Research Article

An efficient content based image retrieval using enhanced multi-trend structure descriptor

S. Sathiamoorthy[1] · M. Natarajan[2]

Abstract

This paper contributes a new variant of multi-trend structure descriptor (MTSD) for efficient content based image retrieval. The proposed variant of MTSD encodes color/edge orientation/texture quantized values versus orientation of equal, small and large trends instead of color/edge orientation/texture quantized values versus equal, small and large trends. In addition, it also encodes color/edge orientation/texture quantized values versus average location of distribution of pixel values for equal, small and large trends at each orientation. To reduce the time cost of the proposed variant of MTSD with the preservation of its accuracy, the image is decomposed into fine level using discrete Haar wavelet transform and the fine level for the decomposition of an image is determined empirically. Comprehensive experiments are conducted using the benchmark Corel-1k, Corel-5k, Corel-10k, Caltech-101, LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI image datasets and the results evident that the proposed variant of MTSD achieves the state-of-the-art performance for natural, textural and biomedical image retrieval. Precision and recall are the measures used to measure the accuracy. Euclidean similarity measure is used to calculate the similarity information between query and target images.

Keywords Multi-trend structure descriptor · Euclidean measure · Local level structure · Precision · Recall

1 Introduction

Today, massive digital image collections are available in almost all domains like medicine, remote sensing, education, multimedia, geology, oceanography and astronomy. Since usage of these massive digital image collections are significantly increasing ever more in our daily life, searching for related images more robustly and efficiently is needed and is the key objective of image retrieval system. Generally, images are searched based on either keyword or visual contents or high level semantics of an image. The keyword based method has limitations like physical annotation of image which flops for massive digital image collections owing to the difficulty in keyword annotation for rich content in an image and lack of sufficient and distinctive discriminatory vocabulary. Since massive and diverse digital image collection involves more number of labors for annotation task and interpretation of keyword differs from labor to labor, physical annotation task results in inappropriate annotation for images [1, 2]. On the other side, researches based on the mapping of low level visual content with high level semantics is also receiving attraction over a decade [3–5]. However, since reducing the semantic gap between the computed low level features of an image and high level semantics is still highly challenging issue owing to semantic labels which fails in expressing the whole visual characteristics of an image. Hence, semantic based image retrieval is so far significantly limited in accuracy [6–8]. However, to address this semantic gap, researchers in the domain of computer vision are working towards biologically inspired feature for better discrimination of an image [7]. Thus, image retrieval based

✉ S. Sathiamoorthy, ks_sathia@yahoo.com; M. Natarajan, mnid2004@gmail.com | [1]Tamil Virtual Academy, Chennai, India. [2]Division of Computer and Information Science, Annamalai University, Annamalai Nagar, India.

on visual contents like color, texture and shape become booming over the past two decades and more vigorous research domain for the multimedia researchers.

The key component of content based image retrieval (CBIR) system is feature extraction and representation of an image [9–12]. In CBIR system, each image in the repository is represented using a feature vector. So, when the system receives query image, feature vector is computed for the image and is compared with the feature vectors of images stored in the repository. Afterwards, the system retrieves the images from the repository based on the high matches found for the feature vector of query image and the images in the repository. Accordingly, image feature extraction and representation plays noteworthy role in the success of image retrieval systems [9–12]. Hence, the extracted feature of an image should produce more accuracy with less storage and time cost as well that should be scaling, rotation, transformations and illumination invariant [7].

Generally, image features are extracted either at local or global level or both. Features characterizing the whole image is termed as global level feature whereas features characterizing the object or region is termed as local level features. Apart from the global or local level, distinctive discrimination capability of feature is strongly depends on the type of features used in the system [7]. However, local image features receives more attention owing to tolerant to illumination changes, occlusion, distortion and image transformations [13].

In recent years, researchers suggested feature combination methods for describing the image more discriminately. Along this direction, combination of various features is recommended by many researchers for better image retrieval and is the lively research in CBIR. However, finding the perfect combination of features or feature with predominant discrimination for image retrieval system is still scanty.

2 Related work

Many state-of-the-art visual features alone and combination of various visual features has been reported over the decades for CBIR and are described in this section. Feng et al. [14], global correlation and directional global correlation descriptor is used to characterize the color and texture features of an image and is used for CBIR system. Zeng [15] introduced local level structure descriptor which extracts color, texture and shape as a single unit for image retrieval. Williams and Yoon [16] described joint autocorrelogram (JAC) that extracts color, texture, gradient and rank and is used effectively for image searching. Color and edge directivity descriptor (CEDD) reported in Chatzichristofis

et al. [17] computes textures from the six-bin histogram of the fuzzy system and color from the 24-bin color histogram formed by the 24-bin fuzzy-linking system. Fuzzy color and texture histogram (FCTH) presented in Chatzichristofis et al. [17] characterize texture from the 8-bin histogram of the fuzzy system and color from 24-bin color histogram formed by the 24-bin fuzzy-linking system. The combination of CEDD and FCTH results in joint composite descriptor (JCD) [18] which comprises of color, texture and shape features for CBIR.

The texture, shape and spatial information is characterized as a single unit by the edge orientation autocorrelogram (EOAC) which is defined by Mahmoudi et al. [19] and it is reported that EOAC outperforms MPEG-7's edge histogram descriptor (EHD) owing to the computation of edge orientations and correlation between neighboring edges, and is translation, viewing position, illumination and small rotation invariant.

Scale invariant feature transform (SIFT) is a 3D histogram of locations and orientations and is robust to scaling and rotation [20], and it is employed in many image retrieval and classification systems. Inspired by the discrimination ability of SIFT, researchers presented several variants of SIFT with diverse ability. In continuation, fusion of SIFT and principal component analysis (PCA) is described by Ke and Sukthankar [21] for the dimensionality reduction of SIFT generated feature vector. Partially inspired by SIFT, another local feature descriptor which is illumination invariant, faster and more robust to transformations than the SIFT is presented in Ahonen et al. [22] and named as speeded-up robust features (SURF) in which keypoints are detected using Hessian blob detector. Later, SIFT and SURF are combined by Ali et al. [23, 24] for achieving better retrieval rate. Later, SIFT is combined with rotation invariant local binary pattern (LBP) [25], where LBP is used to define the local region centered at the keypoints which are identified by SIFT.

On the other side, inspired by the LBP [26], LBP is combined with histogram of gradient for better accuracy in INRIA dataset [27] and numerous variants of LBP has been introduced and some of them are integrated with other features for enhancing the retrieval result. Recently, SIFT and center symmetric LBP is combined by Heikkilä et al. [28] in which intensity of center symmetric pixels are only considered. Along this direction, center-symmetric local ternary patterns (CS-LTP) is introduced by Gupta et al. [29]; Local tetra pattern (LTrP) is introduced by Murula et al. [30] to define the structural formation of local level structure by including all the four directions for center pixel; directional binary wavelet pattern is reported in Murala et al. [30] for biomedical image indexing and retrieval; co-occurrence of similar ternary edges are encoded for CT and MRI image retrieval using the local ternary co-occurrence patterns

(LTCoP) and is suggested in Murala and Wu [31]; Local ternary pattern is described in Srivastava et al. [32]; Local mesh pattern (LMeP) and Local bit-plane decoded pattern for medical image retrieval is presented in Murala and Jonathan [33] and Dubey et al. [34] respectively; Dubey et al. [35, 36] described Local diagonal extrema pattern and Local wavelet pattern for CT image retrieval; Local quantized extrema pattern (LQEP) is introduced by Rao and Rao [37] for natural and texture image retrieval which captures the spatial relation between any pair of neighbors in a local region along the directions 0°, 45°, 90° and 135° for a given center pixel in an image; Directional local ternary quantized extrema pattern (DLTerQEP) is suggested in Deep et al. [38] for CT and MRI image retrieval and it captures more spatial structure information by adopting ternary patterns from horizontal, vertical, diagonal, anti-diagonal structure of directional local extrema values of an image.

Concurrently, Gradient location and orientation histogram with PCA is used in Mikolajczyk and Schmid [39] for retrieving the images; Liu and Yang [40] computes the spatial correlation of textons using texton co-occurrence matrices (TCM) which extracts energy, entropy, contrast and homogeneity to represent the image; attributes of co-occurrence matrix is expressed using histogram based on Julesz's textons theory for analyzing the natural images in Liu et al. [41] and is named as multi-texton histogram (MTH), and the authors confirmed that their approach achieves better performance than the texton co-occurrence matrix and edge orientation auto-correlogram; feature based on edge orientation similarity and underlying colors is described by Liu et al. [42] and named it as Micro-Structure Descriptor (MSD) which captures local level color and texture effectively; Saliency Structure Histogram (SSH) reported in Liu and Yang [43] computes the logarithm characteristics of Gabor energy to describe the image; Structure Element Descriptor (SED) comprising of color and texture information and Structure Element Histogram (SEH) comprising of the spatial correlation of color and texture feature is reported in Xingyuan and Zongyu [44] for image retrieval; Seetharaman and Sathiamoorthy [45] introduced a new variant of EOAC in which edges are identified in HSV color space using a framework based on Full Range Gaussian Markov Random Field (FRGMRF) model that extracts very minute and fine edges from HSV color space and evades loss of edges owing to spectral variations; Gradient field histogram of gradient (GF-HOG) is reported for the retrieval of photo collections [46] and it attains better results than the features like multi-resolution HOG, SIFT, structure tensor, etc. Histograms of triangular regions and relative spatial information for histogram-based representation of the BoVW (Bag of visual words) model are reported in Ali et al. [23, 24] and Zafar et al.

[47–49] respectively. Feature computation based on spatial information is reported in Zafar et al. [47–49], Latif et al. [50] and Ali et al. [51].

Subsequently, deep learning processes are employed in the domain of image recognition [52–56]. Though deep learning approaches are better in performance, their computational cost is too high and requires high configuration machines.

Recently, Zhao et al. [57] introduced a descriptor for CBIR system, called multi-trend structure descriptor (MTSD). The MTSD characterizes the image by exploiting the correlation among the local level structures of color, edge orientations and texture independently then they are integrated together as single feature vector. That is, MTSD is a feature matrix of color/edge orientation/ texture quantized values versus large, small and equal trends respectively where large trend belongs to pixel values from small to large, small trends corresponds to pixel values from large to small and equal trend means pixel values are same along 0°, 45°, 90° and 135° orientations from bottom to top and left to right direction in a 3×3 non-overlapping window. Zhao et al. [57] performed comprehensive experiments using MTSD on benchmark Corel image and Caltech datasets [57] and reported that MTSD is significantly outperforming TCM, MTH, MSD and SSH descriptors for image retrieval owing to effective discrimination of MTSD. We also validated the performance of MTSD for medical images in Natarajan and Sathiamoorthy [58, 59] and it achieves acceptable performance. However, more research should be done with MTSD to attain higher performance in image retrieval because though MTSD captures the local level structures of color, edge orientations and texture along 0°, 45°, 90° and 135° orientations, it lacks in encoding the orientation details of local level structures of color, edge orientations and texture and that might create descend in discriminative capability of MTSD. We strongly believe that combination of trends and its orientation rather than trends only will have a significant impact on the accuracy of the image retrieval system. Thus, the proposed work aims to encode the orientations of trends. In addition, average location of distribution of pixels for locally identified trends at each orientation is also encoded into the proposed variant of MTSD and thus the proposed variant of MTSD is able to achieve high retrieval accuracy owing to its high discrimination capability. However, the computation time of the proposed variant of MTSD is significantly far above the conventional MTSD and thus we utilized discrete Haar wavelet transformation to obtain the multiresolution pyramid image in which we performed decomposition up to the level of optimum. From the level of optimum, we computed the proposed variant of MTSD. The level of optimum is finalized empirically according to the experimental results of

[59]. The proposed variant of MTSD is represented as a histogram of orientations of local level structure and average location of distribution of pixels for locally identified trends at each orientation i.e., it encompasses a matrix of number of equal/small/large trends for quantized color/edge orientation/texture values versus orientations of trends and a matrix of average location of distribution of pixel values in small/large trends for quantized color/edge orientation/texture values versus orientations of trends. Thus, the proposed approach captures the rough spatial arrangement of the local level structure more effectively by encoding the trends and its orientations. Comprehensive experiments on benchmark dataset are carried out using the MTSD, LTCoP, LMeP, LQEP, DLTerQEP and the proposed variant of MTSD to show the superiority performance of the proposed one. Finally, we confirmed that including the orientation of trends and average location of distribution of pixel values in local level structures achieves highest performance for scene and biomedical image retrieval. The architecture of the proposed system is depicted in Fig. 1.

The rest of the paper is framed as follows: Sect. 3 presented the extraction of conventional MTSD. Issues with MTSD are explained in Sect. 4. Proposed variant of MTSD is described in Sect. 5. Section 6 recalls the discrete Haar wavelet transform. Experimental results and discussion is explained in Sect. 7. Section 8 discussed the conclusion and future scope of the proposed novel variant of MTSD method.

3 Multi-trend structure descriptor

In this section, we described the conventional feature descriptor called MTSD [57] which is compared with the performance of the proposed novel variant of MTSD descriptor.

The multi-trend structure descriptor is extracted from HSV color space as it is more natural to human visual perception [57]. Zhao et al. [57], color quantization is performed such that H, S and V components are segregated into 12, 3 and 3 bins and it results in $12 \times 3 \times 3 = 108$ colors and is defined as $0 \leq Q_c \leq 108$ where Q_c is the color quantized value; since V component consists of intensity details of an image, intensity details in V component are quantized into 20 [57] and is defined as $0 \leq Q_t \leq 20$ where Q_t is the intensity quantized value; due to the less computational cost and well performance of Sobel operator, it is used in Zhao et al. [57] for edge extraction and the orientations of extracted edges are computed [45] and are quantized into 9 orientations, and is defined as $0 \leq Q_e \leq 9$ where Q_e is the edge orientation quantized value.

Zhao et al. [57], the image is divided into 3×3 non overlapping blocks. For each quantized color/edge orientation/texture values, three trends namely equal, small and large trends along 0°, 45°, 90° and 135° orientations from bottom to top and left to right direction is considered in each 3×3 non overlapping blocks for extracting the correlation among the local level structure. According to Zhao et al. [57], the large trend belongs to pixel values from small to large, small trends corresponds to pixel values from large to small and equal trend means pixel values are same along 0° or 45° or 90° or 135° orientations from bottom to
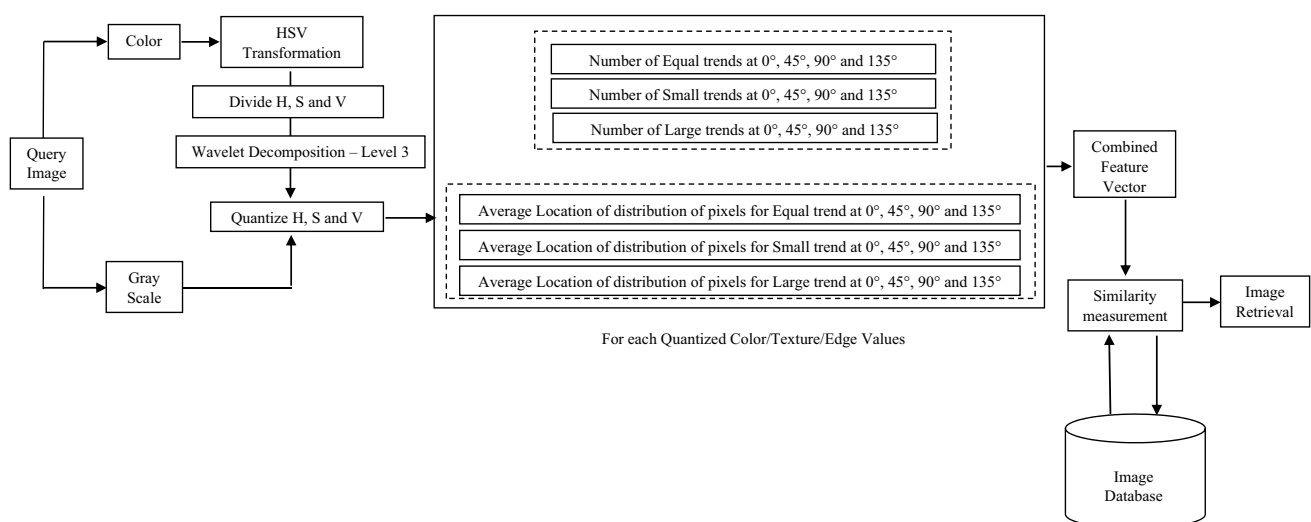

Fig. 1 Architecture of the proposed retrieval system

top or left to right direction in a 3×3 block. The identified trends are expressed as $\left(N_E^{Q_c}, N_S^{Q_c}, N_L^{Q_c} \right)$, $\left(N_E^{Q_t}, N_S^{Q_t}, N_L^{Q_t} \right)$, $\left(N_E^{Q_e}, N_S^{Q_e}, N_L^{Q_e} \right)$ in which N, $Q_c/Q_t/Q_e$, E, S and L represents the number of trends, quantized color/edge orientation/ texture values, equal, small and large trends respectively. That is, MTSD is a matrix of quantized color/edge orientation/texture values versus trends. For example, the color, edge orientation and texture feature matrix captured by the MTSD is shown in Fig. 2a–c respectively. In Fig. 2a, number of equal, small and large trends for quantized color value 0–107 is depicted, where 2, 5 and 0 are the number of equal, small and large trends corresponding to quantized color value 0 and vice versa. In Fig. 2b and 2c, number of equal, small and large trends for quantized texture value 0–19 and quantized edge orientations value 0–8 are represented respectively.

Thus, the authors [57] considered the orientation only for acquiring the local level structures and its correlation. Therefore, MTSD encodes only trends details and not the orientation details of each local level structure which results in less accuracy and is explained in Sect. 4.

The orientations (0°, 45°, 90° and 135°) considered for capturing the local level structure in the MTSD [57] is depicted in Fig. 3. For instance, small and large trends along 0° and 45° and equal trends along 90° and 135° is shown in Fig. 4. For example, process of computing the MTSD is depicted in Fig. 5. Figure 5a, b are the original images. Figure 5c consists of 2 equal, 2 small and 2 large trends for the quantized value 78 whereas Fig. 5d consists of 4 equal, 4 small and 0 large trends for the quantized value 80 respectively. Accordingly, Zhao et al. [57] computed equal, small and large trends for 108 color, 9 edge orientations and 20 texture quantized values respectively and it results in 411 dimensions and its dimensionality is reduced to 137 [57] in order to reduce the time and storage complexity.

4 Issues with multi-trend structure descriptor

For instance, consider the images (a) and (b) in Fig. 6. While computing the MTSD, it results that both images (a) and (b) are having same number of equal, small and large trends that is 2, 3 and 3 respectively. But, when we insight into the local structures acquired by MTSD, we noticed that for the first 3×3 block in Fig. 6a, image consists of one equal and one small trends along 90° and 0° respectively whereas the corresponding block in Fig. 6b also consists of one equal and one small trends but there orientation differs that is 0° and 90° respectively. Similarly, the second 3×3 block of both images contains same number of

	E	S	L
0	2	5	0
1	9	4	1
.	.	.	.
.	.	.	.
.	.	.	.
107	.	.	.

(a)

	E	S	L
0	1	8	11
1	6	14	1
.	.	.	.
.	.	.	.
.	.	.	.
19	.	.	.

(b)

|	E	S	L
0	1	15	10
1	19	3	8
.	.	.	.
.	.	.	.
.	.	.	.
8	.	.	.

(c)

Fig. 2 **a** Conventional MTSD for color feature. **b** Conventional MTSD for texture feature. **c** Conventional MTSD for edge feature

equal and large trends but their orientations also differ. Subsequently, in third 3×3 block of Fig. 6a, b images consists of one small trend along 90° orientation and for the fourth one, both Fig. 6a, b images consists of one small

(a)   **(b)**   **(c)**   **(d)**
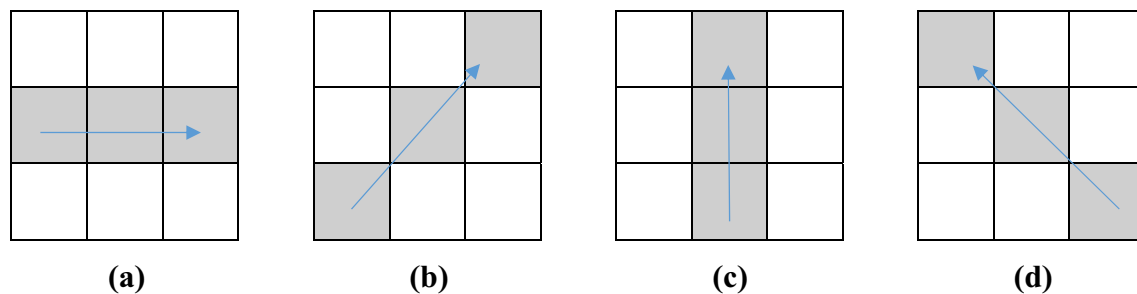
Fig. 3 Orientations and directions (left to right and bottom to top) considered in MTSD for capturing the local level structure, **a** 0°, **b** 45°, **c** 90°, **d** 135°

(a)   **(b)**   **(c)**   **(d)**
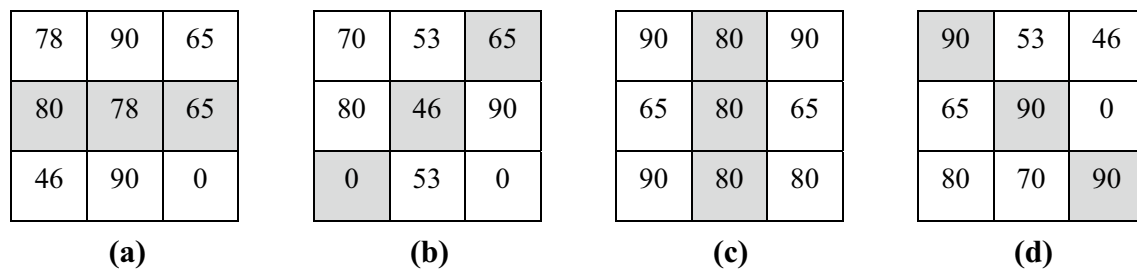
Fig. 4 Example for trends (**a**). Small trend (pixel values from large to small) along 0°; **b** large trend along 45° (pixel values from small to large); **c** equal trend (pixel values are same) along 90°; **d** equal trend (pixel values are same) along 135°

and one large trends but their orientations differ. From the aforesaid understandings, it is summarized that even if the number of equal, small and large trends are equal for images Fig. 6a, b, their orientation is same in some blocks and differ for the other blocks. So, concluding the similarity of images given in Fig. 6a, b as same based on MTSD is completely controversy.
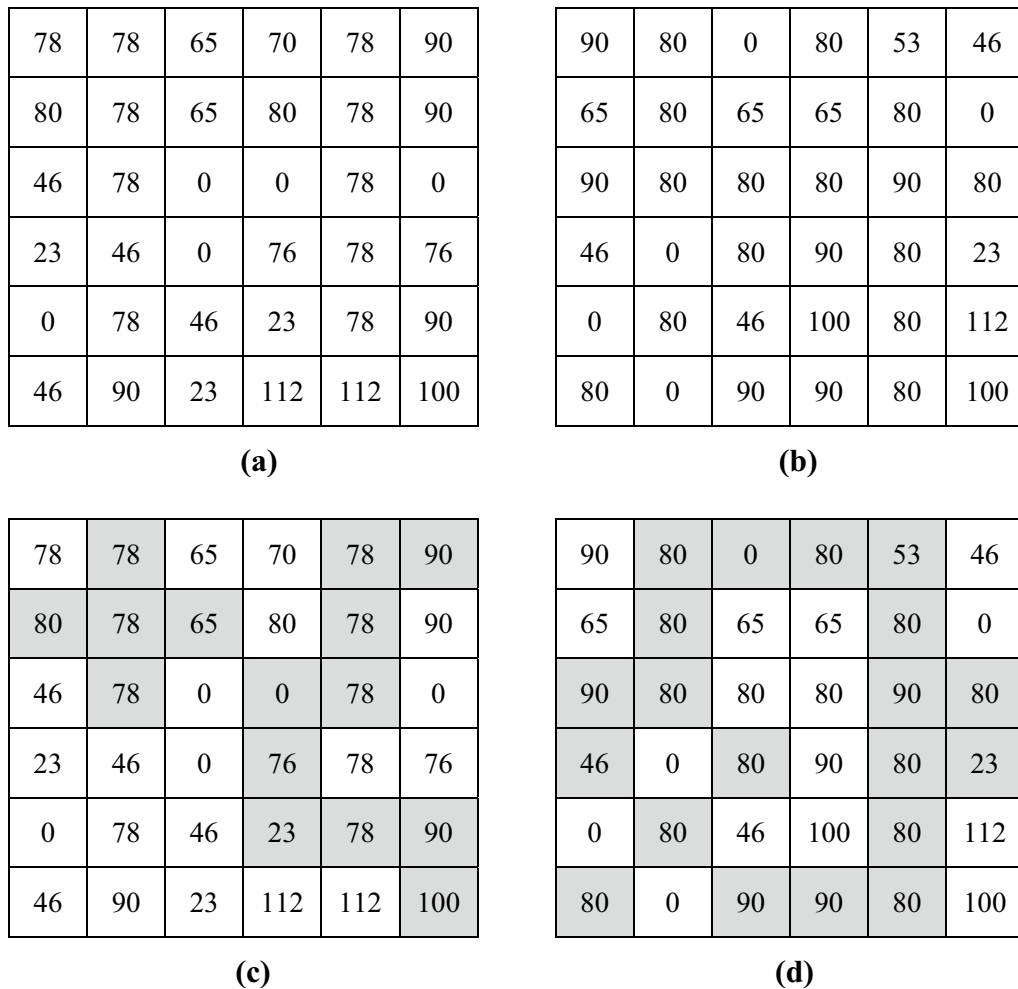
In addition to that even though the trend and its orientation is same in the third 3×3 block of both Fig. 6a, b images, the pixels values along the trends are (112, 78, 65) and (80, 78, 0) respectively which are drastically differ except the quantized pixel value at the center of the 3×3 block and such a drastic difference in the pixel values occurred for the most of the identified local level structures of an image. Therefore, considering both the images as similar is only based on number of small, large and equal trends is not correct. In order to overcome the aforesaid issues, in this paper, we proposed a novel variant of MTSD and it has high discriminative power then the conventional one.

5 Proposed new variant of MTSD

Though MTSD is superior in performance than the more familiar feature descriptors in the literature [57], in line with the aforesaid issues of MTSD, we strongly believe that

encoding the number of equal, small and large trends is not sufficient for effective image retrieval and it insists us to develop a new variant of MTSD with the understanding of above discussed literature. The proposed novel variant of MTSD encodes a feature matrix of color/edge orientation/texture quantized values versus orientations of equal, small and large trends instead a matrix of color/edge orientation/texture quantized values versus equal, small and large trends only, and a feature matrix of color/edge orientation/texture quantized values versus average location of distribution of pixels for small and large trends at each orientation. Since the average locations of distribution of pixels values are same for equal trends in each 3×3 block, it is not computed in the proposed approach. We named the proposed approach as Multi-direction and location distribution of pixels in trend structure (MDLDPTS).

For example, in the proposed approach, the color feature matrix for large trends captured by the proposed approach is represented in Fig. 7a. In Fig. 7A: a, number of large trends for quantized color value 0–107 is depicted and 8, 3, 1 and 5 are the number of large trends at orientations 0°, 45°, 90° and 135° respectively. In Fig. 7A: b, 78, 84, 34 and 92 are the location of distribution of pixels at orientations 0°, 45°, 90° and 135° for large trends respectively. Likewise, the texture and edge orientation feature matrices for large trends are computed and are depicted in Fig. 7b, c respectively. Similarly, feature matrix

78	78	65	70	78	90
80	78	65	80	78	90
46	78	0	0	78	0
23	46	0	76	78	76
0	78	46	23	78	90
46	90	23	112	112	100

(a)

90	80	0	80	53	46
65	80	65	65	80	0
90	80	80	80	90	80
46	0	80	90	80	23
0	80	46	100	80	112
80	0	90	90	80	100

(b)

78	78	65	70	78	90
80	78	65	80	78	90
46	78	0	0	78	0
23	46	0	76	78	76
0	78	46	23	78	90
46	90	23	112	112	100

(c)

90	80	0	80	53	46
65	80	65	65	80	0
90	80	80	80	90	80
46	0	80	90	80	23
0	80	46	100	80	112
80	0	90	90	80	100

(d)

Fig. 5 **a**, **b** The sample image pattern, **c**, **d**. The large, small and equal trends found along 0°, 45°, 90° and 135° for the original pattern (**a**, **b**)

for small and equal trends is also computed by the proposed approach. As above mentioned, the feature matrix of average location of distribution of pixel values of equal trends for quantized color/texture/edge orientation values versus orientations are not computed because pixel values are distributed equally in equal trends.
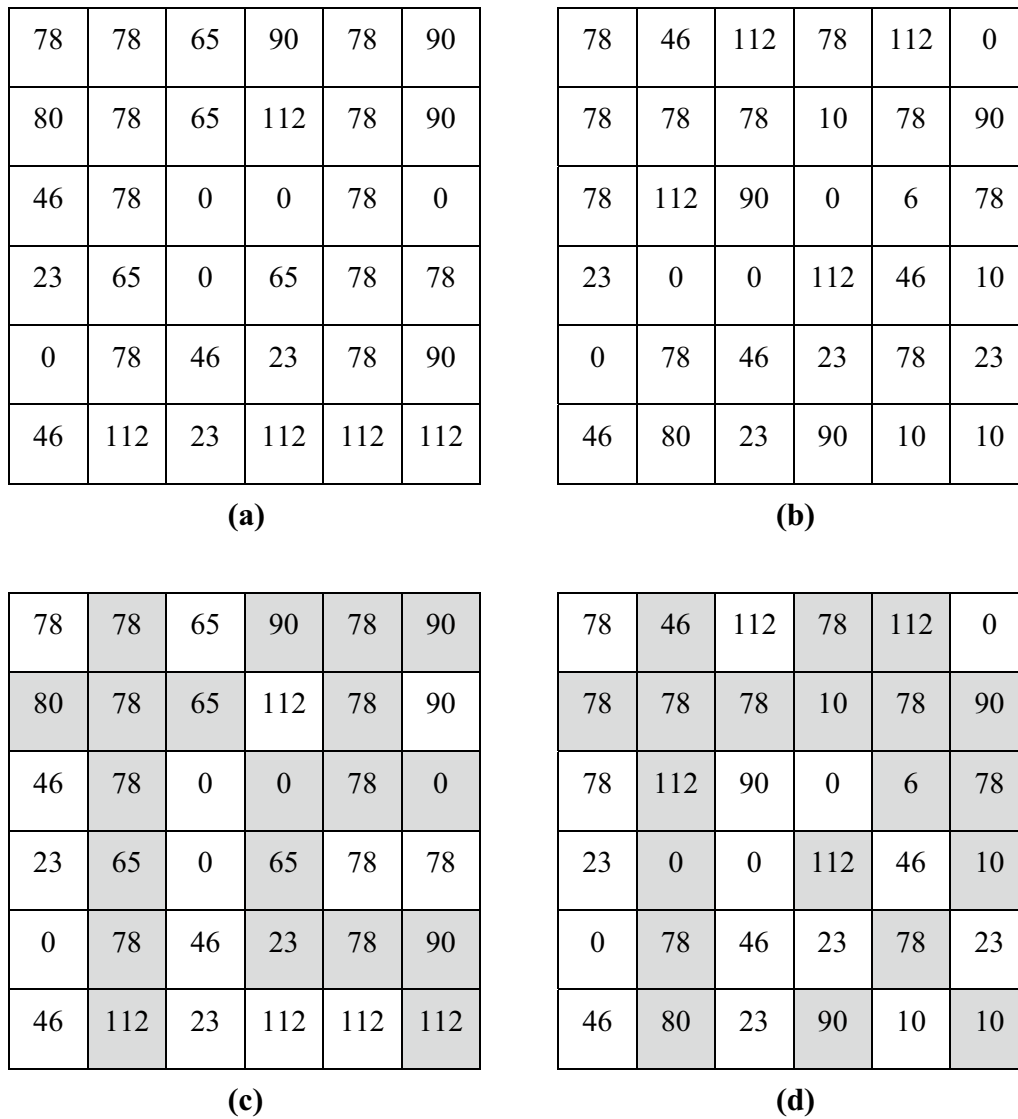
For instance, the computation of the proposed variant of MTSD for the first 3×3 block of image in Fig. 6a acquires one equal trend along 90° and one small trend along 0° and the location of distribution of pixel values in small trend is $74.33 \approx 74$. In Fig. 6b, the first 3×3 block has one equal trend along 0° and one small trend along 90° and the location of distribution of pixel values in small trend is $78.66 \approx 79$. Correspondingly, for other blocks also proposed approach computes the feature matrices as mentioned in Fig. 7. In the proposed system, the feature vector for trends and orientations is described as

$$F_\theta = \left\{ \left(\theta_E^{Q_c}, \theta_S^{Q_c}, \theta_L^{Q_c} \right), \left(\theta_E^{Q_e}, \theta_S^{Q_e}, \theta_L^{Q_e} \right), \left(\theta_E^{Q_t}, \theta_S^{Q_t}, \theta_L^{Q_t} \right) \right\} \quad (1)$$

where θ represents the orientation and $\theta \in \{0°, 45°, 90°, 135°\}$, Q_c represents quantized color value and $Q_c \in \{0, 1, \ldots, 107\}$, Q_e represents edge orientation value and $Q_e \in \{0, 1, \ldots, 8\}$, Q_t represents quantized texture value and $Q_t \in \{0, 1, \ldots, 19\}$, E denotes equal trend, S denotes small trend and L represents large trends. For instance, $\theta_E^{Q_c}$ describes the orientation of equal trend for quantized color value Q_c and the dimension is 108×4 whereas the dimension of $\theta_E^{Q_e}$ and $\theta_E^{Q_t}$ are 9×4 and 20×4 respectively and vice versa. The feature vector for location of the distribution of pixel is expressed as

$$F_\mu = \left\{ \left(\mu_{S_\theta}^{Q_c}, \mu_{L_\theta}^{Q_c} \right), \left(\mu_{S_\theta}^{Q_e}, \mu_{L_\theta}^{Q_e} \right), \left(\mu_{S_\theta}^{Q_t}, \mu_{L_\theta}^{Q_t} \right) \right\} \quad (2)$$

where μ represents the location of distribution of pixel values, θ represents the orientation and

78	78	65	90	78	90
80	78	65	112	78	90
46	78	0	0	78	0
23	65	0	65	78	78
0	78	46	23	78	90
46	112	23	112	112	112

(a)

78	46	112	78	112	0
78	78	78	10	78	90
78	112	90	0	6	78
23	0	0	112	46	10
0	78	46	23	78	23
46	80	23	90	10	10

(b)

78	78	65	90	78	90
80	78	65	112	78	90
46	78	0	0	78	0
23	65	0	65	78	78
0	78	46	23	78	90
46	112	23	112	112	112

(c)

78	46	112	78	112	0
78	78	78	10	78	90
78	112	90	0	6	78
23	0	0	112	46	10
0	78	46	23	78	23
46	80	23	90	10	10

(d)

Fig. 6 **a**, **b** The original pattern (**c**, **d**). The large, small and equal trends found along 0°, 45°, 90° and 135° for the original pattern (**a**, **b**)

$\theta \in \{0°, 45°, 90°, 135°\}$, Q_c represents quantized color value and $Q_c \in \{0, 1, \dots, 107\}$ in accordance with Zhao et al. [57], Q_e represents quantized edge orientation value and $Q_e \in \{0, 1, \dots, 8\}$ as in Zhao et al. [57], Q_t represents quantized texture value and $Q_t \in \{0, 1, \dots, 19\}$ as in Zhao et al. [57], S denotes small trend and L represents large trends. For instance, $\mu_{S_\theta}^{Q_c}$ describes the location of distribution of pixel values of small trend for quantized color value Q_c at orientation θ and the dimension is 108×4 whereas the dimension of $\mu_{S_\theta}^{Q_e}$ and $\mu_{S_\theta}^{Q_t}$ are 9×4 and 20×4 respectively and vice versa. The location of distribution of pixels in each local level structure is computed as expressed below

$$\mu = \frac{1}{M} \sum_{i=0}^{M} P_i \qquad (3)$$

where M is the number of pixels in a trend and P represent the pixel in a trend. Therefore the combined orientation and location of distribution of pixel values information of local level structure is expressed as

$$F = \{F_\theta, F_\mu\} \qquad (4)$$

Consequently, the dimension of proposed variant of MTSD is $(108 \times 4) \times 3 = 1296$ for color, $(9 \times 4) \times 3 = 108$ for edge and $(20 \times 4) \times 3 = 240$ for texture and in total, the dimension of proposed variant of MTSD is 1644. Similarly, the dimension of the location of distribution of pixel values of local level structure based on orientation

Fig. 7 A: a Feature matrix consisting of number of large trends for quantized color values versus orientations; **b** feature matrix consisting of average location of distribution of pixel values of large trends for quantized color values versus orientations. **B: a** feature matrix consisting of number of large trends for quantized texture values versus orientations; **b** feature matrix consisting of average location of distribution of pixel values of large trends for quantized texture values versus orientations. **C: a** Feature matrix consisting of number of large trends for quantized edge orientation values (quantized edge orientations are represented using the numbers from 0 to 8); **b** feature matrix consisting of average location of distribution of pixel values of large trends for quantized edge orientation values versus orientations

A

	0°	45°	90°	135°
0	8	3	1	5
1	16	2	5	3
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
107	.	.	.	.

(a)

	0°	45°	90°	135°
0	78	84	34	92
1	123	42	97	90
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
107	.	.	.	.

(b)

B

	0°	45°	90°	135°
0	13	11	10	8
1	19	22	55	23
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
19	.	.	.	.

(a)

	0°	45°	90°	135°
0	98	104	134	90
1	103	92	87	70
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
19	.	.	.	.

(b)

C

	0°	45°	90°	135°
0	8	10	25	17
1	14	22	33	101
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
8	.	.	.	.

(a)

	0°	45°	90°	135°
0	109	80	84	92
1	109	66	77	101
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
8	.	.	.	.

(b)

is $(108 \times 4) \times 2 = 864$ for color, $(9 \times 4) \times 2 = 72$ for edge and $(20 \times 4) \times 2 = 160$ for texture (equal trends are excluded in the feature matrix of location of distribution of pixel values of local level structure vs. orientations) and the total dimension is 1096. Thus the size of the combined feature vector is 2740 and is reduced to 685 as agreeing with the dimensionality reduction approach of Zhao et al. [57]. Thus, the proposed approach encodes orientations of local structures as well the location of the distribution of pixel values for each local level structure which are enriching the capability of the proposed approach and it succeeds to achieve better results on benchmark datasets than the

existing MTSD. Both the MTSD and proposed variant of MTSD are represented using histograms. In the proposed work, quantization levels for color, edge and texture are agreed with Zhao et al. [57].

6 Discrete Haar wavelet transform

Even though proposed approach is supremacy to conventional MTSD in terms of retrieval rate, its computational time is higher and thus we intended to reduce it by incorporating the discrete Haar wavelet transform [60]. It is obviously confirmed from the literature that researchers working in CBIR, employed wavelets due to its multi-resolution ability and obtained well retrieval performance [60] and thus we employed it, and it performs decomposition by estimating the approximations and details till the optimum level and results in wavelet pyramid image formation [60]. The multi-resolution pyramid image is attained using discrete Haar wavelet transform as in Seetharaman and Kamarasan [60] and is described as follows

$$
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} \Leftarrow \begin{bmatrix} a_0 \\ w_0 \\ a_1 \\ w_1 \\ a_2 \\ w_2 \\ a_3 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} l_0 \\ l_1 \\ l_2 \\ l_3 \\ l_4 \\ l_5 \\ l_6 \\ l_7 \end{bmatrix}
$$

(5)

$l_0, l_1, \ldots, l_k$ are the input images and contains k coefficients in which there are k/2 approximations and k/2 wavelet coefficients and are stored in the upper $[a_0, a_1, .a_3]$ and lower $[w_0, w_1, .w_3]$ arrays respectively and this procedure is repeated till optimum level is achieved. The level in which the numbers of details are very less without dropping the dominant details in an image is called fine or optimum level. In the proposed work, optimum level is determined based on trial and error approach. Multi-resolution pyramid image is framed by a sub-sampling rate of 2 with no overlap approach as in Seetharaman and Kamarasan [60]. Figure 8 depicts the pyramid structure of an example image. The size of an example image is resized to 256×256 in the 0th level and is decomposed up to level 5. At each level of decomposition, the size of an example image is diminishing to half of an example image at previous level. Thus, we obtained 128×128, 64×64, 32×32, 16×16 and 8×8 dimensions at levels 1st to 5 respectively and the level of optimum is chosen as 3 according to the results of our previous work [58, 59]. Therefore, the proposed variant of MTSD is computed from the optimum level i.e., 3 which preserves the accuracy of proposed approach and reduces the computational cost.

7 Experimental results and discussion

The experiments are conducted using the core i3 processor, 4 GB RAM and 64 bit Windows operating system. In our experiments, we test the retrieval performance and time complexity. The algorithm for the computation of proposed approach is described as below.

Algorithm: MDLDPTS (Q, DB$_N$)

Input: Query image (Q) and images in the image database (DB$_N$)

1.   If Image is in gray-scale discrete Haar wavelet is applied upto level 3 then Goto Step 6
2.   Image in RGB color space is converted into HSV color space
3.   Divide image into H, S and V color channels
4.   Discrete Haar wavelet is applied distinctively for each channel and optimum level is set to 3

Color Information:

5.   Color information is uniformly quantized into 12, 3 and 3 bins.
6.   Divide the image into number of non-overlapping blocks of size 3 x 3
7.   Compute (N_c)   // N_c- number of Quantized color

Edge Information:

8.   The Sobel operator is applied to extract the edge from the intensity image and the orientation of edges are computed then the edge orientations are uniformly quantized into 9 bins.
9.   Divide the edge extracted image into number of non-overlapping blocks of size 3 x 3
10. Compute (N_e)   // N_e- number of Quantized edge orientations

Texture Information:

11. Intensity information in V channel is uniformly quantized into 20 bins.
12. Divide the intensity image into number of non-overlapping blocks of size 3 x 3
13. Compute (N_t)   // N_t- number of Quantized textures

Combined Feature Vector:

14. Computed feature vectors are combined then data dimensionality reduction is performed

// Assume that the feature vector is computed for all the images in the datasets as in step 1 to 14.

Similarity Matching:

15. int D[DB$_N$]
16. For i=1 to DB$_N$

    D[i]= Similarity between the feature vector of Q and DB_i using Euclidean distance
    Next i

17. Bubble sort is performed on array D
18. Return k number of similar images in ascending order of similarity values

Compute (int N)

{

        For each c in N

            For i=0 to Number of rows in an image (M) do step3

            For j=0 to Number of Columns in an image (N) do step3

                If the c is equal to the pixel value at the center of the block then

                    a. Computes the number of equal, small and large trends for each orientation

                    b. Computes the location of distribution of pixel values for small and large trends at each orientation.

                End if

            Next

        Next

    Next

}
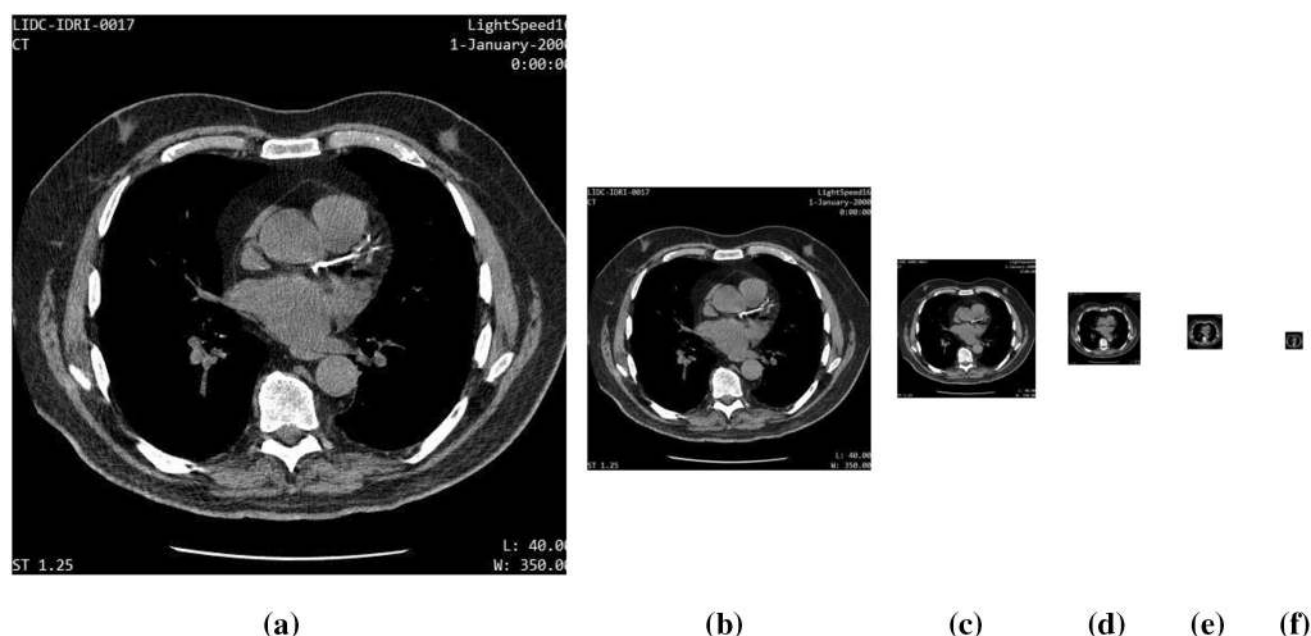
Fig. 8 Multi-resolution pyramid image from level-0 to level-5 (**a**–**f**) utilizing discrete Haar wavelet transform. The input image is of size 256×256 and images in the next level are of size 128×128, 64×64, 32×32, 16×16 and 8×8

7.1 Image datasets

To scrutinize the performance of proposed approach against state-of-the-art feature descriptors, experiments are conducted over the natural and textural image benchmark datasets [42, 57] namely Corel-1k, Corel-5k, Corel-10k, Caltech-101. The Corel-1k consists of 10 categories of 1000 images and each category consists of 100 images. The Corel-5k contains 50 categories of 5000 images and each category consists of 100 images. Corel-10k has 100 categories of 10,000 images and each category consists of 100 images. The Corel-1k and 5k are the part of Corel-10k dataset. Caltech-101 has 101 categories of 9146 of size 300×200 and about 40–800 images per category respectively. These benchmark datasets contains variety of image categories like elephants, mountains, beach, buildings, food, car, door, rhino, men, women, lion, antiques, cat, deer, etc. Few sample images of Corel-1k, Corel-5k, Corel-10k and Caltech-101 datasets are shown in Fig. 9.

Further, in order to estimate the performance of the proposed system for CT and MR images retrieval, experiments are performed on benchmark dataset namely LIDC-IDRI-CT (ftp://medical.nema.org/medical/Dicom/Multiframe/) and VIA/I-ELCAP-CT (http://www.via.cornell.edu/-databases/lungdb.html) and OASIS-MRI [61] which are open accessible. The LIDC-IDRI stands for from Lung Image Database Consortium and Image Database Resource Initiative which consists of 84 cases of lung CT images of size 512×512 in digital imaging and communication format (DICOM)

with the annotations of physicians in XML file and each case contains 100–400 images. The VIA/I-ELCAP-CT stands for vision and image analysis/International early lung cancer action program, which is also a collection of lung CT images of size 512×512 in DICOM format. The open access series of imaging studies (OASIS) is collection of MR images of 421 cases with 4 classes of each 124, 102, 89 and 106 respectively based on the shape of the ventricular. Few sample images of LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets are shown in Fig. 10.

7.2 Performance assessment

In the experiments, from each category of each dataset, images are selected randomly as query image and to assess the performance of the proposed approach, the top 100 retrieval results are considered. In order to assess the performance of proposed approach, we used precision and recall measures and are expressed as in Seetharaman and Sathiamoorthy [8, 45]:

$$\text{Precision (P)} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (6)$$

$$\text{Recall (R)} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in the database}}. \quad (7)$$

Fig. 9 Sample images from Corel-1k, Corel-5k, Corel-10k and Caltech-101 dataset

Fig. 10 Sample images from LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets
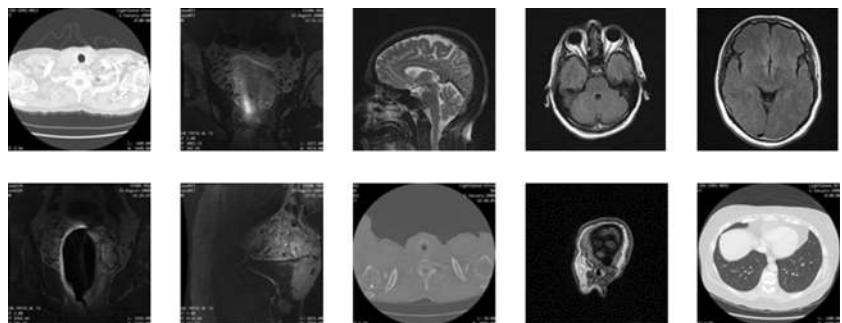
7.3 Similarity measurement

The goal of image retrieval system is retrieving the finest k number of images that are similar to query image. Selecting the most similar k images can be done using the similarity measures which measure the distance between feature vectors of query and target images in the database. The result of similarity measure lies between 0 and 1 where 0 means both query and target images are exactly same and 1 means query and target images are completely different. In this work, the similarity between the query and target images are estimated through more frequently used Euclidean distance [62, 63] which is expressed as

$$S(Q, T) = \sqrt{\sum_{i=0}^{N} \left(|Q_i - T_i| \right)^2} \qquad (8)$$

where Q and T denote the query and target feature vector and N represents the number of features. In the proposed system, the estimated similarity values are ordered in ascending manner using bubble sort method to acquire the similar images at the top level for image retrieval.

7.4 Experiments

According to the results obtained in our previous work [58, 59] the level of decomposition is set to 3 which significantly provides better retrieval accuracy than the image at level 4 and 5. The computational cost at level 3 is less and the retrieval results for decomposed image at level 3 and level 0 to level 2 are more or less similar. But, though the computational cost at level 4 and 5 is too less, its retrieval accuracy drastically differs from level 0–3. As a result, in all our forgoing experiments, we use optimum level as 3 for estimating the proposed approach. For instance, let we assume the size of an image at decomposed level 3 is 32×32 and the proposed approach estimates local level structures from 339 (i.e., $113 + 113 + 113$) non-overlapping windows of size 3×3 whereas estimating the proposed variant of MTSD from the level 0 image of size 256×256 uses 21,843 non-overlapping windows of size 3×3.

Table 1 The average precision of the proposed MDLDPTS and state-of-the-art techniques on various natural and textural datasets

Datasets	Precision (%) (n = 10)						
	MDLDPTS1 (with color, edge and texture information)	MDLDPTS2 (with edge and texture information only)	MTSD	LTCoP	LMeP	LQEP	DLTerQEP
Corel-1k	87.78	78.21	83.23	81.21	83.24	85.45	87.23
Corel-5k	74.23	66.83	69.89	69.10	70.11	71.18	73.21
Corel-10k	72.80	63.34	68.45	67.39	68.93	69.08	70.51
Caltech-101	69.42	57.11	61.99	61.06	63.27	65.93	67.87

Table 2 The average precision of the proposed MDLDPTS and state-of-the-art techniques on various medical datasets

Datasets	Precision (%) (n = 10)						
	MDLDPTS1 (with color, edge and texture information)	MDLDPTS2 (with edge and texture information only)	MTSD	LTCoP	LMeP	LQEP	DLTerQEP
LIDC-IDRI-CT	68.94	67.32	59.32	58.77	61.05	63.43	66.98
VIA/I-ELCAP-CT	67.40	66.27	60.56	59.21	61.22	62.77	64.20
OASIS-MRI	64.83	63.09	56.99	55.89	57.11	60.69	62.02

Table 3 The dimension of the proposed MDLDPTS and state-of-the-art techniques

Methods	Dimension
LTCoP	512
LMeP	768
LQEP	4096
DLTerQEP	2×4096
MTSD	137
MDLDPTS1 (considering color, texture and shape details)	685
MDLDPTS2 (only considering the texture and edge information for CT and MR images)	145

In this section, seven experiments are performed to compare the discrimination capability of the proposed variant of MTSD and conventional MTSD [57], LTCoP [31], LMeP [33], LQEP [37] and DLTerQEP [38] for image retrieval. Since the MTSD is outperforming for scene [57] and medical [58, 59] datasets and proposed approach is a variant of MTSD, in the experiments, we considered Corel-1k, Corel-5k, Corel-10k, Caltech-101, LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets. The retrieval performance of the proposed approach and state-of-the-art feature descriptors are measured in terms of precision and average recall as described in Eqs. (6, 7).

The average precision and average recall [38] for the proposed approach and state-of-the-art feature descriptors for Corel-1k, Corel-5k, Corel-10k, Caltech-101, LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets are depicted in Tables 1 and 2 respectively. The Table 3 illustrated the dimension of the proposed approach and

state-of-the-art feature descriptors. Table 4 depicts the precision for proposed MDLDPTS1 (with color, edge and texture information) and MDLDPTS2 (with edge and texture information) at various levels of decomposition by Haar wavelet transform and Table 5 shown the similarity values between sample input and the target image for the wavelet based proposed approach at various level of decomposition. Since the medical images are in gray scale, color information plays very less role in image characterization and thus MDLDPTS1 and MDLDPTS2 shown very less difference in the retrieval rate for medical images whereas the difference in retrieval rate between MDLDPTS1 and MDLDPTS2 is significantly high for scene image datasets.

In experiment 1, Corel-1k dataset is used. Query image and the top 5 retrieval results of proposed approach is illustrated in Fig. 11a. The performance of the proposed approach and state-of-the-art feature descriptors in terms of average precision and average recall is depicted in Fig. 12. From Fig. 12, it is clearly evident that the proposed approach highly supersedes the state-of-the-art feature descriptors. The performance of DLTerQEP and MTSD is moderate and MDLDPTS2 provides worst performance.

The subsequent experiment is performed for Corel-5k dataset. The plot in Fig. 13 shown the performance of proposed and state-of-the-art approaches. It is obviously demonstrated that the proposed approach significantly achieves better accuracy than the other descriptors. The MDLDPTS2 gives worst performance because it encompasses texture and edge information only and DLTerQEP and MTSD provides moderate performance. Query image

Table 4 Precision (%) for the proposed MDLDPTS1 and MDLDPTS2 at various levels of decomposition by Haar wavelet transform

Datasets	Level of decomposition and methods											
	L_0		L_1		L_2		L_3		L_4		L_5	
	MDLDPTS1	MDLDPTS2	MDLDPTS1	MDLDPTS2	MDLDPTS1	MDLDPTS2	MDLDPTS1	MDLDPTS2	MDLDPTS1	MDLDPTS2	MDLDPTS1	MDLDPTS2
Corel-1k	87.51	78.01	87.63	78.09	87.68	78.16	87.78	78.21	89.14	80.11	92.21	82.16
Corel-5k	73.99	66.66	74.09	66.74	74.17	66.80	74.23	66.83	76.23	68.17	78.12	70.12
Corel-10k	72.64	63.12	72.71	63.23	72.76	63.29	72.80	63.34	74.33	65.11	76.78	68.01
Caltech-101	69.23	56.89	69.31	56.95	69.37	57.05	69.42	57.11	71.90	59.01	74.09	61.24
LIDC-IDRI-CT	68.79	67.11	68.83	67.20	68.88	67.28	68.94	67.32	70.90	69.00	72.67	72.08
VIA/I-ELCAP-CT	67.25	66.02	67.32	66.16	67.37	66.20	67.40	66.27	69.89	68.78	71.89	71.01
OASIS-MRI	64.70	62.89	64.75	62.96	64.80	63.01	64.83	63.09	66.91	65.78	69.16	68.01

and the top 5 retrieval results from the Corel-5k dataset for the proposed approach is illustrated in Fig. 11b.

In the next experiment Corel-10k dataset is employed. Average precision versus average recall graph for the proposed and existing approaches is shown in Fig. 14 and the results clearly shown that the least performer is MDLDPTS2 and proposed approach is superior performer, and MTSD and DLTerQEP are moderate. Query image and the top 5 retrieval results of proposed approach for Corel-10k dataset is demonstrated in Fig. 11c.

Later, Caltech-101 dataset is considered. The plot for average precision versus recall is used to measure the accuracy of the proposed and existing approaches and is depicted in Fig. 15 and for instance, the top 5 retrieval results are illustrated Fig. 11d. The results shown in Fig. 15 describes that the proposed variant of MTSD outperforms the state-of-the-art feature descriptors. Since the proposed approach encodes the orientations and local distribution of pixel values of each identified local level structures of the color, edge orientation and texture information at both local and global level which in turns includes the spatial arrangements of local level structure, it outperforms the conventional feature descriptors for scene image datasets. The average precision performance of the proposed variant of MTSD and existing feature descriptors for randomly selected 10 classes of images from Corel and Caltech-101 dataset is illustrated in Figs. 16 and 17 and it obviously demonstrates that the proposed variant of MTSD obtained the best precision rates for randomly selected categories in both datasets.

Subsequently, we interested to test the efficiency of the proposed approach on medical images. Hence, we performed experiments on medical image datasets namely LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI. The retrieval performance of the proposed and existing feature vectors in terms of average precision versus average recall is illustrated in Figs. 18, 19 and 20 for LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets respectively. The top 5 retrieval results by the proposed variant of MTSD are shown in Fig. 21a–c for LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS-MRI datasets respectively. From the outcomes it is obvious that the proposed variant of MTSD appreciably outperforms the state-of-the-art feature descriptors for medical image retrieval.

Though the computation burden in image matching phase and storage cost of proposed MDLDPTS1 is significantly higher than the MTSD, it is neglected due to the higher accuracy attained for scene image retrieval. Whereas the computation burdens and storage cost of the proposed MDLDPTS1 is appreciably lesser than the more familiar LTCoP, LMeP, LQEP and DLTerQEP feature descriptors as well the retrieval accuracy of proposed MDLDPTS1 is also significantly higher than the others and it is owing

Table 5 Similarity values between sample input and the target image based on Wavelet based MDLDPTS1 at various level of decomposition

Concept and image no.	Level of decomposition and methods					
	L_0	L_1	L_2	L_3	L_4	L_5
	MDLDPTS	MDLDPTS	MDLDPTS	MDLDPTS	MDLDPTS	MDLDPTS
Dino	0.1240	0.1247	0.1253	0.1259	0.3432	0.4567
Image_644021 versus Image_644030						
Deer	0.7311	0.7319	0.7322	0.7328	0.8091	0.9002
Image_77012 versus Image_77033						
Lotus	0.2512	0.2519	0.2525	0.2530	0.3341	0.3923
Image_0012 versus Image_ 0056						
Lung in LIDC	0.1001	0.1010	0.1017	0.1026	0.1014	0.1999
LIDCo000023 versus LIDCo000032						
Car	0.1902	0.1910	0.1919	0.1928	0.2608	0.3004
Image_29044 versus Image_29033						
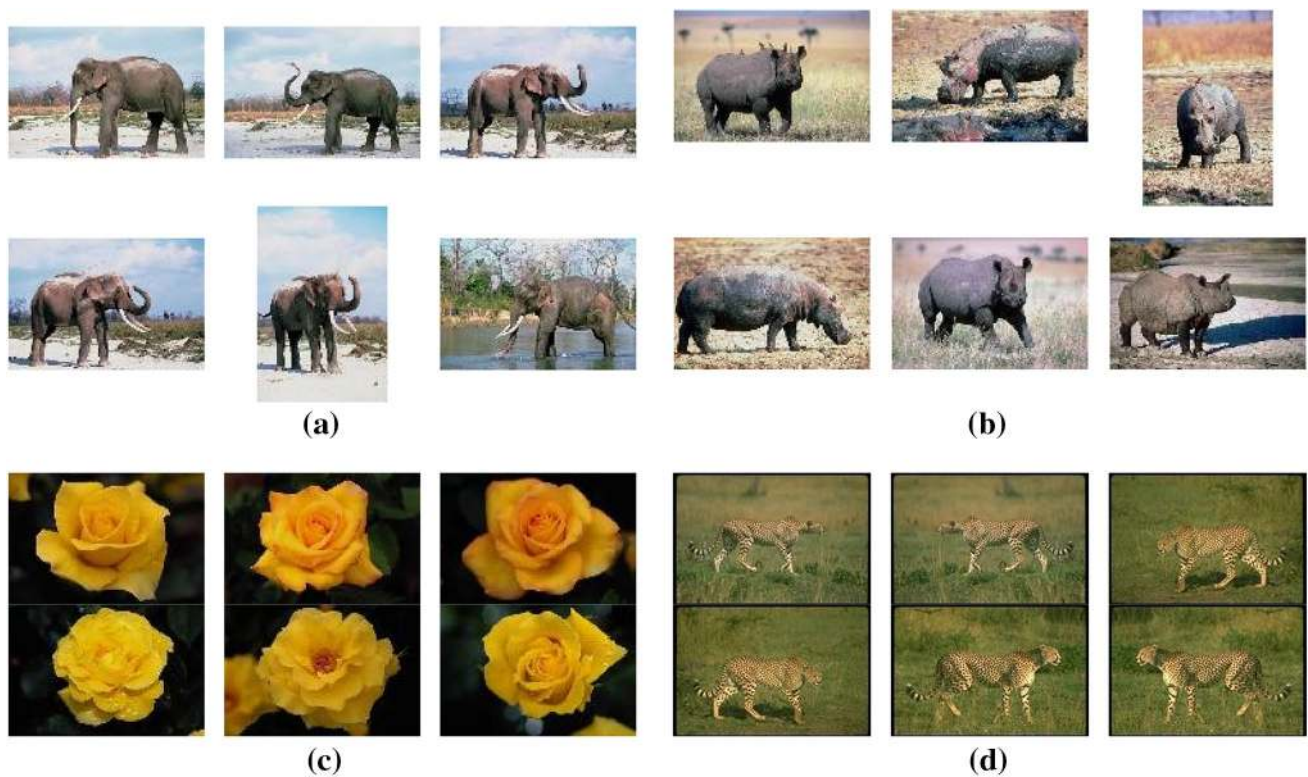
Fig. 11 a–c An example query image and its top 5 retrieval results for Corel dataset using the proposed approach, **d** an example query image and its top 5 retrieval results for Caltech-101 dataset using the proposed approach

to the ability of identifying the very minute and minor deformation in an image by considering the orientation and average location distribution of pixels in the local structure. Subsequently, the computation burden and storage cost of proposed MDLDPTS2 is significantly lower than the MTSD, proposed MDLDPTS1 and other state-of-the-art descriptors. Results of MDLDPTS2 evident that

neglecting the local level structure information of color leads to loss of accuracy for scene database.

As the medical images considered in the experiments are gray scale, both MDLDPTS1 and MDLDPTS2 are performing more or less equally. However, MDLDPTS1 shown quite high in precision versus recall plot because it significantly captures the local level structure from gray color
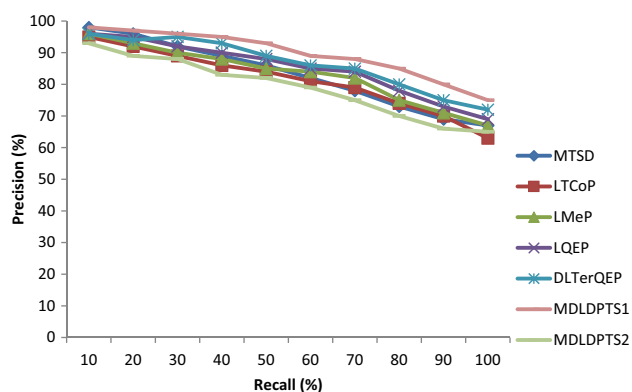
Fig. 12 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on Corel-1k dataset
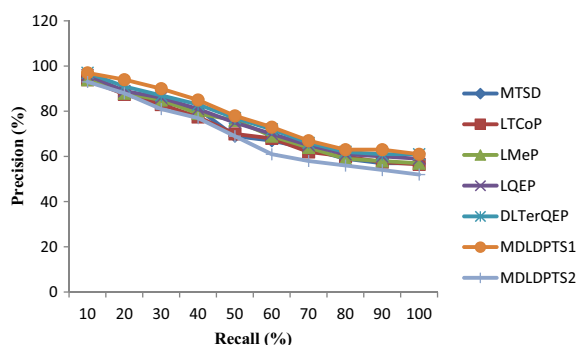
Fig. 13 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on Corel-5k dataset
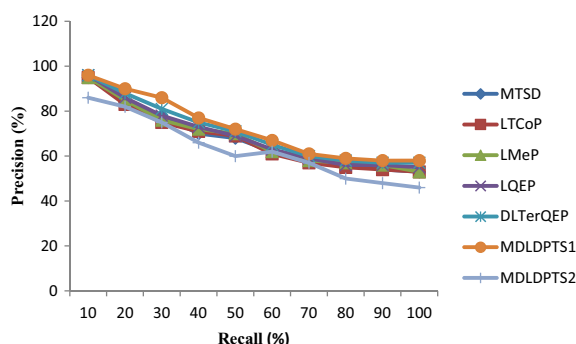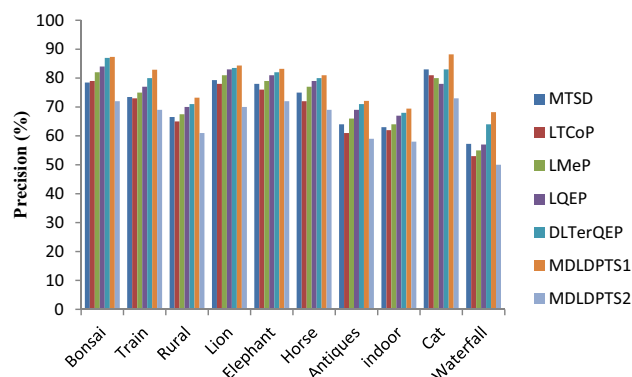

Fig. 14 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on Corel-10k dataset
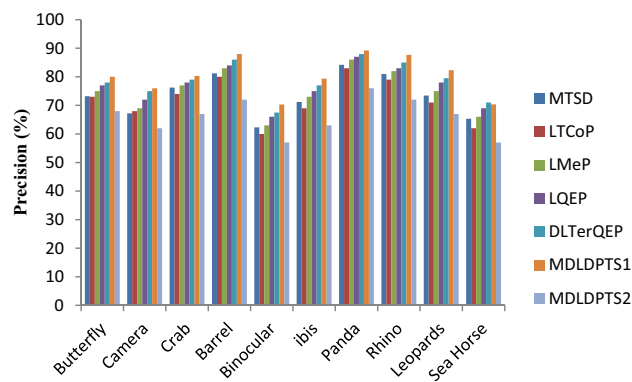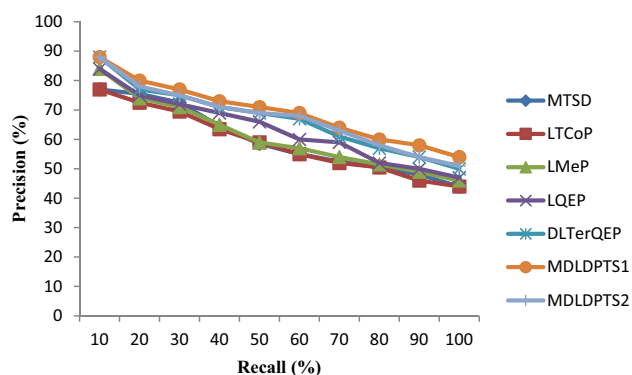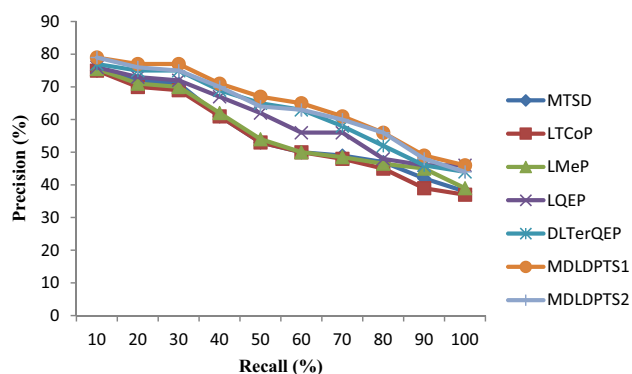

Fig. 15 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on Caltech-101 dataset

Fig. 16 Performance comparison of the proposed variant of MTSD and state-of-the-art techniques for randomly selected 10 different classes in Corel dataset

Fig. 17 Performance comparison of the proposed variant of MTSD and state-of-the-art techniques for randomly selected 10 different classes in Caltech-101 dataset

details of gray scale images. But for such a quite high difference provided by MDLDPTS1 for medical images, high computation cost is required. Thus, from the results we conclude that for scene and medical images MDLDPTS1

and MDLDPTS2 respectively trades off between retrieval rate and time cost.

Fig. 18 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on LIDC-IDRI-CT dataset
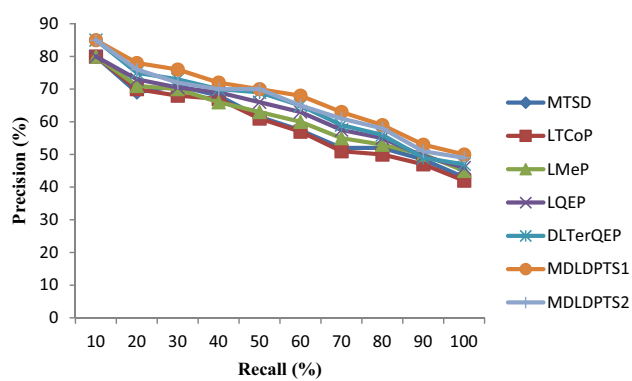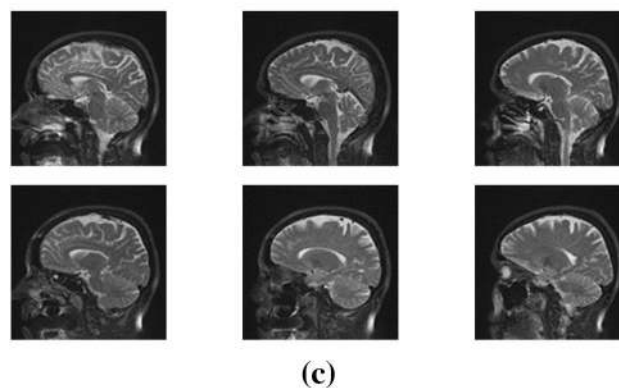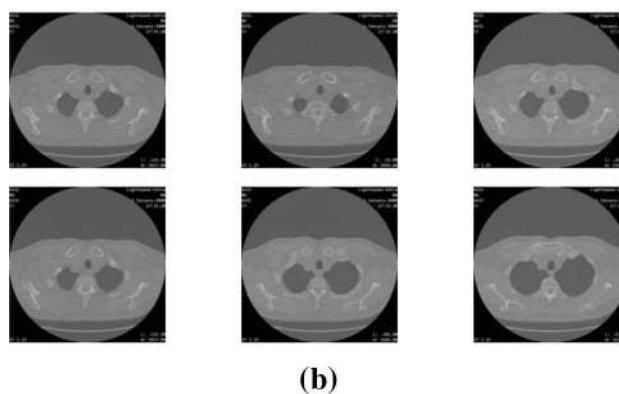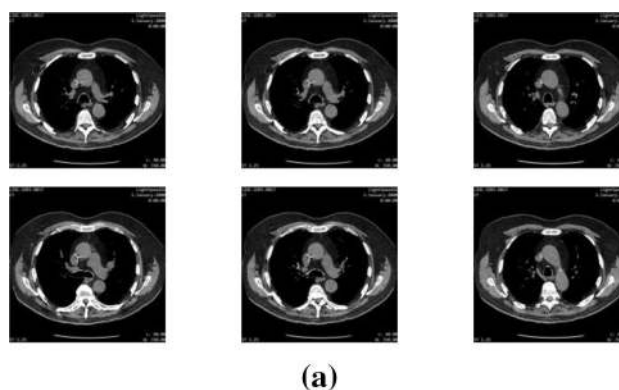

Fig. 19 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on VIA/I-ELCAP-CT dataset

Fig. 20 Precision versus recall graph for the proposed variant of MTSD and state-of-the-art techniques on OASIS-MRI dataset

Fig. 21 **a**–**c** An example query image and its top 5 retrieval results for LIDC-IDRI-CT, VIA/I-ELCAP-CT and OASIS dataset using the proposed approach

8 Conclusion

A novel variant of MTSD for scene and medical image retrieval is proposed which encodes the orientation details of local level structures. In addition, it also incorporates average location of distribution of pixel values of each identified local level structures. To reduce the computational cost of the proposed variant of MTSD, it is computed from the pyramid structure multiresolution domain. The performance of the proposed approach is evaluated using the benchmark datasets in terms of precision and recall. The result after investigation illustrates a considerable improvement as compared to the state-of-the-art descriptors for natural, textural and biomedical image retrieval

and it is because of effective characterization of local level structures and its spatial arrangements. In future, the proposed approach can be implemented for color medical images and combined with efficient machine learning methods to design an effective content based image retrieval system.

Compliance with ethical standards

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

Human and animal rights This article does not contain any studies with human participants or animals.

References

1. Stricker M, Orengo M (1995) Similarity of color images. In: Proceeding of the SPIE storage and retrieval for image and video databases, San Jose, pp 381–392
2. Alsmadi MK (2017) An efficient similarity measure for content based image retrieval using memetic algorithm. Egypt J Basic Appl Sci 4:112–122
3. Jayashree K-C, William H (2008) Effectiveness of global features for automatic medical image classification and retrieval—the experiences of OHSU at ImageCLEFmed. Pattern Recogn Lett 29:2032–2038
4. William H, Antani S, Long LR, Neve L, Thoma GR (2009) SPIRS: a web-based image retrieval system for large biomedical databases. Int J Med Inf 78:S13–S24
5. Burdescu DD, Mihai CG, Stanescu L, Brezovan M (2013) Automatic image annotation and semantic based image retrieval for medical domain. Neurocomputing 109:33–48
6. Zheng Q-F (2008) Constructing visual phrases for effective and efficient object-based image retrieval. ACM Trans Multimed Comput Commun Appl (TOMCCAP) 5(1):7
7. Giveki D, Soltanshahi MA, Montazer GA (2017) A new image feature descriptor for content based image retrieval using scale invariant feature transform and local derivative pattern. Optik 131:242–254
8. Seetharaman K, Sathiamoorthy S (2014) Color image retrieval using statistical model and radial basis function neural network. Egypt Inf J 15(1):59–68
9. Datta R, Joshi D, Li J, Wang JZ (2008) Image retrieval: ideas, influences, and trends of the new age. ACM Comput Surv 40(2):1–60
10. Michael SL, Nice S, Chababe D, Ramsesh J (2006) Content-based multimedia information retrieval: state of the art and challenges. ACM Trans Multimed Comput Commun Appl 2(1):1–19
11. Farhidzadeh H, Goldgof DB, Hall LO, Gatenby RA, Gillies RJ, Raghavan M (2015) Texture feature analysis to predict metastatic and necrotic soft tissue sarcomas. In: IEEE international conference on systems, man, and cybernetics (SMC), pp 2798–2802
12. Farhidzadeh H, Kim JY, Scott JG, Goldgof DB, Hall LO, Harrison LB (2016) Classification of progression free survival with nasopharyngeal carcinoma tumors. In: SPIE medical imaging, international society for optics and photonics, p 97851
13. Charles YR, Ravi R (2016) A novel local mesh color texture pattern for image retrieval system. AEU Int J Electron Commun 70(3):225–233
14. Feng L, Wu J, Liu S, Zhang H (2015) Global correlation descriptor: a novel image representation for image retrieval. J Vis Commun Image Represent 33:104–114
15. Zeng Z (2016) A novel local level structure descriptor for color image retrieval. Information 7(1):9
16. Williams A, Yoon P (2007) Content-based image retrieval using joint correlograms. Multimed Tools Appl 34(2):239–248
17. Chatzichristofis SA, Zagoris K, Boutalis YS, Papamarkos N (2010) Accurate image retrieval based on compact composite descriptors and relevance feedback information. Int J Pattern Recognit Artif Intell 24(2):207–244
18. Savvas A, Chatzichristofis AA (2010) Late fusion of compact composite descriptors for retrieval from heterogeneous image databases. In: SIGIR'10, July 19–23, Geneva
19. Mahmoudi F, Shanbehzadeh J, Eftekhari AM, Soltanian-Zadeh H (2003) Image retrieval based on shape similarity by edge orientation autocorrelogram. Pattern Recogn 36:1725–1736
20. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110
21. Ke Y, Sukthankar R (2004) PCA-SIFT: a more distinctive representation for local image descriptors. In: IEEE conference on computer vision and pattern recognition, pp 506–513
22. Ahonen T, Matas J, He C, Pietikäinen M (2009) Rotation invariant image description with local binary pattern histogram Fourier features. In: Salberg A-B, Hardeberg JY, Jenssen R (eds) Image analysis, SCIA 2009. Lecture notes in computer science, vol 5575. Springer, Berlin, pp 61–70
23. Ali N, Bajwa KB, Sablatnig R, Chatzichristofis SA, Iqbal Z, Rashid M et al (2016) A novel image retrieval based on visual words integration of SIFT and SURF. PLoS ONE 11(6):e0157428
24. Ali N, Bajwa KB, Sablatnig R, Mehmood Z (2016) Image retrieval by addition of spatial information based on histograms of triangular regions. Comput Electr Eng 54:539–550. https://doi.org/10.1016/j.compeleceng.2016.04.002
25. Zheng Y, Huang X, Feng S (2010) An image matching algorithm based on combination of SIFT and rotation invariant LBP. J Comput Aided Des Comput Graph 2:286–292
26. Ojala T, Pietikanen M, Maenpaa T (2002) Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
27. Wang X, Han T, Yan S (2009) An HOG-LBP human detector with partial occlusion handling. In: IEEE 12th international conference on computer vision, pp 32–39
28. Heikkilä M, Pietikäinen M, Schmid C (2009) Description of interest regions with local binary patterns. Pattern Recogn 42:425–436
29. Gupta R, Patil H, Mittal A (2010) Robust order-based methods for feature description. In: Proceeding of the IEEE conference on computer vision and pattern recognition, pp 334–341. https://doi.org/10.1109/cvpr.2010.5540195
30. Murala S, Maheshwari RP, Balasubramanian R (2012) Local tetra patterns: a new feature descriptor for content based image retrieval. IEEE Trans Image Process 21(5):2874–2886
31. Murala S, Wu QMJ (2013) Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval. Neurocomputing 119:399–412
32. Srivastava P, Binh NT, Khare A (2014) Content-based image retrieval using moments of local ternary pattern. Mobile Netw Appl 19(5):618–625
33. Murala S, Jonathan WQ (2014) Local mesh patterns versus local binary patterns: biomedical image indexing and retrieval. IEEE J Biomed Health Inf 18(3):929–938

34. Dubey SR, Singh SK, Singh RK (2015) Local bit-plane decoded pattern: a novel feature descriptor for biomedical image retrieval. IEEE J Biomed Health Inf PP(99):1

35. Dubey SR, Singh SK, Singh RK (2015) Local diagonal extrema pattern: a new and efficient feature descriptor for CT image retrieval. IEEE Signal Process Lett 22(9):1215–1219

36. Dubey SR, Singh SK, Singh RK (2015) Local wavelet pattern: a new feature descriptor for image retrieval in medical CT databases. IEEE Trans Image Process 24(12):5892–5903

37. Rao LK, Rao DV (2015) Local quantized extrema patterns for content-based natural and texture image retrieval. Hum Centric Comput Inf Sci. https://doi.org/10.1186/s13673-015-0044-z

38. Deep G, Kaur L, Gupta S (2016) Directional local ternary quantized extrema pattern: a new descriptor for biomedical image indexing and retrieval. Eng Sci Technol Int J 19:1895–1909

39. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal Mach Intell 27:1615–1630. https://doi.org/10.1109/TPAMI.2005.188

40. Liu GH, Yang JY (2008) Image retrieval based on the texton co-occurrence matrix. Pattern Recogn 41(12):3521–3527

41. Liu G, Zhang L, Hou Y, Li Z, Yang J (2010) Image retrieval based on multi-texton histogram. J Pattern Recognit 43:2380–2389

42. Liu GH, Li ZY, Zhang L, Xu Y (2011) Image retrieval based on micro-structure descriptor. Pattern Recogn 44(9):2123–2133

43. Liu GH, Yang JY (2013) Content-based image retrieval using color difference histogram. Pattern Recogn 46(1):188–198

44. Xingyuan W, Zongyu W (2013) A novel method for image retrieval based on structure elements' descriptor. J Vis Commun Image Represent 24:63–74

45. Seetharaman K, Sathiamoorthy S (2016) A unified learning framework for content based medical image retrieval using a statistical model. J King Saud Univ Comput Inf Sci 28(1):110–124

46. Hu R, Collomosse J (2013) A performance evaluation of gradient field HOG descriptor for sketch based image retrieval. Comput Vis Image Underst 117:790–806

47. Zafar B, Ashraf R, Ali N, Iqbal MK, Sajid M, Dar SH, Ratyal NI (2018) A novel discriminating and relative global spatial image representation with applications in CBIR. Appl Sci 8(11):2242. https://doi.org/10.3390/app8112242

48. Zafar B, Ashraf R, Ali N, Ahmed M, Jabbar S, Chatzichristofis SA (2018) Image classification by addition of spatial information based on histograms of orthogonal vectors. PLoS ONE. https://doi.org/10.1371/journal.pone.0198175

49. Zafar B, Ashraf R, Ali N, Ahmed M, Jabbar S, Naseer K, Ahmad A, Jeon G (2018) Intelligent image classification-based on spatial weighted histograms of concentric circles. Comput Sci Inf Syst 15:25. https://doi.org/10.2298/CSIS180105025Z

50. Latif A, Rasheed A, Sajid U, Ahmed J, Ali N, Ratyal NI, Zafar B, Dar SH, Sajid M, Khalil T (2019) Content-based image retrieval and feature extraction: a comprehensive review. Math Probl Eng. https://doi.org/10.1155/2019/9658350

51. Ali N, Zafar B, Iqbal MK, Sajid M, Younis MY, Dar SH et al (2019) Modeling global geometric spatial information for rotation invariant classification of satellite images. PLoS ONE 14(7):e0219833. https://doi.org/10.1371/journal.pone.0219833

52. Sajid M, Ali N, Dar SH, Ratyal NI, Butt AR, Zafar B, Shafique T, Baig MJA, Riaz I, Baig S (2018) Data augmentation-assisted makeup-invariant face recognition. Math Probl Eng. https://doi.org/10.1155/2018/2850632

53. Jayaraj D, Sathiamoorthy S (2019) Deep learning based depthwise separable model for effective diagnosis and classification of Lung Ct images. Int J Eng Adv Technol (IJEAT). https://doi.org/10.35940/ijeat.A1439.109119

54. Jayaraj D, Sathiamoorthy S (2019) Computer aided diagnosis system using watershed segmentation with exception based classification model for Lung CT images. Int J Sci Technol Res (IJITEE) 8(4):46

55. Jayaraj D, Sathiamoorthy S (2019) Deep neural network based classifier model for lung cancer diagnosis and prediction system in healthcare informatics, intelligent data communication technologies and internet of things. Lect Notes Data Eng Commun Technol. https://doi.org/10.1007/978-3-030-34080-3

56. Ratyal N, Taj IA, Sajid M, Mahmood A, Razzaq S, Dar SH, Ali N, Usman M, Baig MJA, Mussadiq U (2019) Deeply learned pose invariant image analysis with applications in 3D face recognition. Math Probl Eng. https://doi.org/10.1155/2019/3547416

57. Zhao M, Zhang H, Sun J (2016) A novel image retrieval method based on multi-trend structure descriptor. J Vis Commun Image Represent 38:73–81

58. Natarajan M, Sathiamoorthy S (2019) Content based medical image retrieval using multi-trend structure descriptor and fuzzy k-NN classifier. In: 4th international conference on communication and electronics systems (ICCES 2019), 17–19 July 2019, Department of ECE, PPG Institute of Technology, Coimbatore

59. Natarajan M, Sathiamoorthy S (2019) Heterogeneous medical image retrieval using multi-trend structure descriptor and fuzzy SVM classifier. Int J Recent Technol Eng (IJRTE) 8(3):3958–3963

60. Seetharaman K, Kamarasan M (2012), A smart color image retrieval method based on multiresolution features. In: IEEE international conference on computational intelligence and computing research, pp 68–75

61. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL (2007) Open access series of imaging studies (OASIS): cross sectional MRI data in young, middle aged, nondemented, and demented older adults. J Cogn Neurosci 19(9):1498–1507

62. Fazal M, Baharum B (2013) Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain. J King Saud Univ Comput Inf Sci 25(2):207–218

63. Sathiamoorthy S, Kamarasan M (2014) Content based image retrieval using the low and higher order moments of BDIP and BVLC. Int J Innov Res Sci Eng Technol 3(1):8936–8941

64. NEMA-CT Image Dataset. ftp://medical.nema.org/medical/Dicom/Multiframe/

65. VIA/I-ELCAP CT Lung Image Dataset (Online). http://www.via.cornell.edu/-databases/lungdb.html