Computers and Electrical Engineering 000 (2017) 1–11

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Multi-layer multi-level color distribution – User feedback model with wavelet analysis for color image retrieval[☆]

A Kandasamy[a,*], Dr. M Sundaram[b]

[a] Research Scholar, Anna University chennai, Chennai - 600 025, Tamilnadu, India
[b] Professor, Department of Electronics and Communication Engineering, V.P.M.M. Engineering College for Women, Virudhunagar District, Tamilnadu - 626 190, India

ARTICLE INFO

Article history:
Received 7 March 2017
Revised 15 June 2017
Accepted 23 June 2017
Available online xxx

Keywords:
Image retrieval
Large database
Color histogram
Color distribution
Wavelet analysis
MLMLCD

ABSTRACT

To avoid misclassification during retrieval,this paper proposes an efficient multi-layer multi-level color distribution (MLMLCD) approach to improve the image retrieval quality. Here, the MLMLCD Vector Generation stage applies the wavelet transform over the image layers and hence color distribution vectors are generated. In MLMLCD Image Retrieval stage, the similarity measurement of MLMLCD color distribution vector value is made between the proposed technique and the values from larger databases. Finally, the precise retrieval result is produced with user feedback and query model, which is iterated over several runs. The performance of the proposed technique is tested between two datasets namely: McGill and CalTech database. Here, the performance is tested in terms of retrieval efficiency, classification rate and time complexity. A higher retrieval efficiency (98.5%)with less false classification rate(3.4%) is achieved, when compared with conventional techniques and a significance improvement is noted.

1. Introduction

The large graphical database contains several categories of images. Each image has huge similarity to other few images. The images present in the database can be classified into few categories. As of the size of backend is higher, the features present in the similar images has little deviation. For example, the dog image present in the database contains dog feature but the size and color of the dog would be different. Also there will be little deviation of color in hair of dogs in few places. Such micro deviation in the feature would change the result of query. Now a days identifying similar color images from large database are challenges to the researchers. Retrieving related images from large database has been handled with many features like color, shape and texture. However, the color value of the image plays the vital role in all the cases.

Content-based image retrieval is an efficient method which automates retrieval of images with respect to its salient features [23]. The image retrieval has been performed in many ways using above mentioned features and for the classification there are number of algorithms has been used. The popular K-means algorithm computes the image similarity using the color values and edge similarity. Similarly there are number of algorithms has been used in earlier days. The efficiency of image retrieval is highly depending on the feature being considered.

The color distribution is the measure which is computed based on the number of pixels being affected with any particular color. There are number of colors possible for the image pixel, each colors contains unique color value. There may

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. V. Suma.

* Corresponding author.
E-mail addresses: akandasamy052@gmail.com (A. Kandasamy), cm_sundaram2001@yahoo.co.in (Dr.M. Sundaram).

http://dx.doi.org/10.1016/j.compeleceng.2017.06.026
0045-7906/© 2017 Elsevier Ltd. All rights reserved.

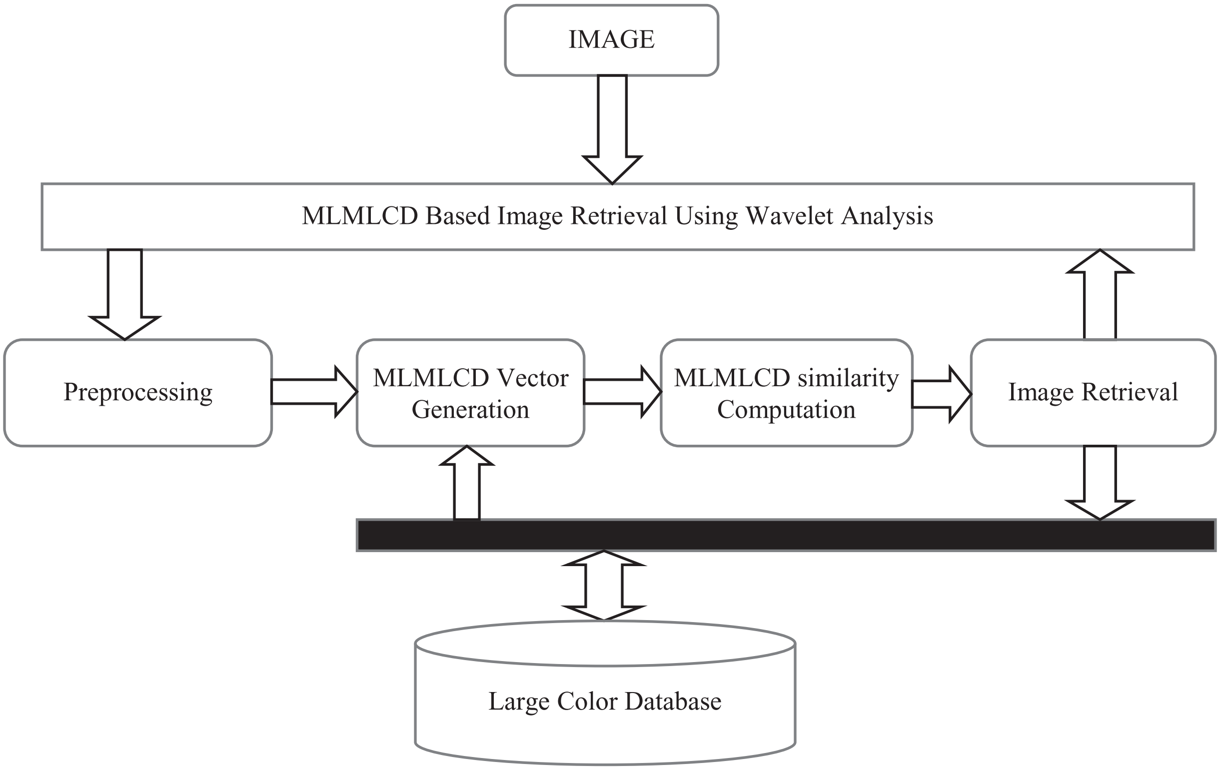

Fig. 1. Architecture of MLMLCD based image retrieval.

be N number of pixel with the same value. By counting the number of pixels affected for the value considered, the feature distribution of the color value can be computed. Each similar image has reasonable closure feature distribution than other images. The colors are basically classified from Red Green and Blue (RGB) consortium. The mixtures of RGB are giving numerous level color combinations. Each combination has unique color codes. Even though the color values are different, they affect only in the red layer not on the blue layer.

In this paper, the color histogram is used for the classification process and generates the feature vector based on the color values of pixels. The images would have any object and the color of the object also vary according to the type. The color histogram is computed by counting the number of color values present in the image for each distinct color values. Then the probability of distribution is computed to restore the pixel with new value [1]. This method concentrates on image layers with resemblance in the bottom layers than the top layer and computation of the color distribution resolves the problems associated with image retrieval process. For any image retrieval algorithm, the efficiency over time complexity is much important and wavelet transform is applied to improve the efficiency. The wavelet transform reduces the signals with less distortion and the images are transformed to different levels by varying the wavelet parameters [4]. The color distribution vector is generated over each wavelet transform output and produces three different distribution vectors for each level. The Multi-Layer Multi-Level Color Distribution (MLMLCD) vector is used in computing the similarity between the images. The proposed Multi-Layer Multi-Level Color Distribution (MLMLCD) method is evaluated over two database for computing its retrieval efficiency, false classification rate and time complexity.

In preprocessing stage, the method generates three level images for a single image from the selected database and the proposed method generates three different color distribution vectors. The entire phase has been split into number of stages namely preprocessing, MLMLCD Vector Generation, MLMLCD Image Retrieval. The different images are generated by applying wavelet transform (stage 1) and using the color distribution vector (stage 2). Hence, the methods computes the color distribution similarity value to perform image retrieval [7]. The architecture of MLMLCD using color image retrieval is shown in Fig. 1.

The outline of the paper is mentioned as follows: The Section 2 provides insights on conventional techniques used in Content Based Image Retrieval. Section 3 gives the proposed content based image retrieval with wavelet analysis. Section 4 evaluates the proposed work with conventional techniques. Finally, Section 5 concludes the paper with future work.

2. Related works

There are number of image retrieval methods have been discussed earlier and this section briefs list of methods relate to the problem.

Content Based Image Retrieval using Color Edge Detection and Discrete Wavelet Transform, generates feature vectors that combines both color and edge features. Here, wavelet transform is used to reduce the feature vector size and simultaneously preserves the content using Manhattan distance [25] measurement [2]. Such system with similarity index measurement [24] helps in improving the image retrieval quality. The robustness of the system is also tested against query image alterations such as geometric deformations and noise addition etc. Statistical Tests of Hypothesis Based Color Image Retrieval examines the input query image using textured or structured information (shape features [21]). In structured image, the shapes are segregated into various regions. The textured image contains no shapes and it is considered as a single object for experimentation purpose. The database in the proposed method uses both the images [3].

Image Retrieval based on the combination of Color Histogram and Color Moment [5] feature vectors are combined. Content-based image retrieval using color moment and Gabor texture feature [6] uses color features, color moments of the Hue, Saturation and Value (HSV) component images in HSV color space [11]. Since, it uses texture features, Gabor texture descriptors are adopted. Here, the weights are assigned to each feature and similarity is calculated using the combined features of color and texture according to normalized Euclidean distance [9] for reduced time complexity.

A Novel Image Retrieval Method using Segmentation and Color Moments [8], propose a color image retrieval method based on the primitives of color moments, which are:

$$u_i = \frac{1}{N}\sum_{j=1}^{N} f_{i,j} \tag{1}$$

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(f_{i,j} - u_i\right)^2} \tag{2}$$

$$s_i = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}\left(f_{i,j} - u_i\right)^3} \tag{3}$$

Here, the image is divided into four segments and the color moments of all segments are extracted and clustered into different classes. A Local Structure Descriptor (LSD) [10], based on a similarity of edge orientation and LSD (effectively combine color, texture and shape as a whole for image retrieval) is constructed for Color Image Retrieval.

A Geolocation-based image retrieval method identifies the geo location or geo-tagged photos [18] of an image with visual attention and color layout descriptors from Flickr database. The method further refines the images through the fusion of unsupervised principal component analysis (PCA). Here, the geo-tagging is used over both high and low level features of an image [15,17]. There exist always a difference between the high level features and low level feature, neural network is used to reduce this difference. To reduce such difference, Content Based Image Retrieval utilizes neural network for interpreting the image contents and match the query image from complex database [16]. Certain image retrieval system using multiple regions is used to identify the watershed regions-of-interest (ROI) within an input image. This helps in finding the relationship between semantics and visual elements using Bayesian learning from positive classification results [19]. Similar to [15], image refining retrieval uses exploitation and fusion using PCA and spectral clustering [20].

3. Content based image retrieval using modified wavelet analysis

3.1. Preprocessing

The input image is read and the color histogram of the each image is computed. Initially, the method identifies the list of unique color values at each layer. The method splits the image into three layers namely RGB and the value of which is found. For each layer, the methods generate the color histogram and the probability of distribution(PDF) values is found. Finally, the method restores the value with the PDF value and the histogram equalization helps in improving the quality of the image, which is given as:

$$T(k) = floor\left((L-1)\sum_{n=0}^{k} p_n\right) \tag{4}$$

where, *floor*() rounds down to the nearest integer, L represent possible pixel intensities and p_n represents the normalized histogram of f. The normalized histogram is defined as the ratio of number of pixels with n intensity to the total pixels. The algorithm (Algorithm1) and flow chart (Fig. 2) express the histogram equalization of color images, which is shown as:

The pre-processing algorithm and flowchart computes the probability of distribution value over each image layer. Here, each layer uses histogram equalization [22] to enhance the quality of the image layers.

Please cite this article as: A. Kandasamy, Dr.M. Sundaram, Multi-layer multi-level color distribution – User feedback model with wavelet analysis for color image retrieval, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.06.026

Algorithm 1 Preprocessing histogram equalization.

Input: Image Im
Output: Enhanced Image EI.
Start
  Read input image Im
  Split image into number of levels.
  Im = $\int_{i=1}^{3} Split(Im, i)$
  For each level l from Im
    Identify unique values uv.
    uv = $\Sigma Distinct(l) \in L$
    For each value v from uv
      Compute probability of distribution Pdf.
      Restore values with the pixels identified.
    End
  End
Stop

Algorithm 2 To calculate color distribution vector.

Input: Image Im
Output: Color Distribution Vector Cdv.
Start
  Read Input image Im.
  Initialize wavelet transform with three levels.
  Generate transform image Tm.
  Tm = $\int_{i=1}^{3} WaveletTransform(Im, i)$
  For each image I from Tm
    Generate multi-layer image Mlm.
    Mlm = $\int_{i=1}^{3} Split(I, i)$
    For each layer l from Mlm
      Compute color histogram Ch.
      Ch = $\sum Distinctvalue(l)\ \nexists ch$
      For each value v from Ch
        Compute distribution value Dv = $\frac{Tn(c)\in l}{size(Ch)} \times size(l)$ (6)
      End
      Generate distribution vector Dvset(I,l) = Dv.
    End
  End
Stop.

3.2. Multi-layer multi-level color distribution vector generations

In this stage, the method reads the input image layers and wavelet transform is applied over it. The wavelet transform [12] is formulated as:

$$DWT(m,k) = \frac{1}{a}\sum_{n=0}^{N-1} s(n) g\left(\frac{k-b}{a}\right) \tag{5}$$

where, a is a scaling parameter and b is a translation parameter, m is the decomposition level index, g is the mother wavelet and $s(n)$ is the original signal.

After the application of wavelet transform in each level, the proposed method generates an image and splits each image into three layers. Then the color histogram is computed over each layer and the distinct value is used to generate the color distribution vectors of each layer and performs the image retrieval process. The color distribution process is computed using the total occurrence of the color value considered and the number of pixels being affected with the same and total number of pixels. The vector generation process is shown in algorithm (Algorithm 2) and flow chart (Fig. 3).

The Color Distribution Vector algorithm and flowchart computes the color distribution vector for each input image of the database. Generated distribution vector is used in performing the image classification and retrieval.

3.3. MLMLCD similarity measurement

To measure the similarity value the method used to compute the multi-layer multi-level color distribution. Each layer color distribution vector of image is considered, the method computes the color distribution similarity with other levels of same distribution vectors [13]. This is iterated to each layer and the method computes the color distribution similarity. The similarity value is computed based on the distance between the values of color distribution vector. Finally a cumulative MLMLCD measure is computed, which is shown in Algorithm 3.

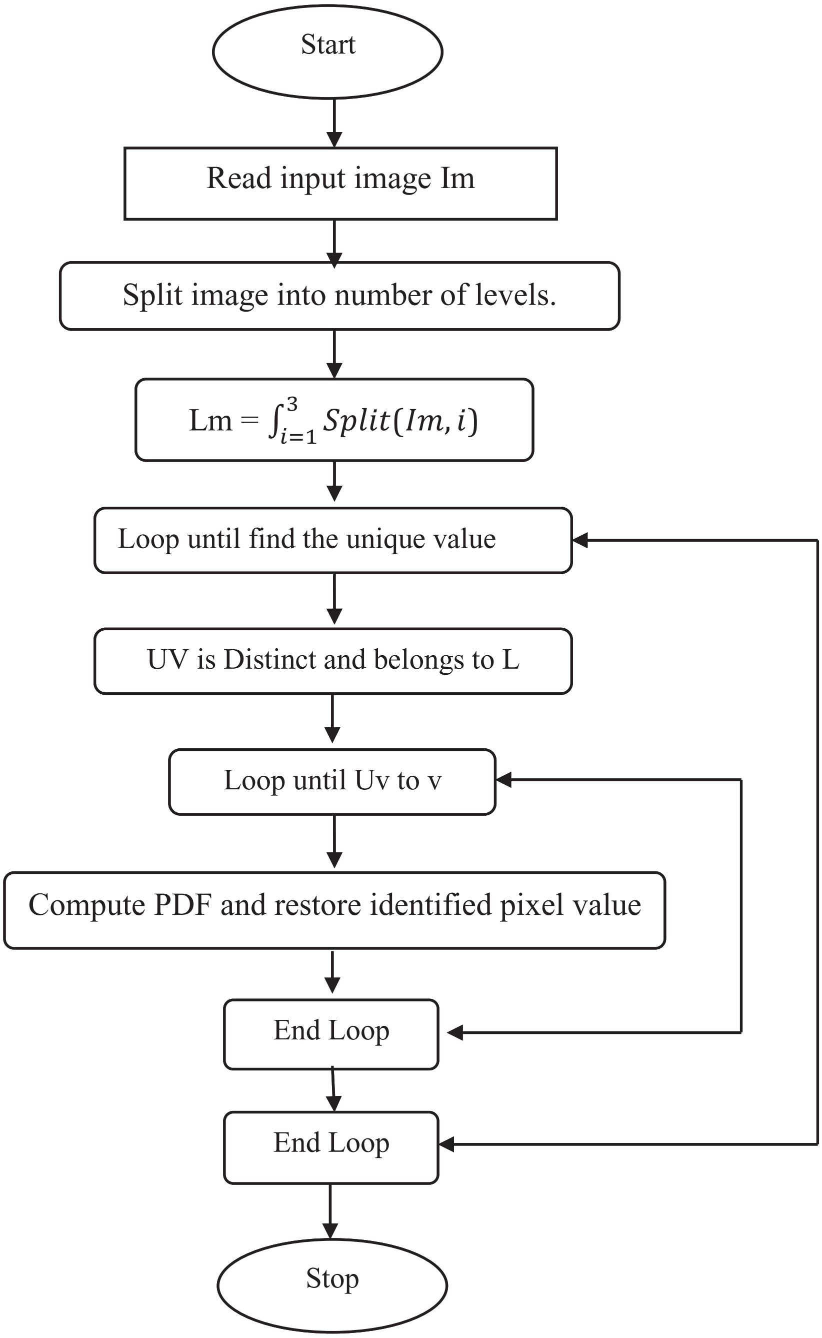

Fig. 2. Preprocessing histogram equalization.

3.4. User feedback model based image retrieval

The image retrieval is performed through the computation of the multi-layer multi-level color distribution value at each image vector. At the training stage, the method computes the multi-layer multi-level color distribution vector and the vector computes the similarity measure [14]. For each images distribution vector, the method computes the similarity value and identifies the most relevant image. The method receives the feedback from the user and based on its value, the method iterates the query to produce more precise results.

Once the input image input image, I is given as input. The image I is enhanced (EI or Enhanced Image) using continuous iteration over pre-processed image I. The MLMLCD vector v is then generated over processing the MLMLCD generation over

Please cite this article as: A. Kandasamy, Dr.M. Sundaram, Multi-layer multi-level color distribution – User feedback model with wavelet analysis for color image retrieval, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.06.026

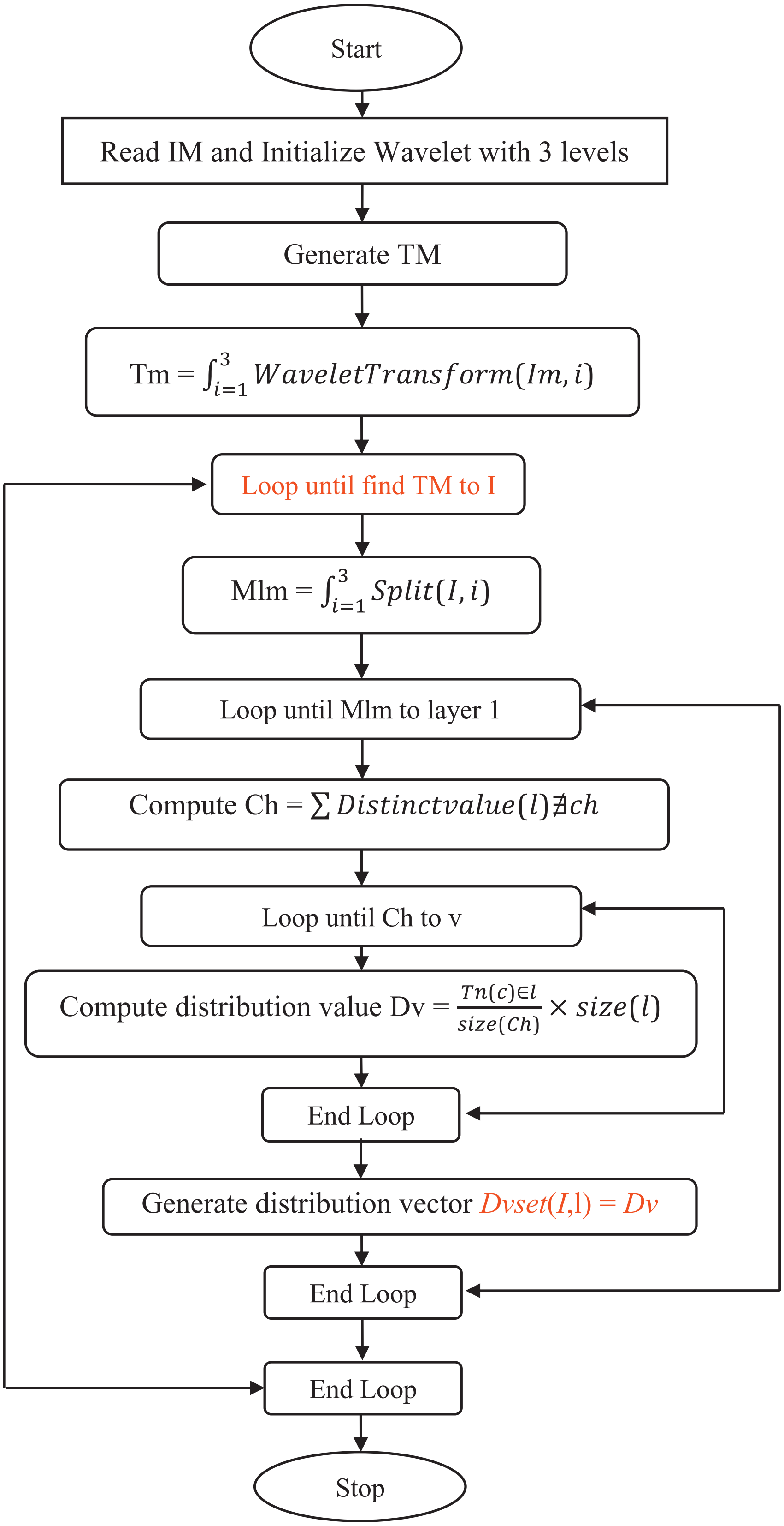

Fig. 3. Calculate color distribution vector.

Algorithm 3 Color distribution.

Input: MLMLCD Vector V, MLMLCD Vector Vset.
Output: MLMLCD-Sim
Start
    For each level vector lv
    For each layer l
        Compute MLMLCD-Similarity.
        MLMLCD-Sim $=\int_{i=1}^{size(lv)} \int_{j=1}^{size(l)} \int_{k=1}^{size(level)} \frac{\sum Distance(v(l)(k),Vset(m)(l)(k)}{size(l)}$ (7)
    End
  End
Stop.

Algorithm 4 Feedback model.

Input: Image I, MLMLCD Vector set Vs.
Output: Relevant image RI
Start
  Read input image I
  Enhanced image EI = Perform pre-processing (I)
  MLMLCD vector v = perform MLMLCD generation (EI)
  For each vector a from Vs
    MLMLCD-sim = Compute MLMLCD-Sim(a,v)
  End
  Choose the images with MLMLCD-Sim> Threshold.
  RI $=\Sigma MLMLCD - Sim(vs)(Image) > Th$ (8)
  While true
    Receive feedback Fb.
      Compute threshold according to Fb.
    Choose images with MLMLCD-sim>Threshold.
  End
Stop.

the EI images. During each vector ranging from *a* and Vs, the MLMLCD similarity value is computed using the similarity measurement between a and v. Once the iteration is completed, the MLMLCD-Similarity value is tested and if the value is greater than threshold then the RI value of the image is computed. Once the value of RI is greater than Th (Threshold), the feedback is received and the threshold is computed as per the feedback value. Finally, the images with lesser MLMLCD-Similarity value than the threshold value is used the final output image. The feedback model is shown in algorithm (Algorithm 4) and flow chart (Fig. 4).

The feedback model algorithm and flowchart computes the similarity value to perform image retrieval and based on computed similarity value the method selects most relevant images from the large database.

4. Results and discussion

The proposed MLMLCD approach has been implemented using Matlab programming tool and for the evaluation the method has been used various databases like McGill database with 1500 images and CalTech database with 186 images.

The retrieval efficiency results from Fig. 5 proved that the proposed MLMLCD method achieves less time complexity (98.5%) than the SURF-DCD (89%), locality sensitive method (87%) and local invariant feature method (83%). Thus the retrieval efficiency of the proposed method is proved better than the other methods due to its simplest computational process. The false classification results from Fig. 6 proved that the proposed MLMLCD method achieves lesser false classification results (3.4%) than the SURF-DCD (11%), locality sensitive method (12%) and local invariant feature method (16%). This leads to higher the retrieval efficiency in the proposed method than the other conventional. The results from Fig. 7 proved that the proposed MLMLCD method achieves less time complexity (34 s) than the SURF-DCD (52 s), locality sensitive method (61) and local invariant feature method (73 s). Here, the less time complexity of the proposed MLMLCD method is due to its simpler computational process. The results proved that the proposed method achieves better performance results than the other conventional methods.

5. Conclusion

This paper provides an efficient multi-layer multi-level color distribution similarity, which retrieves the image in an effective way with high retrieval efficiency, reduced false classification ratio and reduced time complexity. Here, the application of wavelet transform generates multiple image and the application of histogram generates the color distribution vector from the image layers. Finally, multi-layer multi-level color distribution similarity value is computed using color distribution vector and this value retrieves the image with increased accuracy. The accuracy is certainly improved with user relevant

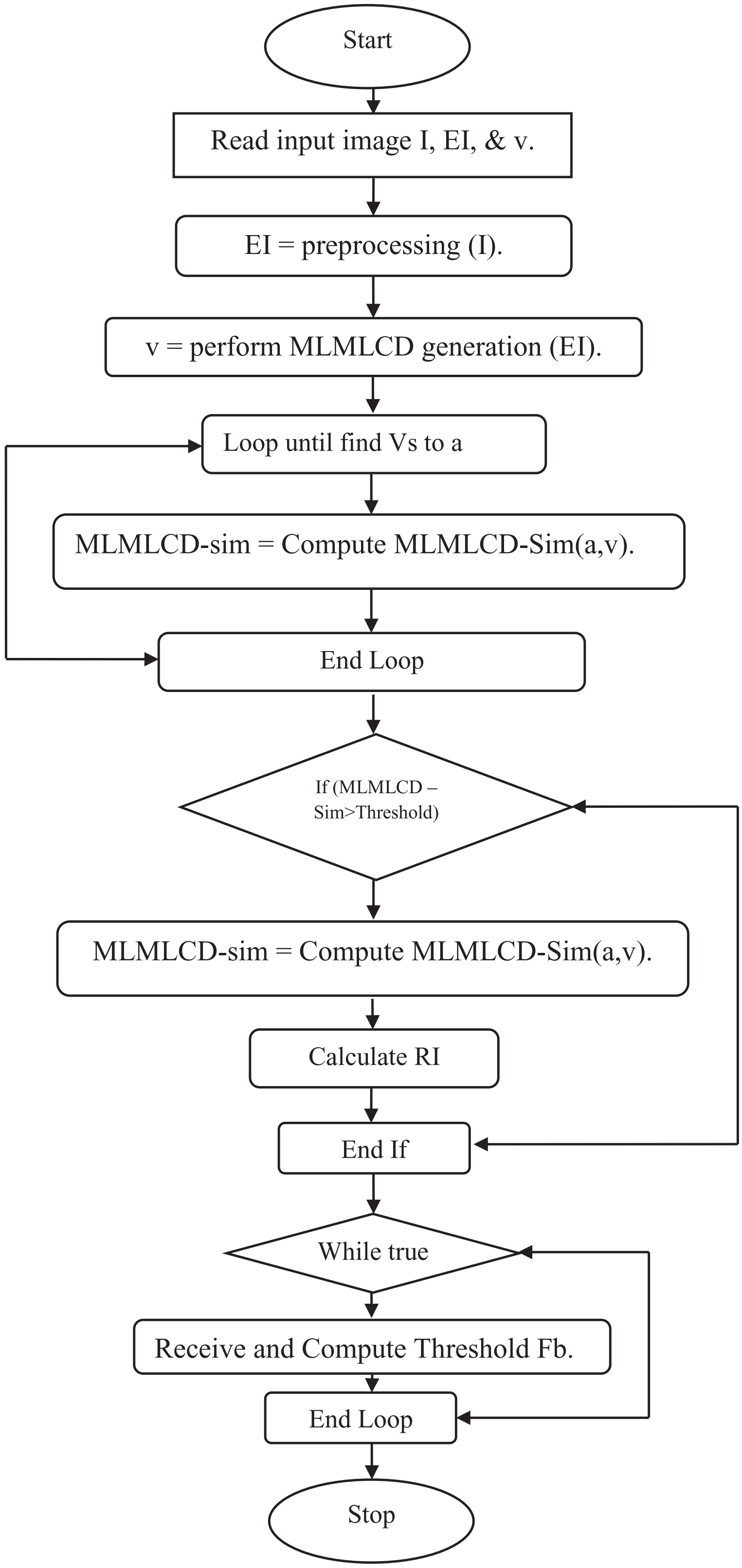

Fig. 4. Feedback model.

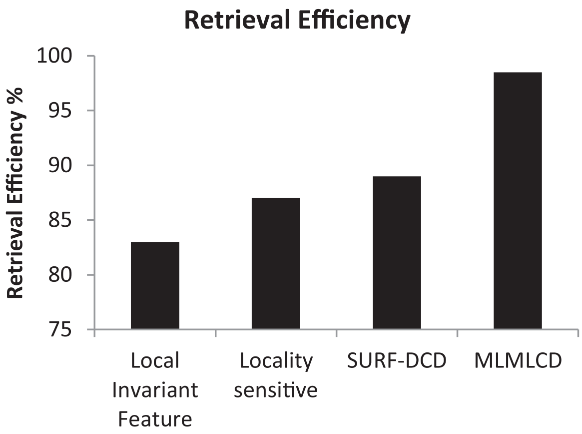

Fig. 5. Comparison of retrieval efficiency.

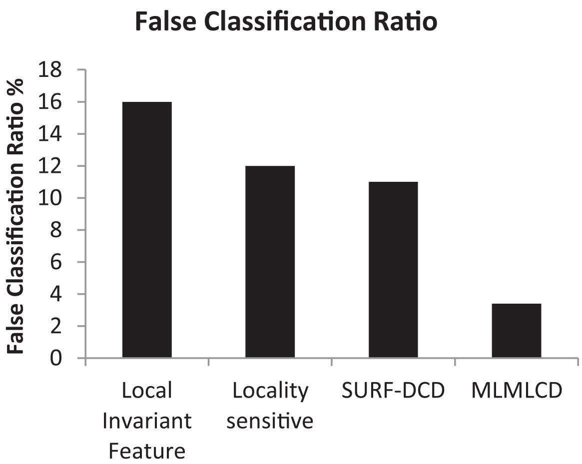

Fig. 6. Comparison of false classification ratio.

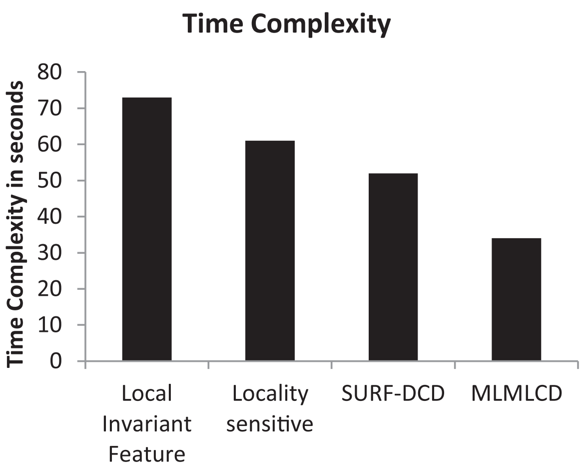

Fig. 7. Comparison of time complexity.

feedback and the results are sorted iteratively with such feedback rate. The retrieval results are efficient enough and further reduces the timing complexity. However, this method provides reduced advantages over noisy images with increased false classification ratio with reduced accuracy. Another limitation is that the presence of in homogeneities in magnetic resonance (MR) imaging with noise. Further, the work can be extended to improve the noisy characteristics to retrieve well the suitable objects using smoothing at individual layer. However, the study on noisy signal is further required, since the number of coefficients at each layer reduces considerably.

References

[1] Olmos A, Kingdom FAA. A biologically inspired algorithm for the recovery of shading and reflectance images. Perception 2004;33:1463–73.
[2] Agarwal Swathi. Content based image retrieval using color edge detection and discrete wavelet transform. In: IEEE transaction on issues and challenges in intelligent computing techniques (ICICT); 2014. p. 368–78.
[3] Seetharaman K, Selvaraj S. Statistical tests of hypothesis based color image retrieval. J Data Anal Inf Process 2016;4:90–9.
[4] Yang Y, Newsam S. Geographic image retrieval using local invariant features. IEEE Trans Geosci Remote Sens 2013;51:818–32.
[5] Singh MangijaoS, Hemachandran K. Article: image retrieval based on the combination of color histogram and color moment. Int J Comput Appl 2012;58(November(3)).
[6] Huang ZC, Chan PPK, Ng WWY, Yeung DS. Content-based image retrieval using color moment and Gabor texture feature. In: Poceedings of the IEEE ninth international conference on machine learning and cybernetics; 2010. p. 719–24.
[7] Kumar DK, Sree EV, Suneera K, Kumar PVCh. Cotent based image retrieval – extraction by objects of user interest. Int J Comput Sci Eng (IJCSE) 2011;3(3):1068–74.
[8] Saikrishna TV, Yesubabu A, Anandrao A, Rani TS. A novel image retrieval method using segmentation and color moments. ACIJ, Adv Comput 2012;3(1):75–80.
[9] Dubey RS, Choubey R, Bhattacharga J. Multi feature content based image retrieval. (IJCSE) Int J Comput Sci Eng 2010;02(06):2145–9.
[10] Zeng Zhiyong. A novel local structure descriptor for color image retrieval. MDPI. Information 2016;7(1):1–14.
[11] Wang XY, Yang HY, Li DM. A new content-based image retrieval technique using color and texture information. Comput Electr Eng 2013;39:746–61.
[12] Lee YH, Kim Y. Efficient image retrieval using advanced SURF and DCD on mobile platform. Multimedia Tools Appl 2015;74:2289–99.
[13] Kafai M, Eshghi K, Bhanu B. Discrete cosine transform locality-sensitive hashes for face retrieval. IEEE Trans Multimedia 2014;16:1090–103.
[14] Yang Y, Nie FP, Xu D. A multimedia retrieval framework based on semi-supervised ranking and relevance feedback. IEEE Trans Pattern Anal Mach Intell 2012;34:723–42.
[15] Memon MH, Li JP, Memon I, Arain QA. GEO matching regions: multiple regions of interests using content based image retrieval based on relative locations. Multimedia Tools Appl 2016:1–35.
[16] Shaikh RA, Deep S, Li JP, Kumar K, Khan A, Memon I. Contemporary integration of content based image retrieval. Wavelet active media technology and information processing (ICCWAMTIP), 2014 11th international computer conference on (pp. 301–304). IEEE; 2014.
[17] Memon MH, Khan A, Li JP, Shaikh RA, Memon I, Deep S. Content based image retrieval based on geo-location driven image tagging on the social web. Wavelet active media technology and information processing (ICCWAMTIP), 2014 11th international computer conference on (pp. 280–283). IEEE; 2014.
[18] Memon I, Chen L, Majid A, Lv M, Hussain I, Chen G. Travel recommendation using geo-tagged photos in social media for tourist. Wireless Pers Commun 2015;80(4):1347–62.
[19] Memon MH, Li JP, Memon I, Shaikh RA, Mangi FA. Efficient object identification and multiple regions of interest using CBIR based on relative locations and matching regions. Wavelet active media technology and information processing (ICCWAMTIP), 2015 12th international computer conference on (pp. 247-250). IEEE; 2015.
[20] Memon MH, Shaikh RA, Li JP, Khan A, Memon I, Deep S. Unsupervised feature approach for content based image retrieval using principal component analysis. Wavelet active media technology and information processing (ICCWAMTIP), 2014 11th international computer conference on (pp. 271–275). IEEE; 2014.
[21] Nikkam PushpalathaShrikant, Reddy EswaraB. A key point selection shape technique for content based image retrieval system. Int J Comput Vision Image Process (IJCVIP) 2016;6(2):54–70.
[22] Wong CY, Jiang G, Rahman MA, Liu S, Lin SCF, Kwok N. Histogram equalization and optimal profile compression based approach for colour image enhancement. J Visual Commun Image Represent 2016;38:802–13.
[23] Desai Rupali, Sonawane Bhakti. Gist, HOG, and DWT-based content-based image retrieval for facial images. In: Proceedings of the international conference on data engineering and communication technology. Singapore: Springer; 2017.
[24] Liang, Ru-Ze, et al. Optimizing top precision performance measure of content-based image retrieval by learning similarity function. (2016) arXiv preprint arXiv:1604.06620.
[25] Anandh A, Mala K, Suganya S. Content based image retrieval system based on semantic information using color, texture and shape features. Computing technologies and intelligent data engineering (ICCTIDE), international conference on. IEEE; 2016.

Please cite this article as: A. Kandasamy, Dr.M. Sundaram, Multi-layer multi-level color distribution – User feedback model with wavelet analysis for color image retrieval, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.06.026

A. Kandasamy, Research Scholar Anna University, Chennai, India. He passed B.E. computer science & engineering from Madurai Kamaraj University Tamil Nadu, India in the year April 1989. He received his master degree M.E computer science & engineering in the year April 1996 form Jadavpur University Calcutta, India. Area of Interest: Image Processing, Data Mining.

Dr. M. Sundaram, Professor/Department of Electronics and Communication Engineering, V.P.M.M. Engineering College for Women, Krishnankoil, Srivilliputtur Taluk, Virudhunagar District, Tamilnadu, India. He received his master degree M.E (Communication Systems) in the year December 2002 form Mepco Schlenk Engineering College, Madurai Kamaraj University, Tamil Nadu, India. He completed his Doctoral degree in the year of 2012, Anna University, Chennai, India. Area of Interest: Digital Image Processing, Medical Image Enhancement and Analysis and VLSI design.

Please cite this article as: A. Kandasamy, Dr.M. Sundaram, Multi-layer multi-level color distribution – User feedback model with wavelet analysis for color image retrieval, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.06.026