels, respectively. Hyperplanes are generated by finding the efficient values of weight vectors 'w' and bias 'b' as follows:

$$w^T \cdot x + b = 0 \tag{25}$$

and finds maximum margin $2/\|w\|$ hyperplanes such that two classes can be separated from each other, i.e.,

$$w^T \cdot x_i + b \geq +1 \tag{26}$$

$$w^T \cdot x_i + b \leq -1 \tag{27}$$

or equivalently

$$y_i(w^T \cdot x + b) \geq +1 \tag{28}$$

then it find outs the solution through kernel version of Wolfe dual problem with the Lagrangian multiplied by α_i [28]:

$$Q(\alpha) = \sum_{i=1}^{m} \alpha_i - \sum_{ij=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j)/2 \tag{29}$$

Subject to $\alpha_i \geq 0$ and $\sum_{i=1}^{m} \alpha_i y_i = 0$. Based on the kernel function, SVM classifier is given by:

$$F(x) = \text{Sgn}(f(x)) \tag{30}$$

where $f(x) = \sum_{i=1}^{l} \alpha_i y_i K(x_i, x) + b$ is the output hyperplane decision function of SVM. High values of $f(x)$ represent high prediction confidence, and low values of $f(x)$ represent low prediction confidence.

3.5.1 Association structure

After generating the sub-repository of images and generating new feature sets through GA, we train category-specific support vector machines on this sub-repository with the concept of one against all classes (OAA) classification. All feature vectors present in positive training set of a specific category are labeled with '1', and all other feature vectors, which do not belong to that specific category are labeled with '0'. In this way, we define training sets for all categories and train SVM classifiers upon them using quadratic programming optimization, and keeping max iterations = 1000. After training of these support vector machines, all images present in image repository are tested against all trained support vector machines, and on the basis of decision function, they are associated with their specific semantic class. Our decision function is as follows:

$$l^* = \operatorname{argmax}(\bar{y} f l) \tag{31}$$

where $l = \{1, 2, \ldots, n\}$ are the total number of support vector machines, $\bar{y} f l$ returns the association of corresponding support vector machine, and l^* represents the obtained associated class.

3.5.2 Class finalization

Due to the object composition present in the images, many images may tend to belong to more than one category or semantic class. So in this case, it is possible that decision function may associate them with undesired classes. Therefore, it is required that the process of association should be further enhanced. This is the reason that for the finalization of semantic class for any input image, we do consider its top K neighbors as well ($K = 5$ in our case) and we are using majority voting rule (MVR) for class finalization purposes [23].

$$L^* = \operatorname{Sgn}\left\{ \sum_i l_i(X) - \frac{N-1}{2} \right\} \tag{32}$$

MVR does not consider any individual behavior of each weak classifier. It only counts the largest number of classifiers that agree with each other [23]. So according to Eq. (32), the class of input image is one on which input image and/or most of its neighbors are agreed.

3.6 Content-based image retrieval

The aforementioned process is applied on all images present in the imagebase and their semantic class is determined. Therefore, when the system suggests an output or semantic class for any query image, only images having the same semantic class are returned to the user after ranking on the base of distance with respect to the query image.

For some query images, semantic association may occur wrongly; in that case, the returned output contains mostly undesired images. To avoid this situation, system also returns query top neighbors to the user, and enable relevance feedback upon them. Relevance feedback gives freedom to the user to guide image retrieval system in case of complex queries to achieve desired output. Both sets of images are returned to the user in the form of representative images. These representative images are selected from both output sets and are the images that appear most similar to the query image in terms of distance.

3.6.1 RF overview

The relevance feedback system generates the initial output against a query image on the basis of Manhattan distance. For this, it uses distance-based image set appeared against the query image. User gives initial feedback by selecting only positive images in set, rest of the images are considered as negative feedbacks. Positive samples are passed to the GA, and it returns genetically diverse chromosomes against those positive samples, and consider negative samples as chromosomes, which are not diverse and after few generations/iterations they will intentionally die. We train support vector machines on these two classes of chromosomes and return the output and need feedback from user. The process remains continued until user become satisfied from the output.

4 Experiment and results

To elaborate the effectiveness of proposed method, we performed extensive experiments on a real dataset and compared against several algorithms for CBIR. Details of the experiments and analysis are presented in following subsections.

4.1 Database description

To show the effectiveness of the proposed method, extensive experiments are conducted on Corel dataset having 10900 images. Corel dataset has two versions, Corel set A and Corel set B. Corel set A has 1000 images divided into 10 categories, namely Africa, Beach, Buildings, Buses, Dinosaurs, Elephants, Flowers, Mountains, Horses, and Food (Fig. 3)

[35]. Corel set B has 9900 images belonging to several groups such as sunset, texture, butterfly, birds, animals, jungle, cars, and boats [2]. As already described in Sect. 3, system returns two image sets against any query image. (1) Images obtained on the basis of semantic association. (2) To avoid the risk of mis-association, it returns top query neighbors for relevance feedback obtained on the basis of distance. To elaborate the effectiveness of proposed method, we used Corel set A to measure the performance of system for semantic associations as described in Sect. 3. To further elaborate the performance of proposed method in case of mis-associations, we used relevance feedback method, and for this, we combined both versions of Corel dataset to bring more diversity. We reorganized the Corel Photo Gallery, because (1) many images with similar concepts were not in the same group and (2) some images with different semantic contents were in the same group in the original database. In the reorganized database, each group includes at least 100 images and the images in the group are category homogeneous [2,23].

4.2 Query examples

As a first step for the implementation of the proposed method, grayscale versions of repository images are generated. This preprocessing is required to perform the feature extraction in a cost-effective way. Therefore, as a backend process feature extraction is performed on grayscale versions of images and to display the results, we display the color versions of the retrieved images. Each image in the image dataset is tested against the proposed method and its semantic class is determined. Results of semantic association are stored in a file, which serves as the association database. The benefit of this scheme is that this process is done only once on the database images, and only the class specific ranking is required for query images after their association with semantic classes.

For performance evaluation of the proposed method, we randomly selected three images belonging to three different semantic classes, namely Buildings, Africa, and Elephants as query images, and then displayed the retrieval results against them. The results of top 20 retrievals against the query images are shown in Fig. 2. The response of the system can be observed from the number of correctly matched images appeared against query images (Fig. 3).

4.3 Retrieval precision and recall evaluation

To evaluate the effectiveness of the proposed method, we determined how many relevant images are retrieved in response of a query image. For this, retrieval effectiveness is defined in terms of precision and recall rates. Precision also known as specificity determines the ability of system to retrieve only those images, which are relevant for any

query image among all of the retrieved images. Precision is defined as:

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of images retrieved}}$$

Recall rate is also known as sensitivity or true positive rate and determines the ability of classifier system in terms of model association with their actual class. Recall rates can be determined by:

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of relevant images}}$$

Experimental results are reported after running five times on twenty query images randomly selected from each image category. For each query image, relevant images are considered to be those images only, which belong to the same category as that of the query image (Fig. 4).

Top 20 retrieved images are used to compute the precision and recall rates. In order to show the superiority of proposed technique, it is compared with Hung and Dai's [50], CTD-CIRS [16], ICTEDCT [17], Babu Rao et al. [20], and IGA [35]. Table 1 describes the class-wise comparison of the proposed system with comparative systems in terms of mean precision values. Same results are graphically illustrated in Fig. 5. Similarly, Table 2 presents comparison of the proposed system with comparative systems in terms of mean recall values. Recall results are graphically illustrated in Fig. 6. From the results, it can be observed that our proposed system is showing promising precision and recall rates in many categories and has the highest overall precision and recall values against comparative systems. By changing the parameters such as population size in GA, we can further improve the results in categories, where any comparative system is giving better results than the proposed method. But this will certainly affect the image retrieval speed, as more time will be required for SVM training.

As in terms of overall precision and recall values, ICTEDCT [17] is appeared as second best system. Therefore, to show the retrieval capacity of the proposed system, we compared it with ICTEDCT [17] on different number of returned images. From results illustrated in Fig. 7, it can be observed that our system has consistently shown better results on different number of returned images. Therefore, on the basis of results, we can say that our system is more consistent and robust toward the image retrieval as compared to the existing comparative systems.

To further elaborate the performance of the proposed method, we compared the proposed method with standard support vector machines without GA for image categorization problem. In image categorization, we want to find out the classification accuracy of any algorithm. Classification accuracy can be defined as:

$$\text{Class Accuracy} = \frac{\text{Total right class retrievals}}{\text{Total class images}} \times 100$$
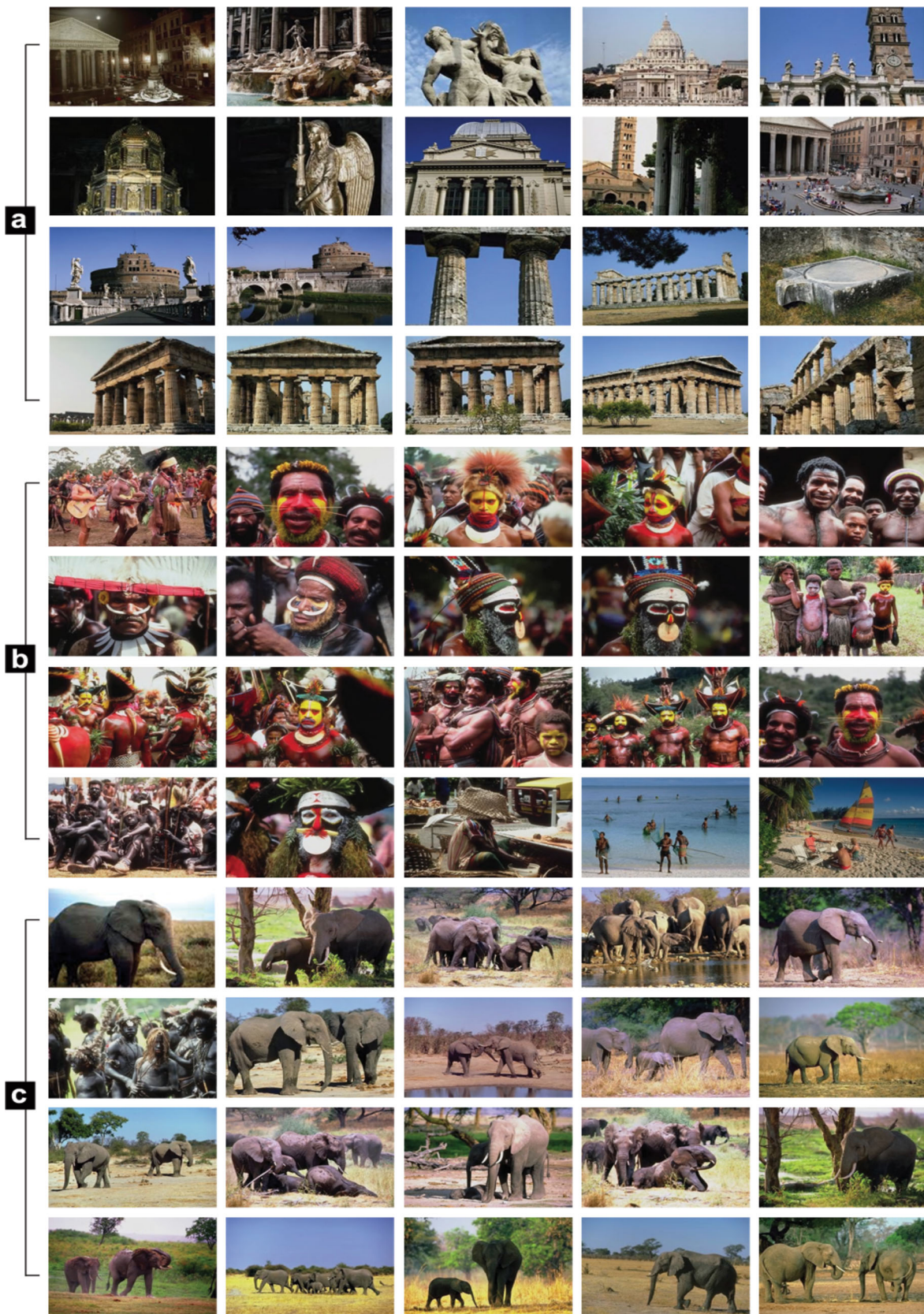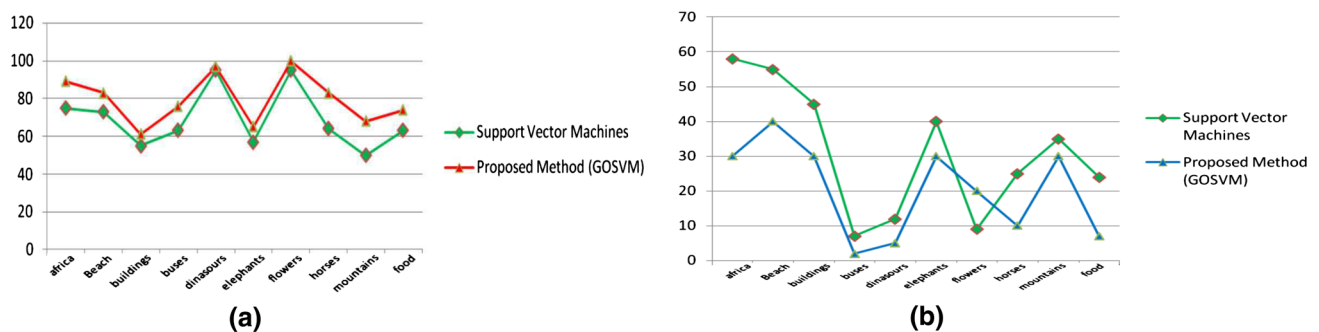
Fig. 2 Image retrieval results for class Buildings, Africa, and Beach. *First image* in every group is the query image, and all other images in that group are retrieved results

Fig. 3 Sample images of each category of Corel set A

(a)

(b)

Fig. 4 Comparison of classification accuracy and false-positive rates using support vector machines and proposed method after application of GA. **a** Class association comparison, **b** false positive comparison

Class	Proposed method	Hung and Dai's [50]	CTDCIRS [16]	Babu Rao [20]	ICTEDCT [17]	IGA [35]
Africa	0.73	0.42	0.56	0.42	0.64	0.61
Beach	0.72	0.45	0.54	0.39	0.64	0.93
Buildings	0.79	0.41	0.61	0.43	0.70	0.85
Buses	1.0	0.85	0.89	0.65	0.92	0.71
Dinosaurs	0.97	0.59	0.98	0.97	0.99	1.0
Elephants	0.75	0.43	0.58	0.63	0.78	0.79
Flowers	0.86	0.90	0.90	0.90	0.95	0.82
Horses	0.82	0.59	0.78	0.65	0.95	0.79
Mountains	0.69	0.27	0.51	0.46	0.74	0.56
Food	0.90	0.43	0.68	0.52	0.81	0.99
Mean	0.823	0.53	0.70	0.60	0.811	0.806

Table 1 Comparison of mean precision obtained by proposed method with other standard retrieval systems [16,17,20,35,50] on top 20 retrievals

As observed from the results illustrated in Fig. 4, classification accuracy of proposed method after the application of GA outperforms that of standard support vector machines.

Any CBIR system may have high correct categorization rate, but it is quite possible that the system is still not effective in image retrieval. This happens because of high false-positive rate or misclassified images. High false-positive rate

is a big reason of low retrieval accuracy of any CBIR system. A good CBIR system must have low false-positive rate as well. As observed from the results illustrated in Fig. 4, false-positive rates of proposed method are lowest as compared to standard support vector machines. Therefore, we can say that proposed GA-based architecture is more effective in image retrieval than standard support vector machines.
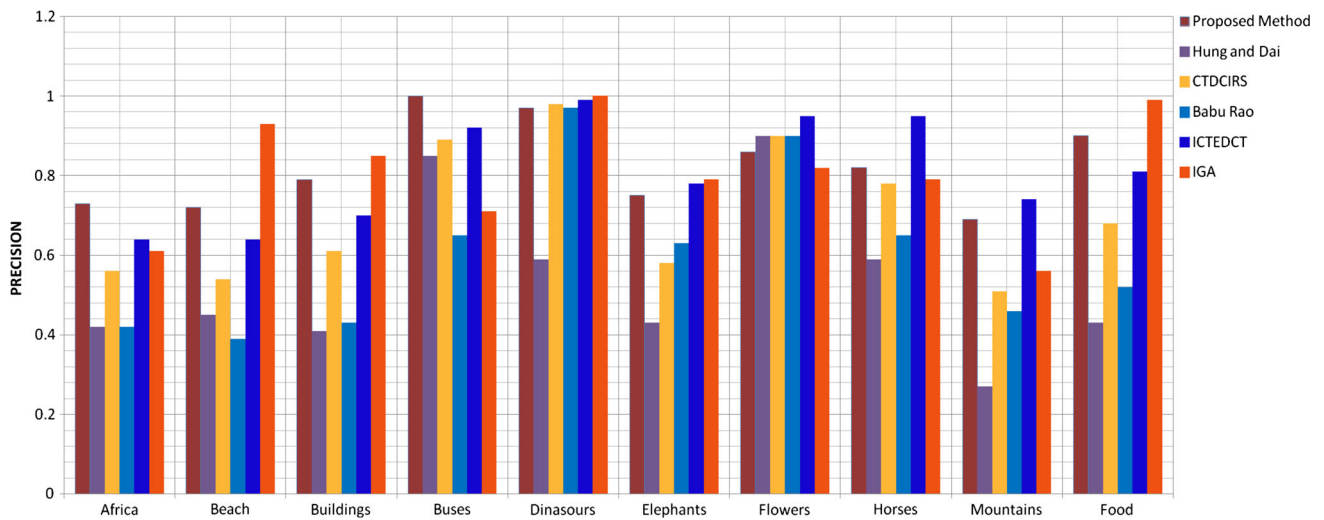
Fig. 5 Comparison of mean precision obtained by proposed method with other standard retrieval systems [16,17,20,35,50]

Table 2 Comparison of mean recall obtained by proposed method with other standard retrieval systems [16,17,20,35,50] on top 20 retrievals

| Class	Proposed method	Hung and Dai's [50]	CTDCIRS [16]	Babu Rao [20]	ICTEDCT [17]	IGA [35]
Africa	0.15	0.08	0.11	0.08	0.13	0.12
Beach	0.14	0.09	0.11	0.08	0.13	0.19
Buildings	0.16	0.08	0.12	0.09	0.14	0.17
Buses	0.20	0.17	0.18	0.13	0.18	0.14
Dinosaurs	0.19	0.12	0.20	0.19	0.20	0.20
Elephants	0.15	0.09	0.12	0.13	0.16	0.16
Flowers	0.17	0.18	0.18	0.18	0.19	0.16
Horses	0.16	0.12	0.16	0.13	0.19	0.16
Mountains	0.14	0.05	0.10	0.09	0.15	0.11
Food	0.18	0.09	0.14	0.10	0.16	0.20
Mean	0.165	0.107	0.141	0.12	0.162	0.161
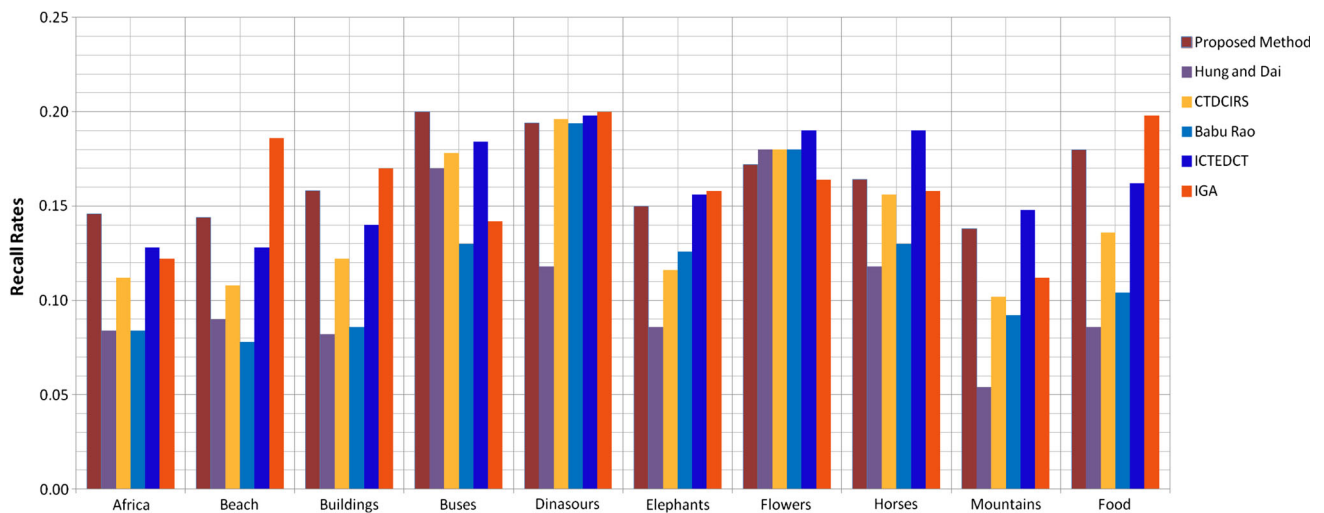
Fig. 6 Comparison of mean recall obtained by proposed method with other standard retrieval systems [16,17,20,35,50]

Fig. 7 Comparison of mean precision versus number of returned images
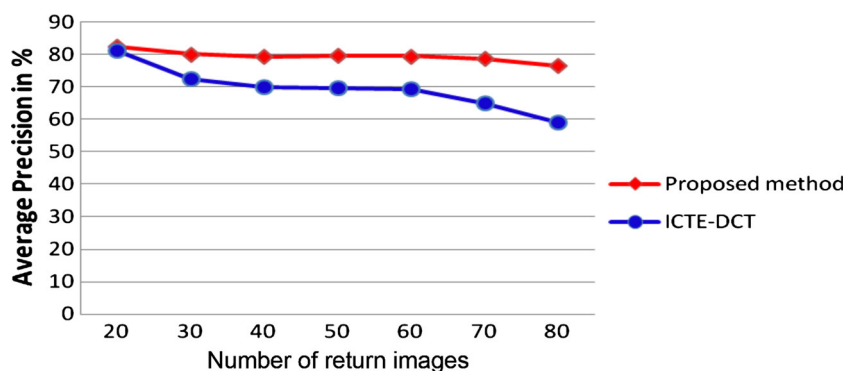
Table 3 Comparison of classification accuracy by proposed method with other standard CBIR systems [3,5,21,26,51]

Class	Proposed method	Simplicity [5]	CLUE [3]	SIFT [21]	Motif co-occurrence [51]	EMSVM [26]
Africa	89	38	49	57	45	50
Beach	83	32	37	58	39	70
Buildings	61	35	43	56	37	20
Buses	76	62	64	93	74	80
Dinosaurs	97	95	95	98	91	90
Elephants	65	28	29	58	30	60
Flowers	100	62	73	83	85	100
Horses	83	60	70	68	57	80
Mountains	68	26	28	46	29	50
Food	74	47	59	53	37	22
Mean	80	49	54	66	52	62

Table 4 Comparison of mean precision obtained by proposed GOSVM model, compared with SEMI-BDEE [2], MBA [52], BDA [53], and KBMCM [54]

Retrieval	GOSVM	S-BDEE	MBA	BDA	KBMCM
Top 1	0.992	0.991	0.988	0.978	0.99
Top 3	0.987	0.986	0.97	0.93	0.96
Top 5	0.986	0.97	0.968	0.85	0.88
Top 10	0.974	0.962	0.935	0.79	0.815
Top 20	0.92	0.94	0.78	0.58	0.775
Top 30	0.912	0.88	0.64	0.448	0.698
Top 40	0.83	0.8	0.54	0.39	0.63
Top 50	0.818	0.715	0.46	0.325	0.57
Top 60	0.801	0.645	0.38	0.29	0.51
Top 70	0.72	0.58	0.355	0.273	0.462
Top 80	0.631	0.544	0.312	0.253	0.42
Top 90	0.603	0.5	0.29	0.23	0.398
Top 100	0.561	0.46	0.262	0.212	0.36

Classification accuracy or correct categorization rate of proposed method is also compared with several existing systems of CBIR. For this, we compared the proposed method with SIMPLICITY [5], CLUE [3], SIFT [21], Jhanwar's motif co-occurrence systems [51], and Ensemble-based SVM [26]. Table 3 describes the class-wise comparison of the proposed method with comparative systems. From the results, it can be observed that our proposed method has highest overall classification accuracy as compared to the comparative systems.

4.4 Experiment for relevance feedback

The purpose of this experiment was to assess the performance of proposed system in case of wrong semantic class association through relevance feedback learning. Relevance feedback schemes based on support vector machines are widely used in content-based image retrieval. However, these schemes fail to give good results when number of positive feedback samples are much less then the negative feedback samples. This is because of following reasons: (1) When we have small-sized training set, SVM classifier becomes unstable, and (2) SVM's optimal hyperplane may be biased when the positive feedback samples are much less than the negative feedback samples. Through our experiments, we proved that our proposed system is capable of handling such situations and can provide good results even when the positive feedback samples are much less then the negative feedback samples.

Table 5 Comparison of mean recall obtained by proposed GOSVM model, compared with SEMI-BDEE [2], MBA [52], BDA [53], and KBMCM [54]

Retrieval	GOSVM	S-BDEE	MBA	BDA	KBMCM
Top 1	0.00992	0.00991	0.00988	0.00978	0.0099
Top 3	0.02961	0.02958	0.0291	0.0279	0.0288
Top 5	0.04933	0.0485	0.0484	0.0425	0.044
Top 10	0.09743	0.0962	0.0935	0.079	0.0815
Top 20	0.1833	0.188	0.156	0.116	0.155
Top 30	0.274	0.264	0.192	0.134	0.2094
Top 40	0.332	0.32	0.216	0.156	0.252
Top 50	0.409	0.357	0.23	0.163	0.285
Top 60	0.481	0.387	0.228	0.174	0.306
Top 70	0.501	0.406	0.248	0.191	0.323
Top 80	0.5046	0.44	0.249	0.202	0.336
Top 90	0.543	0.45	0.261	0.207	0.358
Top 100	0.561	0.46	0.262	0.212	0.36

4.4.1 Experimental details

In our experiments, we randomly selected 300 images from combined version of Corel dataset, and then relevance feedback is automatically performed by the computer as per the work done in [2,23]. All query relevant images (i.e., images with the same concept as the query) are marked as positive feedback samples, and all the other images are marked as negative feedback samples. We tested our method on several top images ranging from top 10 images to top 100 images. For all our reported experiments, we used 9 iterations, in which the 0th iteration returns the results obtained through Manhattan distance. Mean precision and recall are used as the performance measuring criteria of the proposed system. Precision in RF can be defined as the ratio of relevant images with respect to the total retrieved images in one feedback iteration. While recall is the ratio of relevant images retrieved with respect to the total number of relevant images in the database.
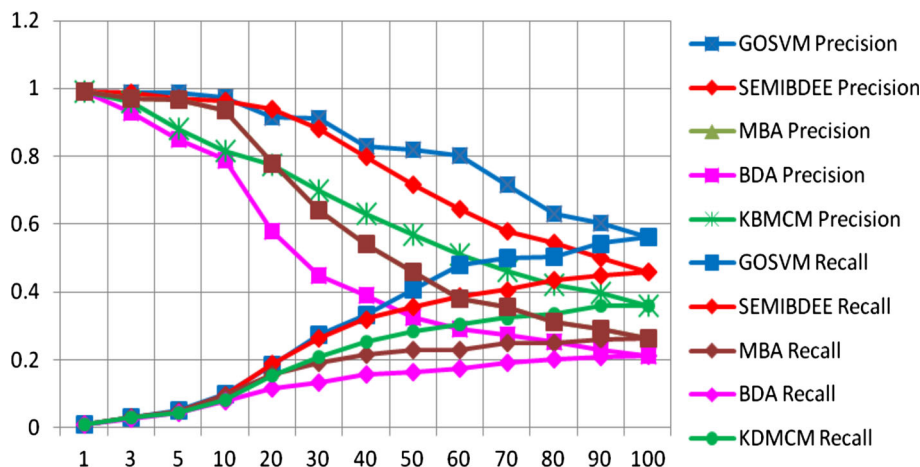
We compared the proposed method against semi-biased discriminative Euclidean embedding (semi-BDEE) [2], marginal-biased analysis (MBA) [52], kernel-biased marginal convex machine (KBMCM) [54], and biased discriminant analysis (BDA) [53]. From the results presented in Tables 4, 5, and Fig. 8, we have following observations: (1) Proposed GO-SVM has outperformed semi-BDEE, MBA, BDA, and KBMCM in all top retrieval results except in top 20, where semi-BDEE has slightly shown better results than our proposed approach. (2) Convergence of our proposed method is much better than all of the comparative techniques, which means that user will experience much better results after one or two feedback rounds. (3) Semi-BDEE has provenly shown better results than many previous SVM-based techniques such as ABRSVM [23] as mentioned in their work [2]; therefore, our results are automatically applicable to their considerations about SVM-based techniques. So on the basis of our experiments, we can say that our proposed technique is much precise and efficient than the other standard RF-based CBIR techniques.

5 Conclusion

In this paper, we introduced a multiple SVM-based architecture for image retrieval purposes, which is empowered by GA. We focused on finding the ways through which we can assure semantically correct retrieval of images against any query image. For this, we introduced a semantic association scheme, which also utilizes the neighborhood of query images and guarantees much consistent image retrieval output. To further enhance the capabilities of proposed system and to avoid the risk of mis-association, relevance feedback is also incorporated in the proposed method. Our RF scheme guarantees the efficient retrieval of images even in the case of imbalanced feedbacks as suffered by most of SVM-based RF techniques. The proposed method has also introduced a texture feature extraction method, which ensures high precision

Fig. 8 Mean precision and recall plotted against number of retrieved images for the proposed GOSVM model, compared with SEMI-BDEE, MBA, BDA, and KBMCM

in results. For this, a feature extraction phenomena is introduced through which images are analyzed by the best nodes of the wavelet packets tree, curvelet transform of images, and Eigen values obtained through Gabor filters.

In the current paper, GAs are utilized for SVM performance enhancement. For this, we resolve the problem of imbalance training for multiple classes, and imbalance feedbacks for relevance feedback through GA. In the proposed method, GA is also utilized for true hyperplane generation for SVM classifier, even when small positive training sets are available. The proposed GA architecture is benefiter for both automatic and interactive CBIR systems, which are based on SVM classifier. While using GA, it is important to note that the performance of GA relies on the model parameters. Non-accurately initialized parameters can directly affect the retrieval performance of CBIR systems; and in case of support vector machines, they can bring classification inefficiency. Therefore, it is necessary that parameters such as population size, mutation rate, crossover probabilities should be carefully defined. Two other problems while using GA are the genetic coding, which is used to define the problem, and evaluation function, which is used to measure the fitness of the solutions, can also cause the performance degradation if they are not closely observed. The proposed GA architecture is able to handle all these issues and ensures high performance in image retrieval.

To elaborate the effectiveness of the proposed method, extensive experiments are conducted on Corel Photo dataset. To measure the semantic association capabilities of proposed method, we compared it with several popular CBIR techniques such as Hung and Dai's method, CTDCIRS, ICTEDCT, Babu Rao's method, IGA, SIMPLICITY, Motif co-occurrence systems, CLUE, SIFT, and Ensemble-based SVM. Reported results shows that the proposed method is very effective in terms of associating images with their respective semantic classes. We also compared the proposed method with several popular RF algorithms, such as semi-BDEE, marginal-biased analysis (MBA), biased discriminant analysis (BDA), and kernel-biased marginal convex machine (KBMCM), on a range of top retrieval rates from small numbers (such as first hit, top 3, top 5) to large numbers (such as top 70, top 80, top 90, top 100). The proposed system has reported higher precision and recall rates compared to the other RF-based CBIR systems in all retrieval values except in top 20, where SEMIBDEE has shown slightly better results than the proposed method.

References

1. Saadatmand-Tarzjan, M., Moghaddam, H.A.: A novel evolutionary approach for optimizing content based image indexing algorithms. IEEE Trans. Syst. Man Cybern. **37**, 139–153 (2007)

2. Bian, W., Tao, D.: Biased discriminant Euclidean embedding for content based image retrieval. IEEE Trans. Image Process. **19**, 545–554 (2010)

3. Chen, Y., Wang, J.Z., Krovetz, R.: CLUE: cluster based retrieval of images by unsupervised learning. IEEE Trans. Image Process. **14**, 1187–1201 (2005)

4. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., et al.: Query by image and video content: the QBIC system. IEEE Comput. **28**, 23–32 (1995)

5. Wang, J.Z., Li, J., Wiederhold, G.: SIMPLIcity: semantics-sensitive integrated matching for picture libraries. IEEE Trans. Pattern Anal. Mach. Intell. **09**, 947–963 (2001)

6. Liane, A., Fan, J.: Texture classification by wavelet packets signature. IEEE Trans. Pattern Anal. Mach. Intell. **15**, 1186–1191 (1993)

7. Andrysiak, T., Chora's, M.: Texture image retrieval based on hierarchal Gabor filters. Int. J. Appl. Math. Comput. Sci. **15**, 471–480 (2005)

8. Gnaneswera Rao, N., Vijaya Kumar, V., Vinkata Karishna, V.: Texture based image indexing and retrieval. Int. J. Comput. Sci. Netw. Secur. **09**, 206–210 (2009)

9. Lei, Z., Fuzong, L., Bo, Z.: A CBIR method based on color-spatial feature. In: TENCON 99. Proceedings of the IEEE Region 10 Conference (2002)

10. Zhou, X.S., Huang, T.S.: Edge based structural features for content based image retrieval. Pattern Recognit. Lett. **22**, 457–468 (2001)

11. Zhang, D.: Improving image retrieval performance by using both color and texture features. In: IEEE Conference on Image and Graphics (2004)

12. Lin, C.-H., Chen, R.-T., Chan, Y.-K.: A smart content based image retrieval system based on color and texture feature. Image Vis. Comput. **27**, 658–665 (2009)

13. Jun, Z.: Robust content based image retrieval of multiple example queries. Doctor of philosophy thesis, School of computer science and software engg. Univ. of Wollongong (2011). http://ro.uow.edu.au/theses/3222

14. Faloutsos, C., Barber, R., Flickner, M., Hafner, J., Niblack, W., Petkovic, D., Equitz, W.: Efficient and effective querying by image content. J. Intell. Inf. Syst. **3**, 231–262 (1994)

15. Lama, M., Disney, T., et al.: Content based image retrieval for pulmonary computed tomography nodule images. In: SPIE Medical Imaging Conference, San Diego (2007)

16. BabuRao, M., Rao, B.P., Govardhan, A.: CTDCIRS: content based image retrieval system based on dominant color and texture features. Int. J. Comput. Appl. **18**, 09758887 (2011)

17. Youssef, S.M.: ICTEDCT-CBIR: integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content based image retrieval. Comput. Electr. Eng. (2012). doi:10.1016/j.compeleceng.2012.05.010

18. Gupta, A., Jain, R.: Visual information retrieval. Commun. ACM **40**, 70–79 (1997)

19. Smith, J.R., Chang, S.-F.: VisualSEEK: a fully automated content based query system. In: Proceedings of 4th ACM International Conference on Multimedia, pp. 87–98 (1996)

20. Babu Rao, M., Prabhakara Rao, B., Govardhan A.: Content based image retrieval system based on dominant color, texture and shape. Int. J. Eng. Sci. Technol. (IJEST) **4**, 2887–2896 (2011)

21. Yuan, X., Yu, J., Qin, Z., Wan, T.: A SIFT-LBP image retrieval model based on bag of features. In: IEEE International Conference on Image Processing (2011)

22. Valmunurang, K., et al.: Content based image retrieval using SURF and color moment. Glob. J. Comput. Sci. Technol. **11**, 1–5 (2011)

23. Tao, D., Tang, X., Li, X., Li, X.: Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1088–1099 (2006)

24. Wang, B., Zhang, X., Li, N.: Relevance feedback technique for content based image retrieval using neural network learning. In: Proceedings 5th International Conference on Machine Learning and Cybernetics, Dublin, August 2006

25. Seo, K.-K.: Content based image retrieval by combining genetic algorithm and support vector machine. Lecture Notes in Computer Science, vol. 37, pp. 537–545 (2007)

26. Yildizer, E., Balci, A.M., Hassan, M., Alhajj, R.: Efficient content based image retrieval using multiple support vector machines ensemble. Expert Syst. Appl. **39**, 2385–2396 (2012)

27. Tong, S., Chang, E.: Support vector machine active learning for image retrieval. In: Proceedings of ACM International Conference on Multimedia, pp. 107–118 (2001)

28. Muneesawang P., Guan L.: A neural network approach for learning image similarity in adaptive CBIR. In: Proceedings of Multimedia, Signal Processing, pp. 257–262 (2001)

29. Su, J.-H., Huang, W.-J., Yu, P.S., Tseng, V.S.: Efficient relevance feedback for content based image retrieval by mining user navigation patterns. IEEE Trans. Knowl. Data Eng. **23**, 360–372 (2011)

30. Azimi-Sadjadi, M.R., Salazar, J., Srinivasan, S.: An adaptable image retrieval system with relevance feedback using kernel machines and selective sampling. IEEE Trans. Image Process. **18**, 1045–1059 (2009)

31. Irtaza, A., Jaffar, A., Muhammad, M.S.: Content based image retrieval in a Web 3.0 Environment. Multimed. Tools Appl. (MTAP) (2013). doi:10.1007/s11042-013-1679-2

32. Khan, A., Ullah, J., Jaffar, M.A., Choi, T.-S.: Color image segmentation: a novel spatial fuzzy genetic algorithm. Signal Image Video Process. doi:10.1007/s11760-012-0347-8

33. Bulo, S.R., Rabbi, M., Pelillo, M.: Content based image retrieval with relevance feedback using random walks. Pattern Recognit. **44**, 2109–2122 (2011)

34. da S. Torres, R., Falco, A.X., Goncalves, M.A., Zhang, B., Fan, W., Fox, E.A., Calado, P.: A new framework to combine descriptors for content based image retrieval. In: Fourteenth Conference on Information and Knowledge Management, Bremen, Germany, pp. 335–336 (2005)

35. Lai, C.-C., Chen, Y.-C.: A user-oriented image retrieval system based on interactive genetic algorithm. IEEE Trans. Instrum. Meas. **60**, 3318–3325 (2011)

36. Wang, S.-F., Wang, X.-F., Xue, J.: An improved interactive genetic algorithm incorporating relevant feedback. In: Proceedings of the 4th International Conference on Machine Learning and Cybernetics, Guangzhou, pp. 2996–3001 (2005)

37. Irtaza, A., Jaffar, A., Aleisa, E.: Correlated networks for content based image retrieval. Int. J. Comput. Intell. Syst. (IJCIS) **6**(6), 1189–1205 (2013)

38. Irtaza, A., Jaffar, A., Mahmood, T.: A statistical image retrieval technique from large image repositories in the domain of partial supervised learning. Inf. J. **16**, No. 9(B), 7145–7152 (2013)

39. Do, M.N., Vetterli, M.: The finite ridgelet transform for image representation. IEEE Trans. Image Process. **12**(1), 16–28 (2003)

40. Sumana, I.J., Islam, M.M., Zhang, D., Lu, G.: Content based image retrieval using curvelet transform. In: IEEE 10th Workshop on Multimedia and Signal Processing, Gippsland Sch. of Inf. Technology Monash University, Churchill, VIC (2008)

41. Banerjee, M., Kundu, M.K.: Content based image retrieval using wavelet packets and fuzzy spatial relations. In: ICVGIP (2006)

42. Irtaza, A., Jaffar, A., Aleisa, E., Choi, T.-S.: Embedding neural networks for content based image retrieval. Multimed Tools Appl (MTAP) 1–21 (2013).

43. Irtaza, A., Jaffar, A., Mahmood, T.: Semantic image retrieval in a grid computing environment using support vector machines. Comput. J. (2013). doi:10.1093/comjnl/bxt087

44. Parra, L., Sajda, P.: Blind separation via generalized eigenvalue decomposition. Int. J. Mach. Learn. Res. **4**, 1261–1269 (2003)

45. Kekre, H.B., Thepade, S.D., Maloo, A.: CBIR feature vector dimension reduction with eigenvectors of covariance matrix using row, column and diagonal mean sequences. Int. J. Comput. Appl. **3**, 39–46 (2010)

46. Davis, L. (ed.): Genetic Algorithms and Simulated Annealing. Pitman, London (1987)

47. Holland, J.H.: Adaptation in Natural and Artificial Systems. Univ. Michigan Press, Ann Arbor (1975)

48. Tavares, A., da Silva, A., Falcão, X., Magalhães, L.P.: Active learning paradigms for CBIR systems based on optimum-path forest classification. Pattern Recognit. **44**, 2971–2978 (2011)

49. Stejic, Z., Takama, Y., Hirota, K.: Genetic algorithm-based relevance feedback for image retrieval using local similarity patterns. Inf. Process. Manag. **39**, 1–23 (2003)

50. Huang, P.W., Dai, S.K.: Image retrieval by texture similarity. Pattern Recognit. **36**, 665–679 (2003)

51. Jhanwar, N., Chaudhurib, S., Seetharamanc, G., Zavidovique, B.: Content based image retrieval using motif co-occurrence matrix. Image Vis. Comput. **22**, 1211–1220 (2004)

52. Seo, K.-K.: Content based image retrieval by combining genetic algorithm and support vector machines. In: Proceedings of the 17th International Conference on Artificial Neural Networks, Berlin, Heidelberg, pp. 537–545 (2007)

53. Xu, D., Yan, S., Tao, D., Lin, S., Zhang, H.-J.: Marginal fisher analysis and its variants for human gait recognition and content based image retrieval. IEEE Trans. Image Process. **16**, 2811–2821 (2007)

54. Ye, J.: Least squares linear discriminant analysis. In: Proceedings of International Conference on Machine Learning, pp. 1087–1093 (2007)