METHODOLOGIES AND APPLICATION

Efficient hybrid multi-level matching with diverse set of features for image retrieval

V. Geetha[1] · V. Anbumani[1] · S. Sasikala[1] · L. Murali[2]

Abstract

Content-based image retrieval has become popular in the retrieval of images from large image database using reduced human intervention. Researchers are still in need to develop effective systems for dealing many of the wide-scope scientific and medical applications. Past research works have faced a problem on differentiating different images by means of using the single features alone. In this paper, a multi-level matching scheme is introduced for retrieval of image based on a hybrid feature similarity integrating local and global features. Both global- and local-level features included in multi-level scheme are used for image representation. From an image, the color information is extracted globally using a new color, edge directivity descriptor and color-based features. Further, the interest of points from each image is detected using local descriptors called local binary pattern and speeded-up robust features. Using two image databases, the improved retrieval accuracy obtained with the combination of global and local features is analyzed. Experimental outcomes have revealed the effectiveness of proposed system on achieving 91% and 92% precision rates over two datasets compared to other existing methods.

Keywords Multimedia · Global · Local features · Similarity · Multi-level matching · Color-related feature · CEDD · LBP · SURF

Abbreviations

CBIR	Content-based image retrieval
CRF	Color-related feature
BoW	Bag-of-words
IR	Image retrieval
VOC	Visual object classes
CCM	Color co-occurrence matrix
SVM	Support vector machine
DWT	Discrete wavelet transform
PNN	Probabilistic neural network
CRF	Color-related feature
MLM	Multi-level matching
K-NN	K-nearest neighbor
FRAR	Full range autoregressive
RBFNN	Radial basis function neural network
EOAC	Edge orientation auto-correlogram
SOM	Self-organizing map
CHF	Color histogram feature
CEED	Color and edge directivity descriptor
TSK	Takagi–Sugeno–Kang
LBP	Local binary pattern
SURF	Speeded-up robust features
NMS	Non-maximum suppression
SIFT	Scale-invariant feature transform
SM	Similarity matching
ED	Euclidian distance
AUC	Area under the precision-recall curve
MD	Manhattan distance
CD	Canberra distance
MLM	Multi-level matching

✉ V. Geetha
  vgeethaid@gmail.com

[1] Department of ECE, Kongu Engineering College, Perundurai, Erode, India

[2] Department of ECE, P.A. College of Engineering and Technology, Coimbatore, India

1 Introduction

In computer vision field, the content-based image retrieval (CBIR) has grown as an advanced research topic (Peng et al. 2019; Zeng et al. 2019). One of the significant tasks

that take a core role in computer vision applications is feature extraction process. Returning the most relevant or similar images for a given query image by means of extracting and formulating discriminative and meaningful image representation is the main goal of feature extraction process. Notably, the low-level image features (spatial information, shape, texture and color) were highly adopted by most of the earlier research works on CBIR-based methods (Somasundaran et al. 2020; Tian et al. 2019). For an image, the whole visual contents can be described well using the global features. In other words, this feature is found to be the best one for not considering alone some specific points of interest to illustrate the complete visual contents of an image (Hongpeng 2019; Tesfaye and Pelillo 2019; Sujatha and Shalini Punithavathani 2018; Vinu 2016).

From the past decades, the researchers in computer vision field have provided their attention on analyzing the local image descriptors, such as histograms of oriented gradients (HOG) (Argyriou and Tzimiropoulos 2017), speeded-up robust features (SURF) (Bay et al. 2008) and scale-invariant feature transform (SIFT) (Lowe 2004). More commonly, the key points of particular image portions (corners/edges, object of interest, and region) are utilized to illustrate local information by the local image descriptors. Toward wide range of current computer vision applications, the local descriptors have proved their strength in image retrieval process, object tracking, visual object classification, panoramic stitching and scene categorization and so on. Most significant advantages of local descriptors on providing reliable matching across wide-scope of various conditions (Amato et al. 2019; Lai et al. 2016) and having the ability to withstand image scale and rotation invariance make them strong enough compared to other conventional global features. For improving the discrimination and robustness during image representation, the exploitation of local and global image features benefits is considered to be a challenging and interesting task in this work (Zheng et al. 2018).

Some of the soft computing methods and several significant features have been developed for computer vision applications in recent years (Sundararaj 2019; Al-Janabi and Alkaim 2019; Al-Janabi and Mahdi 2019; Sundararaj et al. 2018; Vinu 2019). However, the different forms of image deformations (such as viewpoint changes, appearance of noises, image rotations) and scale variances are handled well using local descriptors; thereby, this descriptor has improved the system robustness. In other sense, the spatial relations and objects are closer to the human vision characteristics, which means the complete image structure is considered by the global features. But, the high retrieval accuracy can be significantly achieved by extracting only the proper (accurate) features. Notably, the

CBIR system performance can be improved by considering carefully the dimension of feature vector. In fact, if this step is avoided, then the increased computational cost and memory consumption will degrade the performance of CBIR systems (Gladis 2019).

In this paper, multi-level matching scheme is proposed for content-based image retrieval (CBIR) combining local and global features. Our research contribution includes two significant parts: At first, the local and global information of the image are represented effectively through developing a new multi-level structured representation. Additionally, the complex scenes and events categories can be characterized suitably using this representation. Next, the optimal similarity among images is identified by introducing a multi-level matching technique incorporating the linear programming-solved Euclidean distance formula. Followed by this, the image retrieval accuracy is improved through using a hybrid similarity combining both local and global information. In the CBIR field, the research contributions carried out by this work are explained as follows: (a) Both global and local information are combined to develop hybrid feature information; (b) the retrieval performance is improved by introducing a color-related feature (CRF) combining other relevant features; (c) two-step retrieval steps-combined multi-level matching (MLM) scheme is introduced.

The rest of this paper is organized as follows: Sect. 2 details the related works on CBIR methods. Sections 3 and 4 explain the multi-level feature extractions and hybrid similarity-based multi-level matching scheme. The proposed image retrieval system framework is introduced in Sect. 4. In Sect. 5, the experimental results are analyzed. Section 6 concludes the paper.

2 Review of related works

In current trend, the Bag-of-Words (BoW) representation-included local features are employed by the CBIR methods. Some of the image primitive features (shape, spatial information, texture and color) were highly focused by these CBIR-based studies (Dubey et al. 2014) to attain better image retrieval accuracy. Work in Bagri and Johari (2015) has proposed the texture and shape properties-based feature extraction technique. This technique has used both the shape-invariant Hu-moments and gray-level co-occurrence matrix approaches. Texture and shape features are combined to perform the comparison. The performance metrics such as recall and precision are applied to determine the retrieval accuracy. In object recognition field, the visual cue called shape has played a major role. The classification of binary shapes is performed based on the vector quantization, feature extraction and feature

detection. In Ramesh et al. (2015), the BoW model is used to develop the invariant features-based classification framework. Experimental study is conducted through adopting shape classifier to be used in animal shapes dataset for shape classification. Work in Montazer and Giveki (2015) has introduced image descriptors with two significant methods. Feature matrix is obtained through performing k-means clustering based on SIFT extraction process. Two different forms of dimensionality reduction methods were employed to obtain high precision rate. Li database images and Caltech-101 are used for experimental validation process. Work in Li et al. (2015) has proposed the 3D shape retrieval technique.

Here, 6 and 12 dissimilar 3D shape retrieval techniques are considered to perform the evaluation process. The common benchmark evaluation is adopted to compare 26 retrieval techniques during experimental analysis (Anandh et al. 2016). Furthermore, the Wavelet Transform, Gabor Wavelet, and color auto-correlogram were used for image generation. Initially, the RGB color space is considered for extracting the features combined with color information. Then, the proposed feature extraction technique allows for the extraction of features combined with the texture information. Thirdly, the corner and edge detection process is used for the extraction of shape-based information.

A Color Directional Local Quinary Pattern also called as color-texture features is extracted using the proposed image retrieval method (Vipparthi and Nagar 2014). To the surrounding and reference pixels, the extraction of RGB channel-wise directional edge information is performed. MIT-Color and Core-5000 databases were used to conduct experimental validation process. Work in Iakovidou et al. (2015) has used the four image extraction techniques to extend and simplify the MPEG-7 descriptor functionality. Ultimately, the interest points were generated for an input image. UKbench and UCID are the two databases used to conduct experimental validation. Fuzzy classifiers were generated from the local image features used in object classification (Korytkowski et al. 2016). Local features can be determined using Meta learning approach. PASCAL Visual Object Classes (VOC) dataset with its three classes is used for experimental analysis. Work in Wang et al. (2017) has proposed a technique based on the combination of texture and shape features. Notably, this technique has extracted the texture features through applying the localized angular phase histogram; conversely, the shape features were extracted through applying the exponent moment's descriptor. To satisfy hue saturation intensity (HSI) color space, it is most to extract the texture features, whereas the RGB color space can be satisfied through extracting the shape features. In order to improve the simplified selection process, the feature selection technique was suggested by authors El Alami (2011). Essentially, the

suggested feature selection technique should select the most relevant image features. By that fact, the texture and color features were extracted using both Gabor filter and 3D color histogram process. The feature space complexity is reduced by applying the genetic algorithm. Corel-1000 dataset is used for experimental analysis.

For all complex background images, the average precision rates were reported. Work in Shrivastava and Tyagi (2015) has avoided the feature fusion process and retrieved the images using different steps. Initially, the color features are applied for the retrieval of fixed number of images. Next, the shape and texture features are applied to filter the most relevant images. However, the elimination of normalization and fusion steps has reduced significantly the computational cost. But, the spatial information of an image is not classified accurately. Moreover, the color co-occurrence matrix (CCM) is adopted by the ANN classifier (El Alami 2014). Here, the texture and color features are extracted through computing the scan pattern pixels (DBPSP). Similarity value is computed by means of presenting a feature matching strategy. For all object images, their average rates were reported. Instead of considering the multichannel descriptors, the image properties were captured using two image channels (Xiao et al. 2014). However, the embedded Sobel filter information is used for improving a hyperopponent color space performance. But, this method has not classified accurately the background and foreground objects.

3 Multi-level structured feature extraction

In this section, a multi-level feature extraction scheme is explained using the combination of global and local features. Due to the ability of showing robustness and scalability in global characteristics or information in an image, the local features are selected in this work. Conversely, the global features showing its ability in local characteristics or information are selected. However, the CBIR systems retrieval accuracy is improved with the combination of local and global features. Using CEDD and color-related feature (CRF), the global features are computed. Similarly, local features are calculated using LBP and SURF.

3.1 Global feature extraction

3.1.1 Color-related features (CRF)

To extract color information (features), we have proposed a novel feature named color-related feature (CRF). For an image, their spatial color information can be described effectively using a new image descriptor called color-related feature (CRF). This descriptor works same as that of

color histogram feature (CHF) (Guo et al. 2015). Furthermore, some of the image properties such as color distribution, color information and image brightness are also described using CRF. Max and min-quantizers are used to compute this CRF.

In CHF computation, the color indexing is performed to do the color truncation process through applying balanced-tree clustering method. Figure 1 depicts the CHF computation process (Guo et al. 2015). Several nodes altogether form a balanced-tree. Root node is identified from the top node of the balanced-tree. A set of nodes not including any child is called as leaf nodes (i.e., bottom nodes of the balanced-tree). By the fact, the nodes in a balanced-tree contain left and right child nodes. Among these, the right child will possess higher or equal value as its parent node, while the left child will possess lower value than its parent node. The color codebook of CHF can be followed to develop a balanced-tree.

Furthermore, a balanced-tree is built depending on the norms of all CHF codewords. In other words, their norms are considered to sort and arrange them in ascending order. Hence, using this arranged order, it is easy to develop a balanced-tree. Then, the leaf nodes are assigned with these sorted codewords. A new value is obtained by averaging one of the nodes in the tree and its sibling (i.e., two adjacent nodes) to form a complete balanced-tree. Then, a node using this value can act instead of a root node (parent node). Notably, for all codewords (all leaf nodes), this process is repeated continuously until the root node is reached.

Subsequently, the color quantizers can be made effectively through performing the color truncation process. This can be achieved only with the formation of a complete balanced-tree. For each max and min color quantizer, a single value representation is assigned using this color truncation process. Assume that the balanced-tree formed the color quantizer (i.e., max- and min-quantizers) and has a set of leaf nodes denoted as $T_{\min} = \left\{ \widehat{q}_1, \widehat{q}_2, \ldots, \widehat{q}_{N_{\min}} \right\}$ and $T_{\max} = \left\{ \widehat{q}_1, \widehat{q}_2, \ldots, \widehat{q}_{N_{\max}} \right\}$, respectively.

In which, the min and max sizes of color clusters are indicated as $N_{\min}$ and $N_{\max}$. More significantly, the color truncation process is possible to be performed only with the presence of different balanced-tree in different image databases. Consider an image block (i, j) having $i = 1, 2, \ldots, \frac{M}{m}$, and $j = 1, 2, \ldots, \frac{N}{n}$, then the min- and max-quantizers are represented as $q_{\min}(i, j)$ and $q_{\max}(i, j)$, respectively. For min-quantizer, the color truncation process used is expressed as follows:

$$\xi\{q_{\min}\} \Rightarrow \widehat{q}_a \tag{1}$$

Here, $a = 1, 2, \ldots, N_{\min}$. The color truncation process is denoted using the symbol $\xi\{\bullet\}$. From the leaf of the balanced-tree, the color codeword index is returned using this color truncation process. Leaf nodes that are matching close to the min-quantizer can be considered to be the min-quantizers color truncation process, which is satisfying $\arg\min_{a=1,2,\ldots,N_{\min}} \left\| q_{\min}(i,j), \widehat{q}_a \right\|_2^2$. For max-quantizer, the color truncation process is shortly expressed as,

$$\xi\{q_{\max}\} \Rightarrow \widehat{q}_b \tag{2}$$

Here, $b = 1, 2, \ldots, N_{\max}$. In $T_{\max}$ set, the color cluster and max-quantizers closest matching is expressed using Eq. (2), where it should satisfy $\arg\min_{b=1,2,\ldots,N_{\max}} \left\| q_{\max}(i,j), \widehat{q}_b \right\|_2^2$.

In a traversed fashion, the balanced-tree and color quantizers are closely matched with the aid of CIF performance. Initially, the parent (root) of the balanced-tree and its corresponding color quantizers similarity score are computed using CIF. As and when compared to the similarity score of right child of a balanced-tree, the similarity score of left child is smaller; then, it is important to continue the similarity searching process for left child in the balanced-tree. After reaching the leaf node, the repetition of this process is stopped. Ultimately, the color truncation output is obtained based on its returned index. The steps involved in the color truncation process are indicated in Fig. 2. However, the CHF after using this strategy has faced more computational complexity than the CIF.
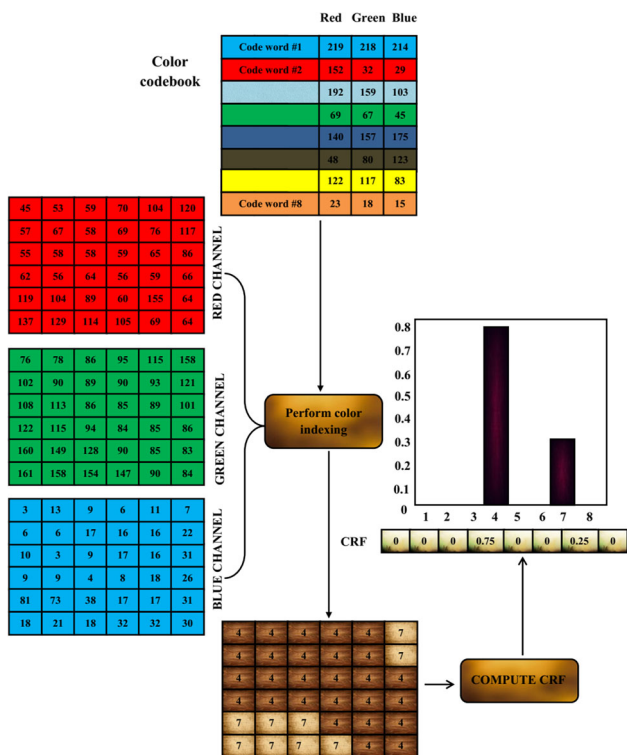

Fig. 1 Example of CHF feature computation

For min-quantizer, the CIF shows the complexity as $O(d \times \log_2 N_{\min})$. Alternatively, the CHF has shown the computational complexity as $O(d \times N_{\min})$, in which the color quantizers are represented using color dimension. By that fact, $d = 3$ is the dimensional value considered for RGB color space. Likely, the same computational complexity is required by the max-quantizer.

Based on the color truncation process, the image feature descriptors in two forms, namely $CIF_{\min}$ and $CIF_{\max}$, can be obtained. They are expressed as follows:

$$CIF_{\min}(a) = \Pr\left[\xi\{q_{\min}(i,j)\} = \widehat{q}_a \middle| i = 1, 2, \ldots, \frac{M}{m}; j = 1, 2, \ldots, \frac{N}{n}\right] \tag{3}$$

$$CIF_{\max}(b) = \Pr\left[\xi\{q_{\max}(i,j)\} = \widehat{q}_b \middle| i = 1, 2, \ldots, \frac{M}{m}; j = 1, 2, \ldots, \frac{N}{n}\right] \tag{4}$$

Here, $a = 1, 2, \ldots, N_{\min}$ and $b = 1, 2, \ldots, N_{\max}$. In whole min- or max-quantizers, the number of codeword index occurrences is computed using the probability factor $\Pr[\bullet]$. $N_{\min}$ or $N_{\max}$ indicating the number of leaf nodes in balanced-tree (i.e., the size of color clusters) equals both the $CIF_{\min}$- and $CIF_{\max}$-related feature dimensionalities.

Fig. 2 Example illustrating number of color clusters **a** 8 and **b** 16 in the color truncation process
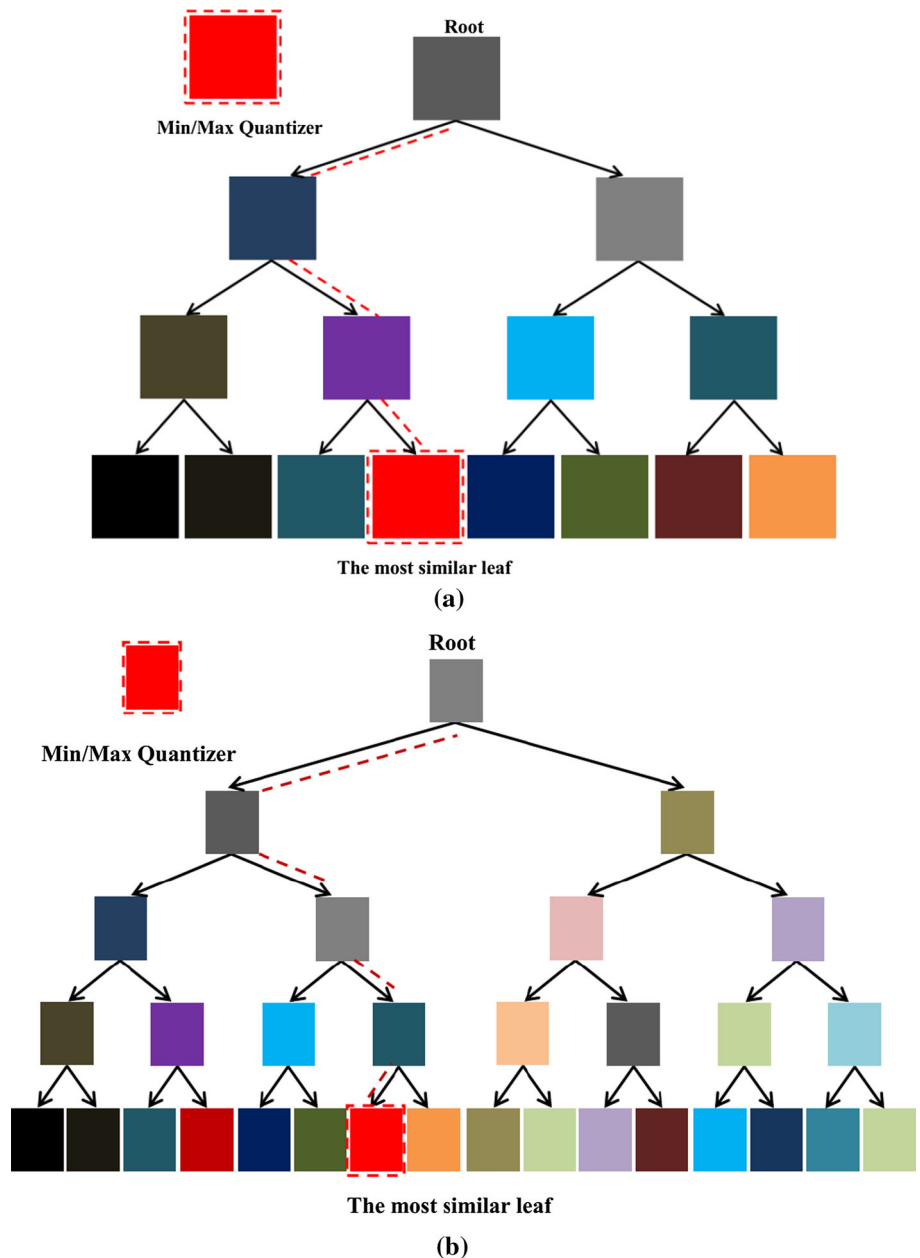
3.1.2 Color and edge directivity descriptor (CEDD)

This section explains about the color and edge directivity descriptor and its structural components. For CBIR function, the invented universal feature image descriptor is characterized using CEDD. Hence, its effective storage basis and discrete mediocre size have significantly attained good system performance. An image is segmented into 1600 rectangular image parts using this CEDD. Then, the texture and color data are extracted through picking the related image blocks. Figure 3 indicates the flow process of CEDD descriptor. Color extraction unit: At first, several image blocks are formed through dividing the input image I. Then, the RGB values of each color unit contained in several image blocks are transformed into the HSV color space. Subsequently, for all that images, a fuzzy linking histogram is generated using a two-staged fuzzy system. The resultant 10-bins histogram is produced from the three mean HSV channels comprised in the fuzzy system output.

By that fact, the fuzzy system has considered hue (H), saturation (S) and channel value (V) as its input and they are explained as follows: 8, 2 and 3 fuzzy regions are formed through dividing the hue (H), saturation (S) and channel value (V). A set of 20 rules is applied to facilitate the fuzzy system output. Then, the limit is fixed from 0 to 1 for generating a crisp value. Ultimately, the first stage of histogram having 10-bins is created using this generated crisp value. The membership functions are illustrated in Fig. 4. Usually, 8 fuzzy regions are formed after dividing hue. Figure 4a indicates the borders of corresponding 8 fuzzy regions. Figure 4b indicates the 2 fuzzy regions formed after dividing saturation. Similarly, Fig. 4c indicates the 3 fuzzy areas formed with the division of channel V. Furthermore, the 2 fuzzy regions generated through separating channel V are indicated in Fig. 4d. To each of the seven colors (including white, gray and black), the brightness value is supplemented at the Takagi–Sugeno–

Kang (TSK) fuzzy linking system's second stage. Again, the image block is assigned with S and V mean values (the given fuzzy inputs). Ultimately, the crisp values containing 3-bin histogram are obtained as the final output. This represents that the color either can be dark-hued, normal or light. At last, a 24-bin histogram is produced by means of pooling both first- and second-stage histograms (i.e., the two outputs).

3.1.2.1 Texture extraction unit In color and edge directivity descriptor (CEDD), the texture has played a main feature role. Initially, the YIQ color space is formed through converting the image block. Subsequently, the texture unit is obtained followed with the conversion of given image block. Moreover, the MPEG-7 Edge Histogram Descriptor-EHD suggested five digital filters are employed to this texture unit. This process is done for grouping the extra non-edge filter with edges (isotropic, 135 diagonal, 45 diagonal, horizontal and vertical). Then, the non-edge filter is used for the segmentation of each given image block into 4 sub-blocks. Thirdly, each image block has produced the specific edge types after performing the fuzzy mapping process among them. For each image block, it is significant to achieve 6-bin vector output as the final process. Thereby, the non-edge case is symbolized using one-bin vector, whereas the textures are symbolized using the other first five bin vectors. Label '1' is assigned for the relative bin, if identified that an image block has included any of the given edge types. If this is not the case, the binary image-block texture vector is generated by labeling the relative bin as '0.'

3.1.2.2 CEDD descriptor Firstly, this descriptor has obtained 144 bin vectors. Then, 24-bins containing six regions are formed through dividing these bin vectors. Furthermore, a divergent texture is illustrated using these six regions. The most relevant regions of 144 bins vector

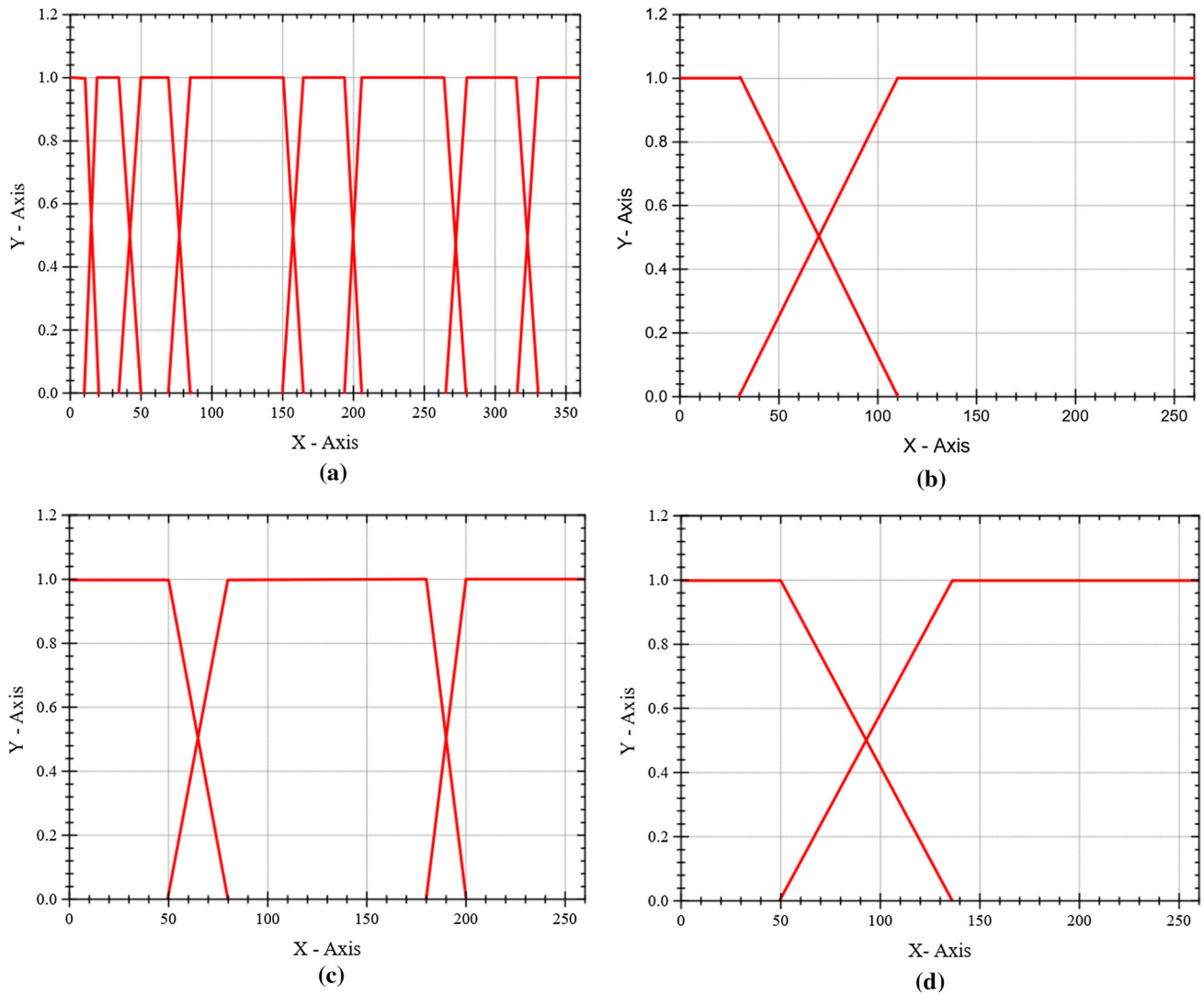Fig. 3 Block diagram of CEDD descriptor
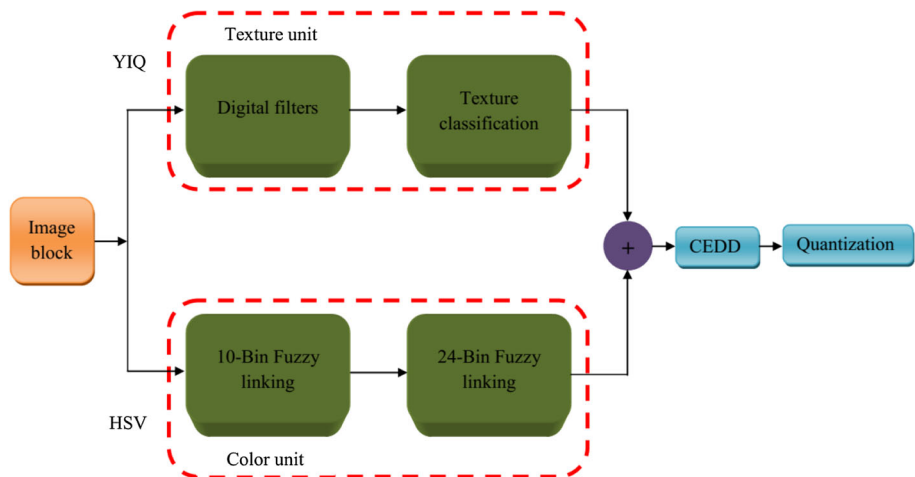
Fig. 4 Membership functions for **a** hue, **b** saturation, **c** value, **d** saturation and value for 24-bins expansion

are represented as '1' and the image block that computes the 24-bins color histogram were used to fill these relevant regions.

Followed by this, the image descriptor is generated by adding all together the whole image-block descriptors. Ultimately, about 8 pre-decided levels are formed through quantizing and standardizing these vectors. Followed with the completion of this process, the visual contents of images are characterized in a distinctive and compressed style by the formulation of EDD descriptor of an image.

3.2 Local feature extraction

3.2.1 Local binary pattern

Local structure of an image is defined using a nonparametric descriptor called LBP (Ojala et al. 2002). An operator value is assigned to each pixel of an image. This value is obtained through considering neighborhood round pixel (neighborhood threshold value fixed using the center pixel value). When compared to the center pixel value, if found the neighboring pixel rate is higher, then the value of this pixel is fixed to 1; otherwise, the value is fixed to 1. LBP formulation is described as follows: The resultant decimal form of LBP obtained for the input pixel at (u_c, v_c) is expressed as:

$$h = (i_p - i_c) \tag{5}$$

$$LBP(u_c, v_c) = \sum_{p=0}^{p-1} sign(h) \cdot 2^p \tag{6}$$

The central pixel located in the neighborhood circle of radius R having its gray-level values is indicated as i_p and i_c, respectively. Here, the surrounding pixels are denoted as

p. Equation (7) indicates the numerical expression of function $S_c(x)$.

$$S_c(x) = \begin{cases} 0, & \text{if} \quad x < 0 \\ 1, & \text{if} \quad x \geq 0 \end{cases} \tag{7}$$

The limited structural features surrounding the static pixel are characterized using the unit value obtained from the binary value. The way of 256 various patterns performance can be indicated using LBP image histogram. However, the whole image structure is defined using this pattern distribution. More information loss is avoided by us through selecting only reduced number of patterns (i.e., the uniform pattern included in LBP histogram alone is selected). The other form of unvarying pattern is LBP pattern. In other words, during binary classification, this LBP pattern has undergone maximum dual bitwise switching varying from 0 to 1 or 1 to 0. For instance, a non-uniform shape is observed with four changes (10110111) and uniform shape is seen with two changes (11100000), respectively. However, there includes possibly 198 unique non-uniform patterns and 58 different uniform patterns in 8-bit LBP representation. Thereby, the histogram representation requires only 59 bins instead of considering 256 bins for texture representation.

3.2.2 Speeded-up robust features (SURF)

Herbert Bay et al. (2008) have explained deeply an inventive scale- and rotation-invariant interest point detector and descriptor called SURF algorithm. Detector and descriptor are the two main phases included in SURF algorithm.

3.2.2.1 Detector An integral image and a basic Hessian matrix approximation were basically used by this detector. Four main steps included in this detector are: (a) integral image, (b) Hessian matrix-based interest points, (c) scale-space representation and (d) interest point localization.

Step 1: Integral Image

Initially, the SURF method is improved through using integral images $M_\Sigma(u)$ to further enhance the speed of local feature extraction. The pixels of an input image M are all summed to represent the presence of an integral image $M_\Sigma(u)$ at a position $u = (u, v)^T$. However, a rectangular region formed using origin u is used to hold this input image M.

$$M_\Sigma(u) = \sum_{k=0}^{k \leq u} \sum_{l=0}^{l \leq v} M(k, l) \tag{8}$$

The process of integral image calculation is illustrated in Fig. 5. After completing the integral image computation, the rectangular area consisting sum of the intensities is

calculated using three additions. Hence, based on the rectangle size, the computational process may show changes.

Step 2: Hessian matrix-based interest points

Here, the interest points are determined through employing the Hessian matrix $H(i, \sigma)$. Equation (9) has defined the Hessian matrix $H(i, \sigma)$ in i at scale σ

$$H(i, \sigma) = \begin{bmatrix} L_{ii}(i, \sigma) & L_{ij}(i, \sigma) \\ L_{ij}(i, \sigma) & L_{jj}(i, \sigma) \end{bmatrix} \tag{9}$$

where the Gaussian second-order derivatives $\frac{\partial^2}{\partial l^2} g(\sigma)$ convolution is indicated as $L_{ii}(i, \sigma)$. The similarity for an image M at a given point i is indicated as $L_{ii}(i, \sigma)$ and $L_{jj}(i, \sigma)$. Furthermore, the approximation of $H(i, \sigma)$ is used by the SURF to minimize the computational cost.

$$H_{\text{approx}} = \begin{bmatrix} D_{ii} & D_{ij} \\ D_{ij} & D_{jj} \end{bmatrix} \tag{10}$$

The location having maximum determinants is analyzed to detect the blob-like structure. It is expressed in (11)

$$\det(H_{\text{approx}}) = D_{ii}D_{ij} - (wD_{ij})^2 \tag{11}$$

Here, the expression $\det(H_{\text{approx}})$ is balanced using the relative weight w. Thus, the relation among approximated Gaussian kernels and kernels can be improved using this relative weight by further enhancing the energy conservation.

Step 3: Scale-space representation

Interest points included in images are extracted from the scale space consisting of different filter size levels. In other words, this extraction process is done by applying Gaussian approximation filters to each of the filter size levels. Also, the scale-invariant feature transform (SIFT) algorithm has also adopted the scale-space representation notion.

By that fact, the image size is gradually reduced using SIFT algorithm. Conversely, the integral images are used by the SURF algorithm to allow filter upscaling in a reduced cost. It is evident that the components without requiring any aliasing and having high-frequency with
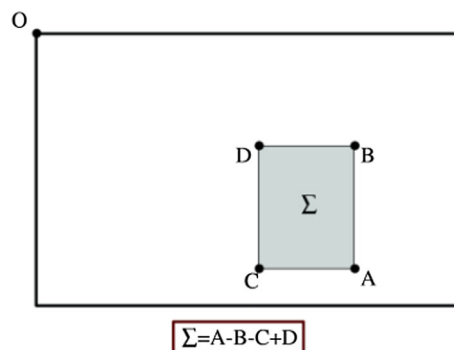

Fig. 5 Representation of integral image

more computational efficiency are offered by the SURF algorithm.

Step 4: Localization of Interest point

Using three neighborhood pixels ($3 \times 3 \times 3$ neighborhood scales), the non-maximum suppression (NMS) is applied to perform the interest point detection process. NMS has considered the feature points from the Hessian matrix with maxima determinant points. Figure 6 indicates the interest points detected from the input images using SURF algorithm.

3.2.2.2 Descriptor This descriptor requiring every interest points should carry their own indicator to perform the assignment of invariability to the interest points. Two most significant steps are included in the SURF descriptor process: (a) assignment of orientation and (b) the descriptor using sum of Haar wavelet responses.

Step 1: Orientation assignment

Considering the image rotation, the invariability of interest points is recognized through applying image orientation. For the Gaussian-weighted Haar wavelet responses summation, the dominant vectors are detected to compute the orientation. This process is done based on split circle region of sliding window using *pi/3* (Schnorrenberg et al. 2000). This is due to the fact that the interest points corresponding directional property and strength are included in vertical and horizontal responses of Haar wavelet transform. Based on the image rotation, the most significant image points are represented effectively using image orientation.

Step 2: Descriptor using the sum of Haar wavelet responses

The interest points of the descriptor are identified using the square regions obtained around the interest points and the selected orientation from the orientation assignment step. However, the 4×4 smaller sub-regions are formed through splitting each square region. At a sample points having 5×5 regular space, the vertical Haar wavelet response d_j and the horizontal Haar wavelet response d_i are determined for each sub-region. Then, the 4D description

Fig. 6 Interest points detected from the input images using SURF feature

vector is formed through utilizing d_i and d_j of each subregion as follows

$$v = \left(\sum d_i, \sum d_j, \sum |d_i|, \sum |d_j| \right) \qquad (12)$$

3.3 Multi-level matching scheme

This section explains the process of a multi-level matching scheme based on content-based medical images. From database D^D, the query image Q is retrieved depending on the matching scheme. This work has used both the local and global features (described in Sect. 3) to enhance the speed of query processing unit in relevant image searches. It is also important to use similarity matching (SM) along with these features. In order to handle good shape representation, the benefits of both local and global features are adopted by the SM. Thereby, the objects from the query images are identified directly. Subsequently, the database images (retrieved similar images) have been verified to determine whether the accurate words for an input query image are placed or not. The specific similarity measures have provided better outcome and motivate us to hybridize both the global and local features.

To compute global features Q_{Global} and local feature Q_{Local}, the query image Q is considered in matching scheme. Likewise, both the global feature vector $I\left[I^1_{\text{Global}}, I^2_{\text{Global}}, \ldots, I^n_{\text{Global}}\right]$ and local feature vector $I\left[I^1_{\text{Local}}, I^2_{\text{Local}}, \ldots, I^n_{\text{Local}}\right]$ are represented for each image of the database. The main aim of this scheme is to choose n number of optimal (best) images that are showing resemblance to the specific input query image. Hence, the distance between the input query image and image contained in the database (DB) is measured to select n top matching images. The steps given below explain the multilevel matching process.

Step 1: Consider Q, Q_{Global} and Q_{Local}

Step 2: To the database image features ($I\left[I^1_{\text{Local}}, I^2_{\text{Local}}, \ldots, I^n_{\text{Local}}\right], I\left[I^1_{\text{Global}}, I^2_{\text{Global}}, \ldots, I^n_{\text{Global}}\right]$), the query image features $Q[F](Q\left[Q_{\text{Local}}, Q_{\text{Global}}\right])$ are matched.

Step 3: Initially, the global similarity between the query image and input image is computed in this matching scheme. Equation (13) indicates the global similarity between images.

$$S_{\text{Global}} = d_{ij} = f\left(Q^{\text{Global}}, I^{\text{Global}}\right) \qquad (13)$$

Here, the database image D^D and query image Q related using a Euclidean distance are indicated as $f\left(Q^{\text{Global}}, I^{\text{Global}}\right)$.

Step 4: In between database image and query image Q, the local similarity evaluated using ED is expressed as:
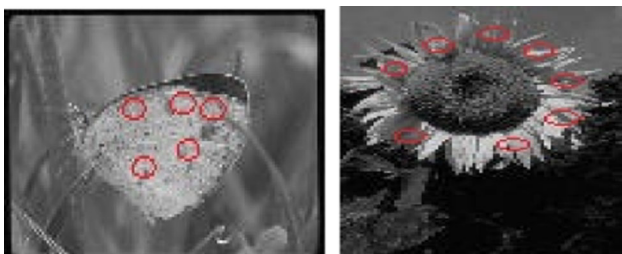
$$S_{\text{Local}} = d_{ij} = f\left(Q^{\text{local}}, I^{\text{local}}\right) \tag{14}$$

$$d_{ij} = \sqrt{\sum_{i=1}^{n}\left(Q^{\text{local}} - I_i^{\text{local}}\right)^2} \tag{15}$$

Step 5: Both distance measures are given by Eqs. (13) and (14) and provide the normalized distance. Ultimately, the global and local similarity synthesized using the hybrid similarity measure is as follows:

$$S_{\text{hybrid}} = C \times S_{\text{Global}} + (1 - C)\, S_{\text{Local}} \tag{16}$$

Here, the global and local similarity measures corresponding significances are adjusted using the weight C. Based on the user's expectations, this hybrid local and global similarity measures are balanced by means of altering the value of weight (C). Hence, it is evident that the system has offered good user flexibility.

Step 6: For the query image, the top n best images are selected from the database images through sorting the hybrid feature score value. The operations of proposed multi-level matching scheme are indicated in Fig. 7.

4 Proposed approach

This work mainly aims to introduce a new content-based medical image retrieval system based on the hybrid similarity measure and multi-level matching scheme. Two significant modules included in the proposed system are: (a) feature extraction and (b) multi-level matching. Firstly, the global and local features are extracted from the images contained in DB. This process is done through converting the color image into a grayscale image. Then, the multi-level matching-based similarity measure is performed to retrieve all the relevant images with respect to the given query image from the database DB. Based on the given query image, the most relevant images are retrieved initially only by means of using global similarity. Secondly, the local similarity measure is used to filter out the too far relevant images from the query image. Figure 8 illustrates the entire implementation process of the proposed framework. Algorithm 1 has explained the step-by-step working procedures of the proposed framework.

Algorithm 1: Proposed CBIR system

Input: Query input image,

Output: Retrieval list

1. Convert the database image into Grayscale images.
2. Extract the global features using CRF and CEDD.
3. Extract the local features using LBP, SURF.
4. Steps 1, 2 and 3 are repeated for an Input query image Q.
5. Based on equation (13), the global similarity is compared to match the database image with the given query image.
6. Sort the database image in ascending order and add the first predefined $R_{\text{Re}t}^1$ images into the first-step retrieval list. When compared to the number of images in the database D^D, the retrieved images $R_{\text{Re}t}^1$ are smaller in number. However, based on the images in database, the determination process is performed.
7. Using Eq. (14), the local similarities among the retrieved image and query image is computed. Then, the hybrid similarity of equation (16) is used to arrange the retrieved images in ascending order.
8. To the retrieval list of second step, the relevant images are added until appending all the user-defined R_{ret} images. But, when compared to $R_{\text{Re}t}^1$, the user-defined R_{ret} should be higher.
9. Return to the users, the initially obtained R_{ret} images.

---

5 Results and discussion

In this section, we discuss the result obtained from the proposed hybrid similarity-based multi-level matching scheme for color image retrieval system. Implementation is done using MATLAB® version (R2017a). Intel Core i5 processor with speed 1.6 GHz and 4 GB RAM which is equipped in the windows machine is used by the proposed technique to perform its operation. A general purpose database called SIMPLIcity image database (http://wang.ist.psu.edu) (Wang et al. 2001) is adopted to test the proposed CBIR system. Conversely, the images are retrieved through adopting a medical image database. However, the elephants, dinosaurs, buses, villages and African people were categorized to form 10 semantic groups. About 100 sample images are included in each individual category. Another one is the gastro-intestinal database (http://www.gastrolab.net) which is comprised of endoscopic images of gastro-intestinal disorders, partially shown in Fig. 9. The samples images that relate to cancer affected regions are partially shown in Fig. 9.

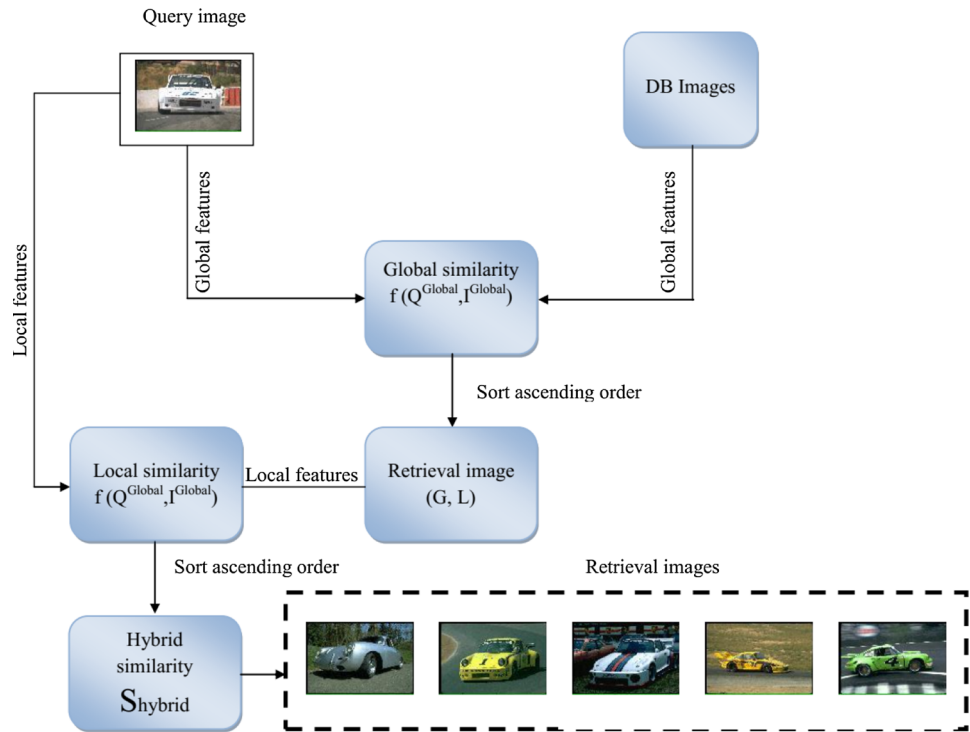**Fig. 7** Multi-level matching scheme used in proposed model

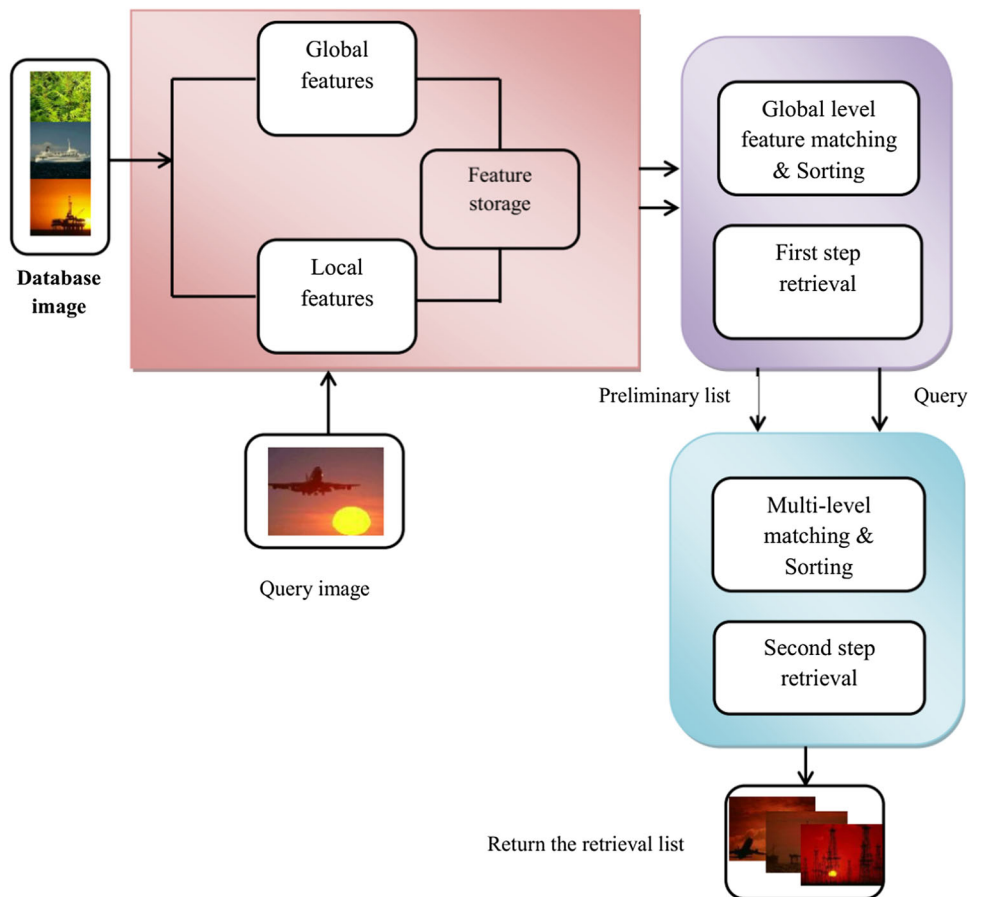**Fig. 8** Flow diagram of the proposed system

Fig. 9 Sample images used in experimentation **a** planet, **b** sunrise from a general database and **c** small bowel category from the medical database
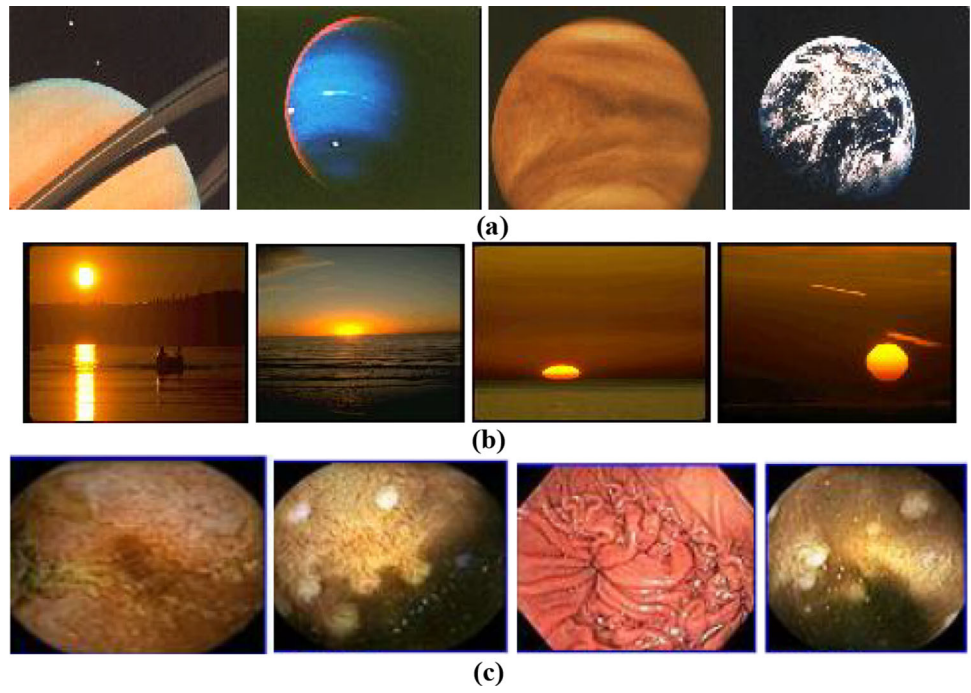

(a)

(b)

(c)

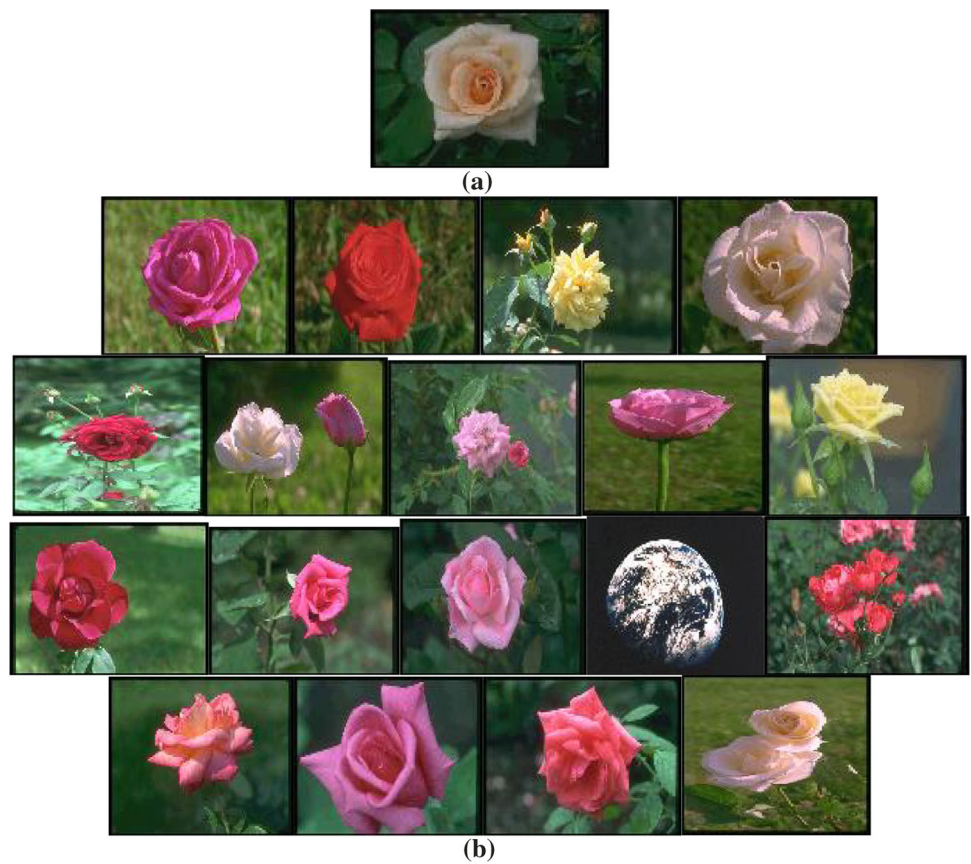**Fig. 10** Retrieved results for general database **a** query image 'rose' and **b** retrieved images

(a)

(b)

Fig. 11 Retrieved results for general database **a** query image 'butterfly' and **b** retrieved images

(a)

(b)

5.1 Performance measures

Some of the commonly known evaluation metrics such as F-measure (F), recall (R) and precision (P) are applied to analyze the system performance. Ratio of total number of recovered images to the number of recovered images (such as query image) defines a precision rate. Ratio of total number of query images in the database to the number of recovered images defines the recall rate. F-measure (F) is obtained by taking the harmonic mean of precision and recall rate. Numerical expressions of (P), (R) and (F) are expressed as follows:

$$P = \frac{N^Q}{T^Q} \times 100 \tag{17}$$

$$R = \frac{N^Q}{D^Q} \times 100 \tag{18}$$

$$F = 2 \times \frac{P.R}{P + R} \tag{19}$$

Here, the number of images in the database relevant to the query image Q is indicated as D^Q, the total number of retrieved images as T^Q, and the number of relevant images retrieved from the database is denoted as N^Q, respectively.

In order to determine the weight C influence, the performance metrics called 'area under the precision-recall curve' (AUC) is applied as follows:

$$AUC(C) = \sum_{i=2}^{R_{max}} \frac{(P_C(i) + P_C(i-1)) \times (R_C(i) - R_C(i-1))}{2} \tag{20}$$

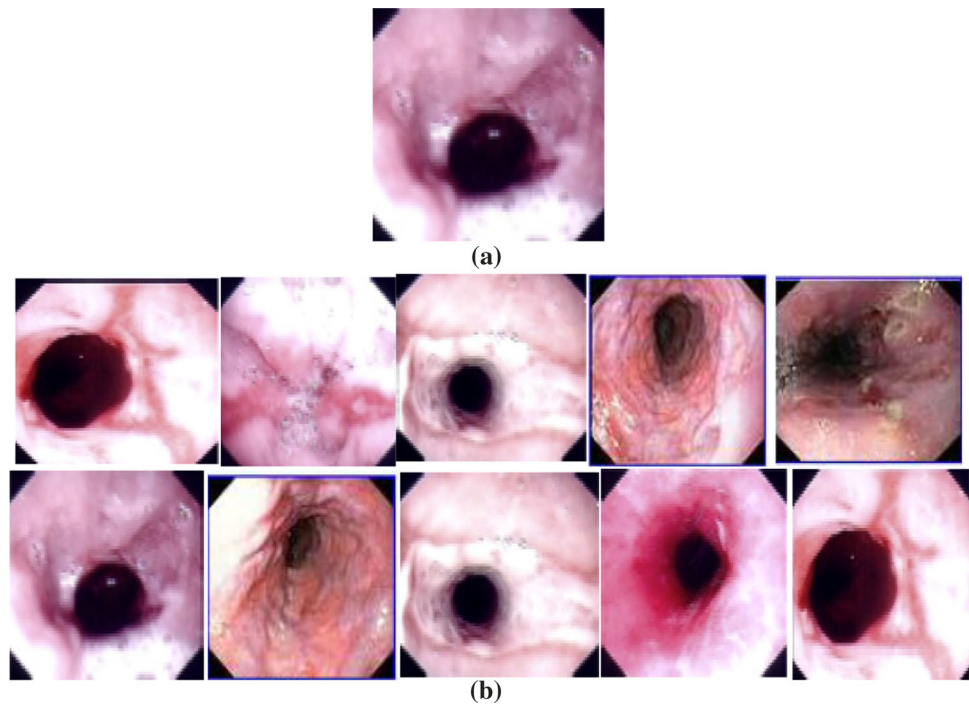**Fig. 12** Retrieved results for medical database **a** query image 'Oesophagitis' and **b** retrieved images

(a)

(b)

Fig. 13 Retrieved results for medical database **a** query image 'Cardia' and **b** retrieved images

(a)

(b)
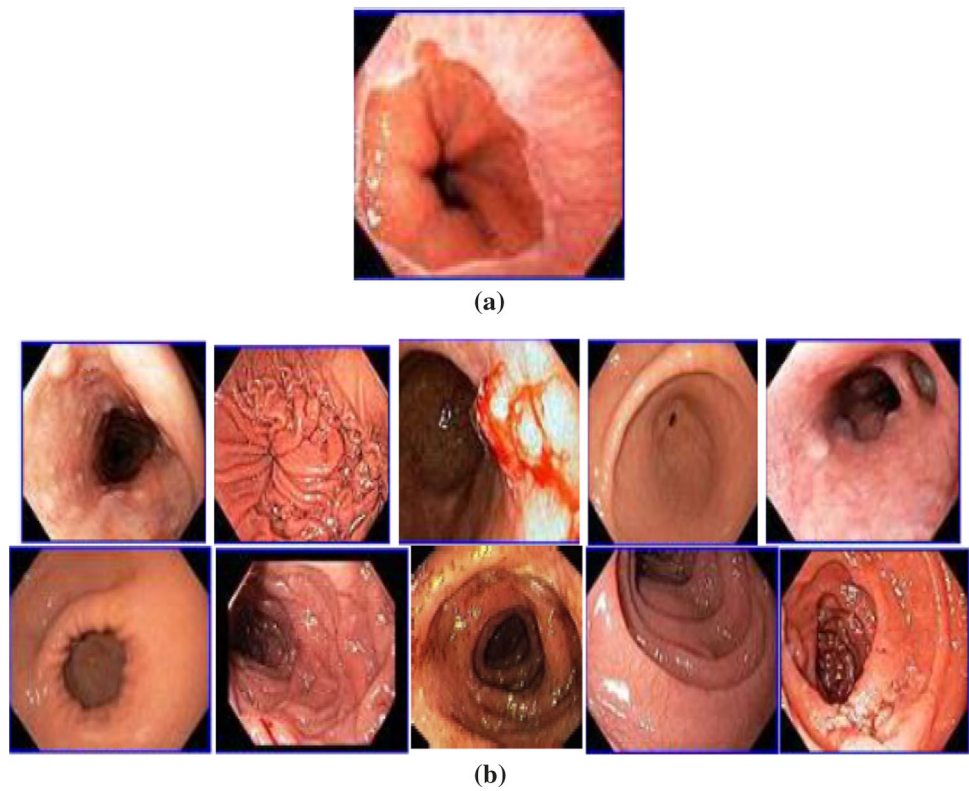
In Eq. (20), the terms $R_C(i)$ and $P_C(i)$ indicate the recall and precision values with the ith image retrieval and the maximum number of images retrieved is denoted as $R_{\max}$.

5.2 Experimental results

Different performance metrics such as recall, precision and F-measure are used to analyze the performance of proposed medical image retrieval system. The retrieved images of

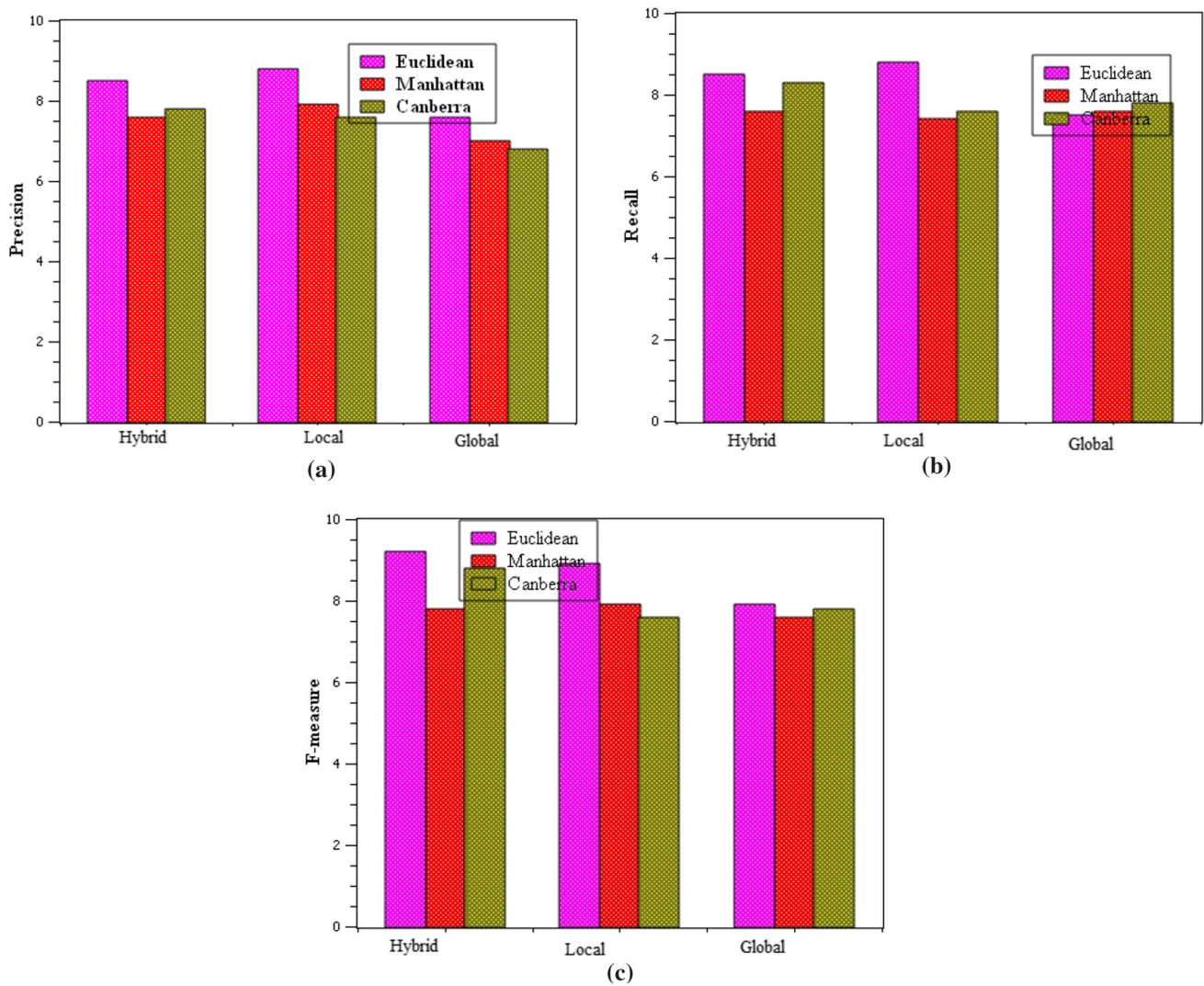**Fig. 14** Results obtained with general database on using different distance formulas in multi-level matching scheme **a** precision, **b** recall and **c** F-measure (for one-step retrieval process)

the proposed model obtained based on the input query image are shown in Figs. 10, 11, 12 and 13.

5.3 Comparative analysis

Improvement in image retrieval rate is considered as the main objective of proposed system. In between the images, their overall similarities are computed through applying Eq. (13), which provides the similarity measure. This computation is done only after achieving global features with the usage of both local and global features. Most relevant images indicate the images having minimum distance. In query-based image retrieval process, the multi-level matching step has played a significant role. However, the matching scheme includes various distance measure types. In this paper, the Euclidean distance (ED) measure is applied to enhance the image matching system. The two

conventional distance measures such as Canberra Distance (CD) and Manhattan distance (MD) are used to compare the effectiveness of proposed ED over two databases (Figs. 14, 15, 16, 17).

Different distance measures used in one-step retrieval steps of proposed approach are shown in Figs. 13 and 14. When used CD and MD, the maximum precision rate obtained by the proposed model is 75% and 80%; but the proposed model has achieved 88% higher precision rate on using ED shown in Fig. 14a. The F-measure and recall rate obtained on using one-step retrieval process with different distance measures are shown in Fig. 14b, c. Similarly, the precision, recall and F-measure rate obtained using one-step retrieval process for medical database are illustrated in Fig. 15. Different distance measures used by the one-step retrieval process of the proposed model are illustrated in Figs. 16 and 17.
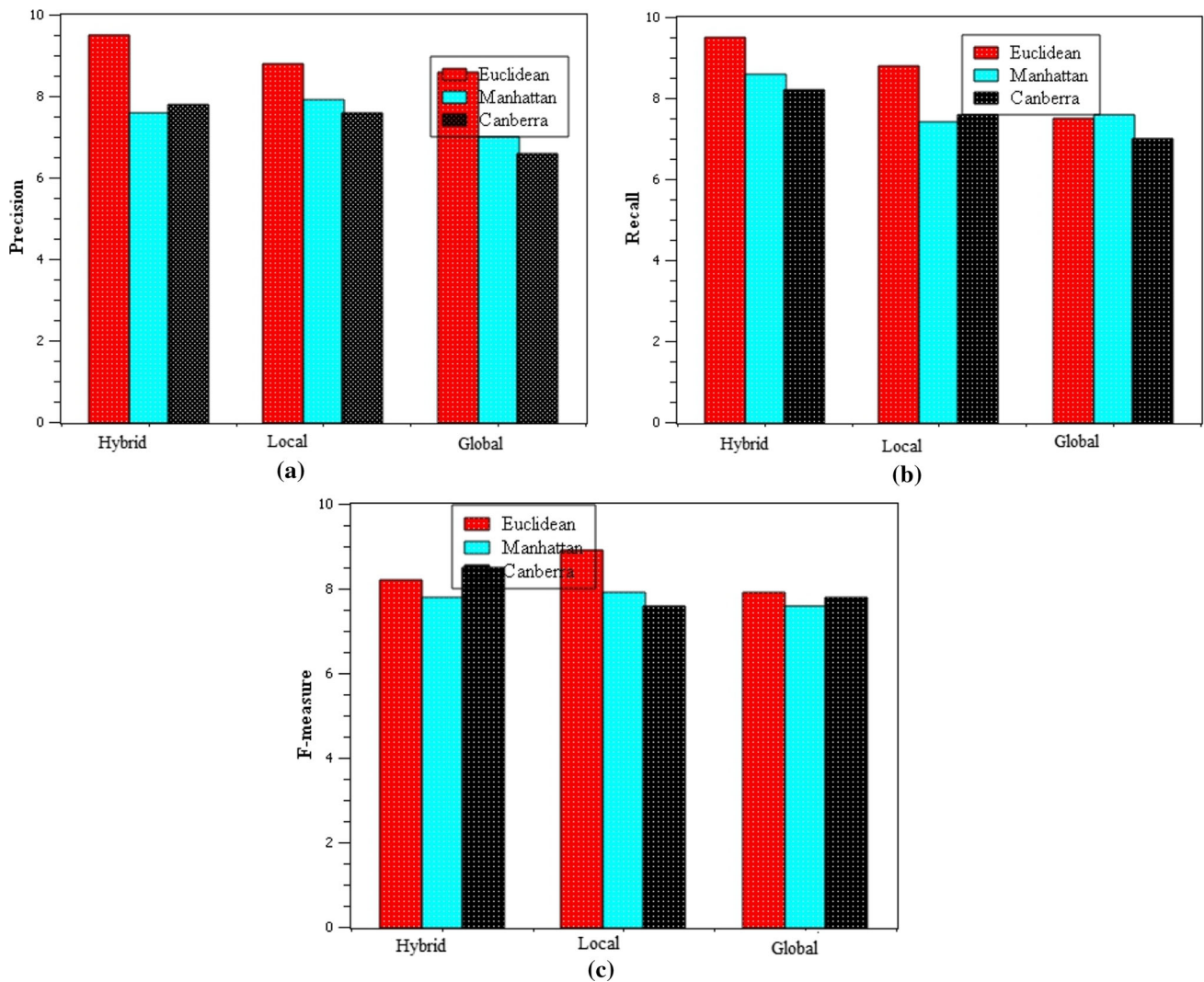
Fig. 15 Results obtained with medical database on using different distance formulas in multi-level matching scheme **a** precision, **b** recall and **c** F-measure (for one-step retrieval process)

Figure 16a indicates the higher precision rate (92%) achieved by the proposed model by using ED measure in two-step retrieval process. Depending on different distance measures, the two-step retrieval process applied in proposed model achieving the recall rate is plotted in Fig. 16b. From the results, it is evident that the proposed model has yielded better performance compared to other traditional approaches.

For two-step retrieval process, the F-measure rate achieved by the proposed model is shown in Fig. 16c. Similarly, the precision, recall and F-measure rate obtained using two-step retrieval process for medical database are illustrated in Fig. 17. From the results, it is evident that the ED measure used on multi-level matching scheme has yielded better performance compared to other conventional measures.

For general purpose database, the AUC values generated are indicated in Fig. 18. These values are produced by varying the value of weight from 0 to 1 using 0.02 increments in terms of considering precision and recall curves of the MLM-hybrid technique. For each dataset, the significances of both local and global information are balanced using an optimal weight value. MLM-hybrid method will consider that setting the weight values to 0.3 is a suitable choice to test the CBMIR system. Based on feature extraction method, the comparative analysis performed for the retrieval process is shown in Fig. 19. From the figures, it is evident that higher precision rate is achieved by our proposed approach compared to other traditional approaches.

For one-step retrieval system, we have to utilize the global features and query image alone. Followed by this, the similarity score is obtained by combining the global
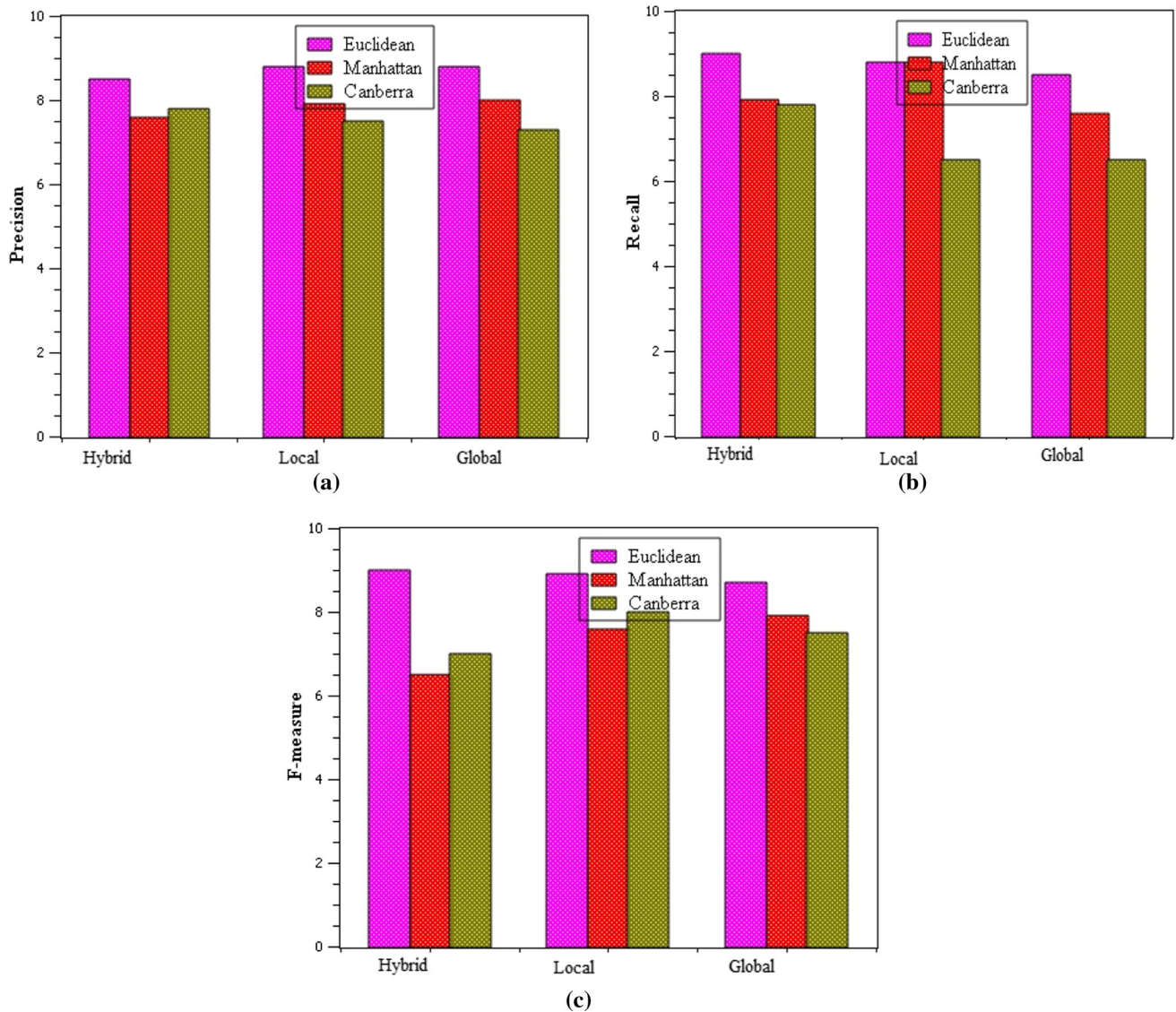
Fig. 16 Results obtained with general database on using different distance formulas in multi-level matching scheme **a** precision, **b** recall and **c** F-measure (for two-step retrieval process)

and local features. Finally, the corresponding image is retrieved using the similarity score. In this work, the global and local features-dependent hybrid similarity measure is adopted.

5.3.1 General database

For the query image 'butterfly,' the performance of one-step retrieval step is illustrated in Table 1. When used the MLM-local features-based image retrieval and MLM-global features-based medical image retrieval process, the proposed model has achieved 86% precision rate. But, the proposed model has achieved 90% of precision rate after using a multi-level matching scheme.

Form the analysis, it is inferred that, when compared to image retrievals based on both global and local feature-based methods, the better image retrieval performance is yielded by the hybrid MLM methods. For a given query image 'butterfly,' the two-step retrieval performance is analyzed, as shown in Table 2. As and when compared to the computational efficiency of one-step image retrieval process, the two-step retrieval process has achieved good performance. Furthermore, the efficiency of two-step retrieval process not only stops with providing accurate relevant retrieval outcomes but also supports the users by providing rapid query response through enabling the system performance. For shape representation, the benefits of both local and global features are adopted during similarity measure computation. As and when compared to the global

(a)

(b)

(c)
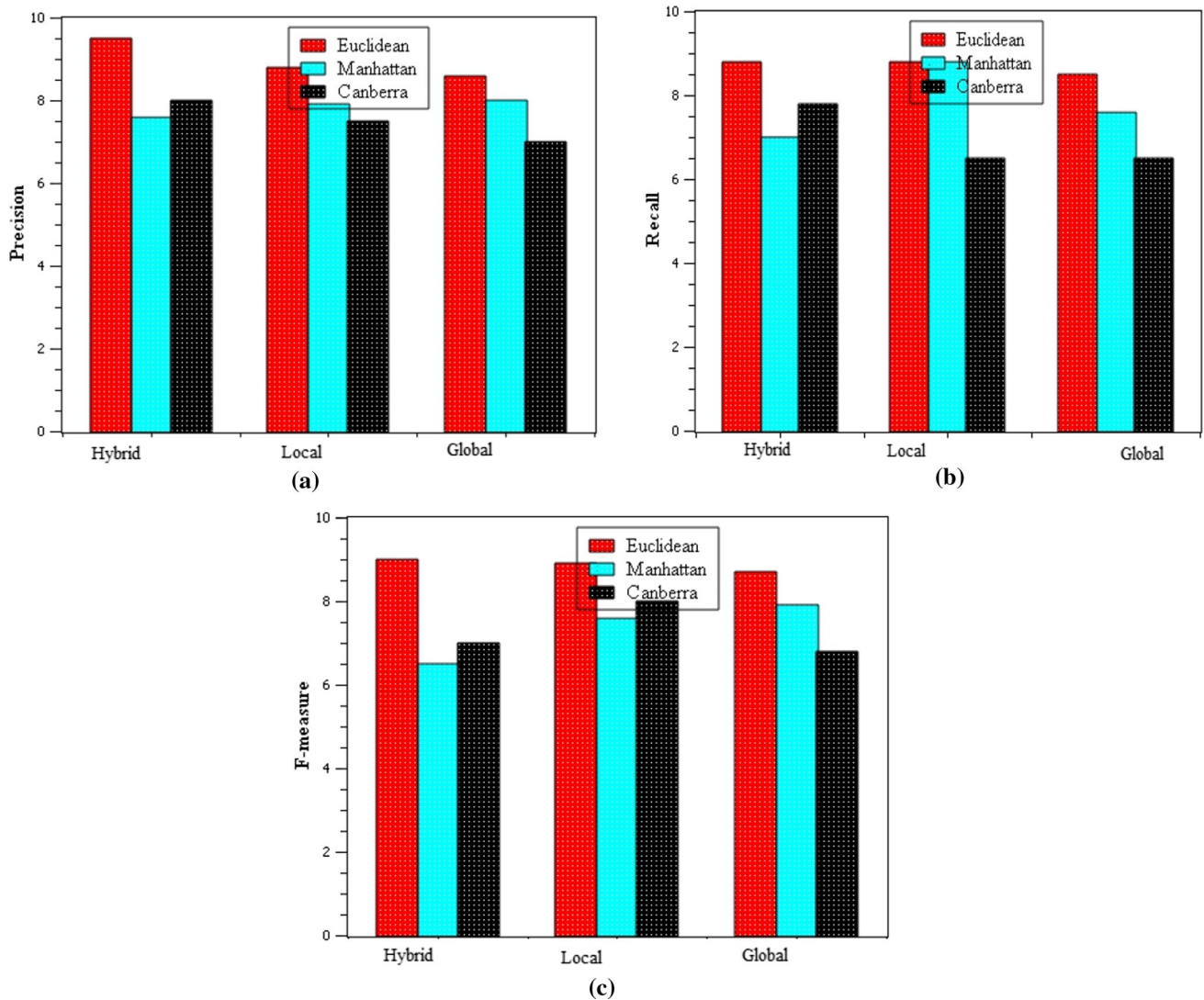
Fig. 17 Results obtained with medical database on using different distance formulas in multi-level matching scheme **a** precision, **b** recall and **c** F-measure (for two-step retrieval process)
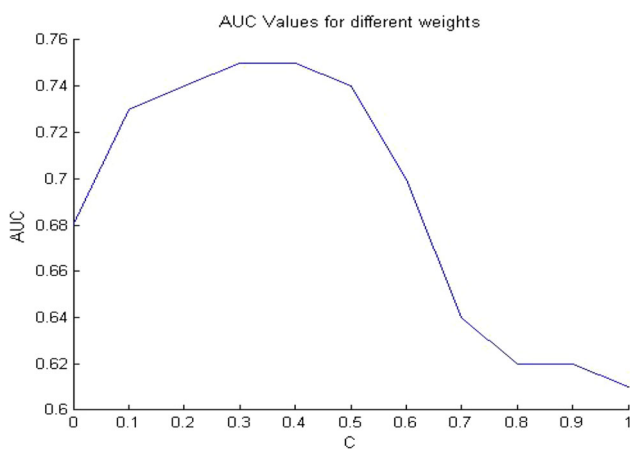
Fig. 18 Different weights of AUC measure in general image database

feature-based image retrieval and individual local feature-based image retrieval, the proposed approach has achieved higher precision rate of 93% (Table 6). For the query image 'rose,' the performance of two-step retrieval and one-step retrieval process is indicated in Table 3.

From the table, it is observed that the one-step retrieval process has achieved lower retrieval performance than the two-step retrieval process (Table 4). For the query image 'Oesophagitis,' the retrieval performance achieved with two-step retrieval and one-step retrieval process is shown in Tables 5 and 6. From the results, it is evident that higher precision value is achieved with our proposed model than the other global features and individual local-based retrieval system.

Fig. 19 Feature extraction-based proposed method comparison **a** general dataset and **b** medical database

5.4 Comparison with other published approaches

In this section, the efficiency of proposed approach is analyzed with different existing works. In medical image retrieval system, the existing works of Kumar et al. (2014)

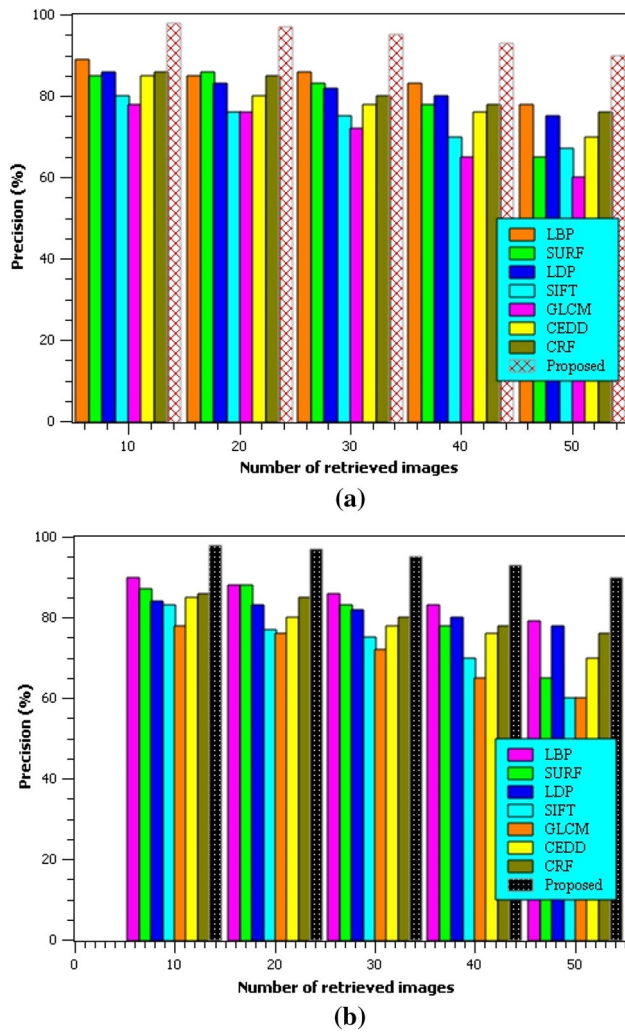and Srinivas et al. (2015) have proved the whole strength on image retrieval. Based on various representations, they have characterized the global and local features of an image. One can observe easily from Table 7 that the better performances are yielded using our proposed approach (Kumar et al. (2014)). This goodness is observed because the visual features of an image are described clearly using these methods. Compared to the aforementioned conventional systems, our method has shown better image retrieval performance. Dictionary learning method was used on the clustering method-based image retrieval process of Srinivas et al. (2015).

In which, the classical dictionaries are matched with a query image to further use the Orthogonal Matching Pursuit (OMP) algorithm for identifying the sparsest representation dictionary. However, the multi-modality medical images were retrieved using the graph-based approach of Kumar et al. (2014). In addition, the comparison was made with some published results. The precision rate achieved on using the techniques of El Alami (2011) and Kumar et al. (2014) is 79.2% and 71%; but the proposed approach has achieved higher precision rate of 93% on using multi-level matching scheme (Table 7). From the result, it is evident that the proposed approach has yielded higher precision rate compared to other traditional methods.

6 Conclusion

In this study, a multi-level matching scheme is introduced for retrieval of image based on a hybrid feature similarity integrating local and global features for an input image. During retrieving the target objects, the complexity is reduced effectively by the local features. For an image, the whole data such as color, texture and shape are captured using global features. Merits and demerits are equally balanced in both global- and local-based features. But, the discrimination ability of local-based features is relatively high as and when compared to the global-based features. To ease the retrieval process difficulty, the similarity measures were adopted in this work to hybridize both the

Table 1 Performance efficiency of one-step retrieval process in a general database for a given query image 'butterfly'

Methods	Number of retrieved images																	
	10			20			30			40			50			60		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	85	86	86.6	78	74	75.8	72	73	70.6	67	64	55.9	61	60	58.7	55	62	57.6
MLM-global	75	74	73.6	73	74	71.4	68	68	68	68	64	62	58	59	58.4	54	58	55.4
MLM-hybrid	89	87	87	85	78	78.5	74	77	74.5	69	70	68.5	66	69	64.6	56	63	56.8

Table 2 Performance efficiency of two-step retrieval process in a general database for a given query image 'butterfly'

Methods	Number of retrieved images																	
	10			20			30			40			50			60		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	84	86	84	85	82	84	85	77	77	67	66	72	68	58	62	62	52	56
MLM-global	83	82	82	83	74	77	81	74	75	70	64	65	66	63	64	67	62	63
MLM-hybrid	90	88	87	86	78	81	88	78	79	76	69	74	69	68	67	55	63	58

Table 3 Performance efficiency of one-step retrieval process in a general database for a given query image 'rose'

Methods	Number of retrieved images											
	10			20			30			40		
	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	84	84	84	82	81	83	81	76	77	76	65	71
MLM-global	82	82	81	78	76	78	76	72	74	69	65	66
MLM-hybrid	90	87	87.5	87	85	85.88	87	84	84	84	78	79.55

Table 4 Performance efficiency of two-step retrieval process in a general database for a given query image 'rose'

Methods	Number of retrieved images											
	10			20			30			40		
	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	85	83	83	82	82	81.5	79	76	76.5	76	65	68.8
MLM-global	83	82	80.56	83	75	76.9	72	72	72	69	63	65.7
MLM-hybrid	91	88	87.96	88	84	85	86	83	84	85	78	81

Table 5 Performance efficiency of two-step retrieval process in a medical database for a given query image 'Oesophagitis'

Techniques	Number of retrieved images											
	5			10			15			20		
	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	84	82	82.96	83	82	81	80	76	77	77	70	73.1
MLM-global	82	80	81.95	79	79	80	73	72	72	72	69	70
MLM-hybrid	88	86	86.98	85	81	82.95	84	83	83	82	79	80

Table 6 Performance efficiency of one-step retrieval process in a medical database for a given query image 'Oesophagitis'

Techniques	Number of retrieved images											
	10			20			30			40		
	P	R	F	P	R	F	P	R	F	P	R	F
MLM-local	86	87	85.51	82	82	82	78	76	75.7	75	69	70.86
MLM-global	84	85	83.57	79	77	77.10	74	72	71.8	72	67	67.95
MLM-hybrid	93	90	89.50	89	85	86	87	83	84.9	85	78	80.46

Table 7 Comparative analysis of the state-of-the-art methods

Techniques	Precision (%)	Recall (%)	F-measure (%)
Srinivas et al. (2015)	79.2	65.8	71.83
Kumar et al. (2014)	71	62	65.74
MLM-hybrid (proposed)	93	90	90.45

global and local features of an image. Using a medical image database, the performance of the proposed method is determined in terms of certain evaluation metrics such as F-measure, recall and precision. Also, one-step retrieval and two-step retrieval process are applied to evaluate the retrieval efficiency of proposed approach. Experimental results have evidently proved that higher precision rate of 91% and 92% is achieved with proposed approach over two different databases. In the future, we intend to extend our dataset scale, further normalizing the image tags, and enriching the query terms.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Al-Janabi S, Alkaim AF (2019) A nifty collaborative analysis to predicting a novel tool (DRFLLS) for missing values estimation. Soft Comput. https://doi.org/10.1007/s00500-019-03972-x

Al-Janabi S, Mahdi MA (2019) Evaluation prediction techniques to achievement an optimal biomedical analysis. Int J Grid Util Comput 10(5):512–527

Amato G, Carrara F, Falchi F, Gennaro C, Vadicamo L (2019) Large-scale instance-level image retrieval. Inf Process Manag, In press, corrected proof, Available online 29 Aug 2019, Article 102100

Anandh A, Mala K, Suganya S (2016) Content based image retrieval system based on semantic information using color, texture and shape features. In: 2016 International conference on computing technologies and intelligent data engineering (ICCTIDE'16), pp 1–8. IEEE

Argyriou V, Tzimiropoulos G (2017) Frequency domain subpixel registration using HOG phase correlation. Comput Vis Image Underst 155:70–82

Bagri N, Johari PK (2015) A comparative study on feature extraction using texture and shape for content based image retrieval. Int J Adv Sci Technol 80(4):41–52

Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359

Dubey SR, Singh SK, Singh RK (2014) Rotation and illumination invariant interleaved intensity order-based local descriptor. IEEE Trans Image Process 23(12):5323–5333

El Alami ME (2011) A novel image retrieval model based on the most relevant features. Knowl-Based Syst 24(1):23–32

El Alami ME (2014) A new matching strategy for content based image retrieval system. Appl Soft Comput 14:407–418

Gladis KPA (2019) Integration of global and local features based on hybrid similarity matching scheme for medical image retrieval system. Int J Biomed Eng Technol 31(3):292–314

Guo J-M, Prasetyo H, Chen J-H (2015) Content-based image retrieval using error diffusion block truncation coding features. IEEE Trans Circuits Syst Video Technol 25(3):466–481

Hongpeng Z (2019) Massive-scale image retrieval based on deep visual feature representation. J Vis Commun Image Represent, In press, journal pre-proof, Available online 6 Dec 2019, Article 102738

Iakovidou C, Anagnostopoulos N, Kapoutsis A, Boutalis Y, Lux M, Chatzichristofis SA (2015) Localizing global descriptors for content-based image retrieval. EURASIP J Adv Signal Process 2015(1):80

Korytkowski M, Rutkowski L, Scherer R (2016) Fast image classification by boosting fuzzy classifiers. Inf Sci 327:175–182

Kumar A, Kim J, Wena L, Fulham M, Feng D (2014) A graph-based approach for the retrieval of multi-modality medical images. Med Image Anal 18:330–342

Lai H, Yan P, Shu X, Wei Y, Yan S (2016) Instance-aware hashing for multi-label image retrieval. IEEE Trans Image Process 25(6):2469–2479

Li B, Lu Y, Li C, Godil A, Schreck T, Aono M, Burtscher M, Chen Q, Chowdhury NK, Fang B, Fu H (2015) A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries. Comput Vis Image Underst 131:1–27

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110

Montazer GA, Giveki D (2015) Content based image retrieval system using clustered scale invariant feature transforms. Optik 126(18):1695–1699

Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987

Peng X, Zhang X, Li Y, Liu B (2019) Research on image feature extraction and retrieval algorithms based on convolutional neural network. J Vis Commun Image Represent, In press, journal pre-proof, Available online 11 Nov 2019, Article 102705

Ramesh B, Xiang C, Lee TH (2015) Shape classification using invariant features and contextual information in the bag-of-words model. Pattern Recognit 48(3):894–906

Schnorrenberg F, Pattichis CS, Schizas CN, Kyriacou K (2000) Content-based retrieval of breast cancer biopsy slides. Technol Health Care 8(5):291–297

Shrivastava N, Tyagi V (2015) An efficient technique for retrieval of color images in large databases. Comput Electr Eng 46:314–327

Somasundaran BV, Soundararajan R, Biswas S (2020) Robust image retrieval by cascading a deep quality assessment network. Signal Process: Image Commun 80:115652

Srinivas M, Naidu rr, Sastry CS, Mohan CK (2015) Content based medical image retrieval using dictionary learning. J Sci Direct, pp 1–19

Sujatha K, Shalini Punithavathani D (2018) Optimized ensemble decision-based multi-focus imagefusion using binary genetic Grey-Wolf optimizer in camera sensor networks. Multimed Tools Appl 77(2):1735–1759

Sundararaj V (2019) Optimised denoising scheme via opposition-based self-adaptive learning PSO algorithm for wavelet-based ECG signal noise reduction. Int J Biomed Eng Technol 31(4):325–345

Sundararaj V, Muthukumar S, Kumar RS (2018) An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. Comput Secur 77:277–288

Tesfaye AL, Pelillo M (2019) Multi-feature fusion for image retrieval using constrained dominant sets. Image Vis Comput, In press, journal pre-proof, Available online 12 Dec 2019, Article 103862

Tian X, Zhou X, Ng WWY, Li J, Wang H (2019) Bootstrap dual complementary hashing with semi-supervised re-ranking for

image retrieval. Neurocomputing, In press, corrected proof, Available online 31 Oct 2019

Vinu S (2016) An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. Int J Intell Eng Syst 9(3):117–126

Vinu S (2019) Optimal task assignment in mobile cloud computing by queue based ant-bee algorithm. Wirel Pers Commun 104(1):173–197

Vipparthi SK, Nagar SK (2014) Color directional local quinary patterns for content based indexing and retrieval. Hum-Centric Comput Inf Sci 4(1):6

Wang J, Li J, Wiederhold G (2001) Simplicity: semantics-sensitive integrated matching for picture libraries. IEEE Trans Pattern Anal Mach Intell 23(9):947–963

Wang XY, Liang LL, Li YW, Yang HY (2017) Image retrieval based on exponent moments descriptor and localized angular phase histogram. Multimedia Tools Appl 76(6):7633–7659

Xiao Y,Wu J,Yuan J (2014) mCENTRIST: a multi-channel feature generation mechanism for scene categorization. IEEE Trans Image Process 23(2)

Zeng X, Zhang Y, Wang X, Chen K, Li D, Yang W (2019) Fine-grained image retrieval via piecewise cross entropy loss. Image Vis Comput, In press, corrected proof, Available online 1 Nov 2019, Article 103820

Zheng Y, Jiang Z, Zhang H, Xie F, Ma Y, Shi H, Zhao Y (2018) Histopathological whole slide image analysis using context-based CBIR. IEEE Trans Med Imaging 37(7):1641–1652