8-1-2019

A Novel Adaptive LBP-Based Descriptor for Color Image Retrieval

Mohammad Sotoodeh
Utah State University, mahmood.sotoodeh@aggiemail.usu.edu

Mohammad Reza Moosavi
Shiraz University

Reza Boostani
Shiraz University

Recommended Citation

Sotoodeh, Mahmood, Moosavi, Mohammad Reza, Boostani, Reza. "A Novel Adaptive LBP-Based Descriptor for Color Image Retrieval." Expert Systems with Applications, 127, 2019. https://doi.org/10.1016/j.eswa.2019.03.020

A Novel Adaptive LBP-Based Descriptor for Color Image Retrieval

Mahmood Sotoodeh [1, 2], Mohammad Reza Moosavi [1], Reza Boostani [1]

mahmood.sotoodeh@aggiemail.usu.edu, smmosavi@shirazu.ac.ir, boostani@shirazu.ac.ir
[1] Department of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran.
[2] Department of Computer Science, Utah State University, Logan, UT 84322-4205, USA.

Abstract

In this paper, we present two approaches to extract discriminative features for color image retrieval. The proposed local texture descriptors, based on Radial Mean Local Binary Pattern (RMLBP), are called Color RMCLBP (CRMCLBP) and Prototype Data Model (PDM). RMLBP is a robust to noise descriptor which has been proposed to extract texture features of gray scale images for texture classification.

For the first descriptor, the Radial Mean Completed Local Binary Pattern is applied to channels of the color space, independently. Then, the final descriptor is achieved by concatenating the histogram of the CRMCLBP_S/M/C component of each channel. Moreover, to enhance the performance of the proposed method, the Particle Swarm Optimization (PSO) algorithm is used for feature weighting.

The second proposed descriptor, PDM, uses the three outputs of CRMCLBP (CRMCLBP_S, CRMCLBP_M, CRMCLBP_C) as discriminative features for each pixel of a color image. Then, a set of representative feature vectors are selected from each image by applying k-means clustering algorithm. This set of selected prototypes are compared by means of a new similarity measure to find the most relevant images. Finally, the weighted versions of PDM is constructed using PSO algorithm.

Our proposed methods are tested on Wang, Corel-5k, Corel-10k and Holidays datasets. The results show that our proposed methods makes an admissible tradeoff between speed and retrieval accuracy. The first descriptor enhances the state-of-the-art color texture descriptors in both aspects. The second one is a very fast retrieval algorithm which extracts discriminative features.

Keywords: Image Retrieval; Color Radial Mean Completed Local Binary Pattern (CRMCLBP); Prototype Data Model (PDM); Particle Swarm Optimization (PSO); Feature weighting; K-means clustering.

1 Introduction

In recent years, with rapid expansion and growing popularity of digital technologies such as social networks and photo sharing platforms, people daily produce millions of images with different topics and contents and upload them on hosting services (Lin, Chen, & Wu, 2014), (Furht, 2008). Searching and indexing the huge amount of available images is only possible with promising information retrieval methods. Content Based Image Retrieval (CBIR) system is one of the most popular fields in information retrieval used for retrieving digital images (Hiwale & Dhotre, 2015). Therefore, designing accurate and fast CBIR systems has been become a hot research topic in the field of pattern recognition and artificial intelligence.

Descriptors of the early CBIR systems were developed to extract features from gray scale images. Since the use of color images on the Internet is increasing, extraction color information along with other features like texture and shape from these images can improve the performance of retrieval systems.

The CBIR system receives query image from user and then it uses a descriptor to extract features from the query image. By means of the extracted features, the query image is compared with the set of images in an image database and the most similar images from the database are chosen to be delivered to the user.

The main elements of a CBIR system are feature extraction and feature matching algorithms. The goal of the feature extraction algorithm is to extract high discriminative features in feasible time. Appropriate features should be robust to photometric and geometric deformations such as changes in viewpoint, scale, lighting conditions and occlusion. On the other hand, the feature matching algorithm uses a similarity measure, such as Euclidian, Chi-Square or Canberra distance, to compare the images based on the extracted features.

In general, feature extraction methods are divided into two groups: global and local. In the global approach, the features describe the whole image, without using any sliding local window on pixels. The global methods are robust to noise and have satisfactory computation time. On the other hand, they ignore the local information between neighboring pixels. These algorithms are sensitive to image variations like illumination, viewpoint, occlusion and background clutter. The well-known global feature descriptors are color histogram (Swain & Ballard, 1991), color moments (Stricker & Orengo, 1995), edge histogram(Park, Jeon, & Won, 2000) and texture co-occurrence matrix (Tuceryan & Jain, 1993).

Local descriptors consider the local regions of the image to extract features. These regions are commonly detected by special patches or set of key points whose size is smaller than the size of the image. In these local methods, the locality of data is preserved. The local descriptors are commonly sensitive to noise. Some popular local methods are SIFT (Lowe, 2004), HOG (Dalal & Triggs, 2005), SURF (Bay, Ess, Tuytelaars, & Van Gool, 2008), BRIEF (Calonder et al., 2012), LBP (Timo Ojala, Pietikainen, & Maenpaa, 2002) and so on.

Color is an intrinsic and obvious feature of the images. The histogram based color features are the well-known global descriptors. They are not only simple to implement but also robust to rotation and translation. In addition, some properties such as scale invariant is added to histogram based color features after normalizing by size of image (Manjunath, Ohm, Vasudevan, & Yamada, 2001; X. Y. Wang, Yu, & Yang, 2011).

Texture features are local patterns which are repeated in the images (Faugeras & Pratt, 1980). They are powerful features for describing images which can be represented as uniform and non-uniform patterns (Timo Ojala et al., 2002). In the last few decades, several methods have been presented to extract texture features from images. Gray Level Co-occurrence Matrices (GLCM) is one of the most well-known texture-based global methods for feature extraction (Haralick, Shanmugam, & Dinstein, 1973). Gabor filter is another successful global descriptor for texture features extraction (Manjunath & Ma, 1996). Rotation and scale invariant Gabor filter is one of the robust version of this descriptor which was proposed by Han and Ma (Han & Ma, 2007). High computation time for extracting features is the disadvantage of Gabor filter-based methods (Chen, Lu, & Zhang, 2004).

As mentioned above, an image contains various features such as color, shape and texture. The classic methods only use one type of features to describe an image. Moreover, most of them have been proposed for gray scale images and ignore the color information. In recent years, proposed methods consider the combination of these features to provide more powerful descriptors. Color SIFT (Burghouts & Geusebroek, 2009) is color version of SIFT. It is proposed to extract texture of color images. Although Color SIFT performs better than other color descriptors, it has a high computation time, especially when the size of image or size of database increases.

Li et al. (C. Li, Huang, & Zhu, 2017) proposed a texture retrieval method which is constructed by using Copula model (Kwitt, Meerwald, & Uhl, 2011) and the Gabor wavelets (T. S. Lee, 1996). In this method, copula is used to capture the color dependence and Gabor filter is utilized to model the cells of visual cortex of human. Vassou et al. (Vassou, Anagnostopoulos, Amanatiadis, Christodoulou, & Chatzichristofis, 2017) proposed a low level descriptor named Composite Moment (CoMo) for image retrieval. The method uses the combination of color information with seven statistical invariant moments and edge directivity descriptor (CEDD) (Chatzichristofis & Boutalis, 2008) to extract texture feature from image. Aggarwal et al. (Aggarwal, Sharma, Singh, Singh, & Kumar, 2019) used an orthogonal Fourier-Mellin moments (OFMMs)-based descriptor to extract powerful effective features to achieve an efficient biomedical image retrieval system.

One of the successful methods for texture feature extraction of gray scale images is local binary pattern (LBP) that was first proposed by Ojala et al. (Timo Ojala, Pietikäinen, & Harwood, 1996). Later, rotation invariant, uniform and completed versions of this method were introduced by them and other researchers who followed this approach (T Ojala, 1997), (Timo Ojala et al., 2002), (Guo, Zhang, & Zhang, 2010). LBP-based methods have several advantages. They are fast, easy to implement and invariant to monotonic intensity and illumination changes. These methods also have the ability to extract local information with high precision compared to other local texture descriptors.

In the recent years, several color versions of LBP descriptors have been proposed for extracting texture features of color images. Mäenpää et al. (Maenpaa, Pietikainen, & Viertola, 2002) applied the gray scale LBP descriptor on each channel of the color images independently in order to extract texture features. Later, they added the six sets of LBPs opponent color to the three channel set of color image to extract the cross correlation between them. The efficiency of this method is good when the dimension of feature vectors increases. After that, Mäenpää and Pietikäinen found out that it is not necessary to use all six components to obtain cross correlation between three channels and only three pairs of them is sufficient (Mäenpää & Pietikäinen, 2004). Choi et al. (Choi, Plataniotis, & Ro, 2010) chose the YC_bC_r color space and then used the LBP histogram of each channel to extract texture features. They applied PCA to reduce dimension of feature vector. This method was proposed for face recognition application.

Local color vector binary patterns (LCVBP) is a color descriptor which was proposed by Lee et al. (S. H. Lee, Choi, Ro, & Plataniotis, 2012). LCVBP utilized histogram of color norm patterns and color angular patterns to extract discriminative features for face recognition. The color vector in each specific location on a defined neighborhood pixel have to be constructed by concatenating all of the components of the color image. The norm of this vector is used in uniform LBP to compute color norm pattern. For the color angular pattern, first, the ratio of pixel values between a pair of spectral-band images (i.e.… between R and G or R and B in RGB space) is computed to obtain directional information of color vector effectively. Then, the color angle is calculated by taken the inverse tangent of this ratio and the uniform LBP of this angle is considered as a color angular pattern.

OC-LBP operator is an effective version of LBP operator to reduce the dimensionality of LBP features (Zhu, Bichot, & Chen, 2013). First, the neighborhood of the corresponding pixel is divided into two non-overlapped orthogonal groups: diagonal and horizontal-vertical. Then, the original LBP operator is performed on each of the groups separately, and the results are concatenated together. For example, eight neighbor pixels are separated into two sets of four pixels. Each set contains 16 binary patterns, hence the total number of patterns is 32, which is much less than the original LBP (i.e., 256 patterns).

Different color models have been proposed by Zhu et al. (Zhu et al., 2013) as the extensions of the OC-LBP. One of the successful models is RGB-OC-LBP which is applied on three channels of color image. The total number of patterns for this model is 96 (32 patterns for each channel) versus 768 patterns of original LBP. Zhu stated that the RGB-OC-LBP is not only more efficient than the original LBP but also has high discriminative power.

Quaternion local ranking binary pattern (QLRBP) is local color descriptor which has been presented by Lan et al. (Lan, Zhou, & Tang, 2016). Quaternion (Hamilton, 1866) is a complex number with one real and three imaginary parts. In this method, the imaginary parts of this four dimensional number has been used to represent a color pixel in an image. They applied a window with 3×3 neighborhood of color pixel. In this window, a reference color pixel (r', g', b') and a color pixel (r, g, b) are utilized to derive the QLRBP operator. The Clifford Translation of Quaternionic (CTQ) and a rank based LBP method are used in the 3×3 window to code and rank the color pixels. Since the reference vector (r', g', b') is considered for the whole image, the local information could not be completely described. Therefore, the performance of this method is not high for image retrieval.

Another method named Multispectral Local Binary Pattern (MSLBP) (Maenpaa et al., 2002) applied LBP operator on each spectrum of the color image in RGB space independently. It also utilizes LBP on the cross-correlation of six pairs of opponent colors to capture the spatial relationship between spectra. Although, computation time of this method is very high for image retrieval application, it provides good recognition rate.

Dubey et al. (Dubey, Singh, & Singh, 2016) proposed two multichannel decoded local binary pattern methods which use two transformation functions named adder and decoder to encode the relationship between local binary patterns of channels. They are named multichannel adder based Local Binary Pattern (maLBP) and multichannel decoder based Local Binary Pattern (mdLBP). For obtaining these descriptors in RGB space, the local binary pattern of each channel is computed as a $LBP_j^i(x, y)$ which indicates the i^{th} bit of LBP code of j^{th} channel at a pixel location (x,y). The $LBP_j^i(x, y)$ is 0 or 1 for any bit of binary pattern with length 8 in 3 channels. Thus, the four and eight distinct values are generated for $maLBP^i(x, y)$ and $mdLBP^i(x, y)$ respectively. The histogram of values of two operators are computed in each channel as a feature vector. It should be noted that the histogram bins can have values between 0 to 256. Although this scheme has a good performance for recognition application, the length of feature vector is too long.

Completed local similarity pattern (CLSP) was introduce by Lie et al. (J. Li, Sang, & Gao, 2016) to extract features of color images. Two main parts of this method are color labeling and local similarity pattern. In the first part, standard k-means clustering is applied to color feature vectors of color image to generate a k-color words dictionary $W = \{w_1, w_2, \dots, w_k\}$. It is noted that each of the elements of this dictionary indicates a center of cluster words which has three dimensions $w_i = (r_i, g_i, b_i), i = 1, 2, \dots, k$ in RGB space. For encoding the color words of each pixel *p*, the localized soft-assignment coding approach (L. Liu, Wang, & Liu, 2011) is used. Notice that before applying this function, first the Euclidean distance between each pixel of color image and each cluster center should be calculated. The goal of second part of CLSP method (i.e. local similarity pattern) is to encode the similarity between center pixel and its neighbor pixels in a 3×3 window.

First, the color distance between a certain pixel and its neighbors is computed to find the nearest neighbors of that pixel. Then, the local similarity pattern (LSP) is obtained by using these nearest neighbors. After computing the color label and LSP part, the joint distribution (2D histogram) of their values is constructed.

Finally, the 1D histograms of this 2D matrix are considered as feature vectors. In spite of having such a large size feature vector, the accuracy of this method is not very high.

Singh et al. (Singh, Walia, & Kaur, 2018) proposed a color version of LBP operator named LBPC to extract texture color of color images. This operator is suggested in a vector space with dimension *DIM* to partition color pixels by using an appropriate hyperplane in a local window. The size of local window is determined by $(2R+1)\times(2R+1)$ formula and $\vec{v}_c = (r_c, g_c, b_c)$ is considered as a corresponding vector for window center c (which is called reference point). In the local window, a neighbor pixel p is indicated by $\vec{v}_p = (r_p, g_p, b_p)$. Therefore, a color plane Q in the color space is defined by a reference point $\vec{v}_0 = (r_0, g_0, b_0)$ on a plane and a normal vector $\vec{u} = (u_1, u_2, u_3)$ which is perpendicular on color plane. The connected line between black and white pixels in RGB space determines the normal vector of the color plane. According to this plane, the neighbor pixels $\vec{v}_p,\ p = 1, 2, \dots, P$, are classified into two groups: those which are placed on or above the plane and those which are located below the plane. The histogram of this representation is used as LBPC features.

In this paper, we propose two local texture descriptors for image retrieval named Color Radial Mean Completed Local Binary Pattern (CRMCLBP) and Prototype Data Model (PDM). These methods are based on Radial Mean Local Binary Patterns (RMLBP), a robust to noise method proposed by Shakoor and Boostani to classify texture of gray scale images in appropriate time (Shakoor & Boostani, 2018). The RMLBP method does not consider color textures and their relations.

First, we present the color version of this method by applying it to three channels of RGB space independently which is called CRMCLBP. We use the CLBP with riu2 mapping instead of uniform LBP method to obtain features with high discriminative power.

CLBP framework generates six output operators. The three main operators of this method are CLBP_S, CLBP_M and CLBP_C (Guo et al., 2010). CLBP_S and CLBP_M are built by comparing the sign and magnitude of the gray value of central pixel of each local window with its neighbors respectively. CLBP_C is constructed by comparing the gray values of each central pixel with the average gray value of the whole image. The rest of the operators are constructed by two ways of combination of these three basic operators: in concatenation and jointly. They are CLBP_S_M/C, CLBP_S/M and CLBP_S/M/C. In the combination operators, the sign "/" indicates joint and the "_" shows the concatenation. For example, to construct CLBP_S_M/C, first, the histogram of CLBP_S is calculated. Then it concatenates to 2D joint histogram of CLBP_M and CLBP_C. Geo et al. have indicated that the CLBP_S/M/C operator is more powerful than the others for feature extraction of texture images (Guo et al., 2010). It is built by 3D joint histogram of three main output operators, which is called CLBP_S/M/C. In our research, this operator is integrated into our proposed method to extract more powerful discriminative features. We would like to point out that the CLBP in the proposed CRMCLBP indicates CLBP_S/M/C. When applying the proposed method for each channel of RGB space, three components are generated. They are $\mathrm{CRMLBP}^{\mathrm{Red}}\mathrm{S/M/C}$, $\mathrm{CRMLBP}^{\mathrm{Green}}\mathrm{S/M/C}$ and $\mathrm{CRMLBP}^{\mathrm{Blue}}\mathrm{S/M/C}$ (i.e., one component for each channel). The final descriptor is achieved by concatenating these three components.

The second proposed descriptor, named Prototype Data Model (PDM), provides a compact image descriptor. This low-dimensional representation of color images could be stored in a database which can be used in real-time and large-scale online applications. To construct the PDM descriptor, the three outputs of CRMCLBP (CRMCLBP_S, CRMCLBP_M, CRMCLBP_C) are used as discriminative features for each pixel of a color image. In this method, each pixel is a feature vector with 9 features. The idea is to use a few representative feature vectors (called prototypes) to describe each image in the database. The k-means clustering method is applied to select k-best porotypes from whole pixels which are the good candidates from among all the pixels. We introduce a new similarity measure to compare the content of the images (based on the extracted prototypes).

Many feature weighting algorithms are successfully applied in many distance based learning frameworks in the literature. Feature weighting algorithms improve the performance by controlling the contribution of each feature in the distance function (Moosavi, Jahromi, Ghodratnama, Taheri, & Sadreddini, 2012). To our knowledge, the concept of feature learning is not well studied in LBP based image retrieval methods. Hence, to increase the performance of the proposed methods, we utilize PSO algorithm to generate optimum weights for extracted features.

We assess the performance of our descriptors in terms of image retrieval accuracy and speed. The first proposed method improves the state-of-the-art methods in retrieval performance and has comparable computation time. On the other hand, the second method (PDM) does not provide competitive accuracy but is the fastest method among the state-of-art methods. The number of features in PDM method is drastically less that other methods.

The rest of this paper is organized as follows. In section 2, our proposed methods are explained. Section 3 presents the results of our experiments. Finally, conclusions are remarked in section 4.

2. The proposed method

In this section, first, the Radial Mean Local Binary Pattern (RMLBP) is briefly explained. Then, the proposed color texture descriptors (CRMCLBP and PDM) and their weighted versions are elaborated.

2.1 Radial Mean Local Binary Pattern (RMLBP)

Radial Mean Local Binary Pattern is a robust to noise operator which is proposed by Shakoor and Boostani to extract texture features of gray scale images(Shakoor & Boostani, 2018). In this method, the average of points on each radial is considered as a corresponding neighbor of each center. RMLBP is computed as follows:

$$RMLBP_{P,R,m,d}^{riu2} = \begin{cases} \sum_{n=0}^{P-1} S(\bar{g}_n - g_c) \times 2^n & if\ U(RMLBP_{P,R}) \leq 2 \\ P+1 & Otherwise \end{cases} \tag{1}$$

$$S(\bar{g}_n, g_c) = \begin{cases} 1 & \bar{g}_n \geq g_c \\ 0 & Otherwise \end{cases}$$

$$\bar{g}_n = \sum_{j=-(m-1)/2}^{m-1/2} \bar{g}_n(R + j \times d) \big/ m \tag{2}$$

$$U(RMLBP_{P,R}) = |S(\bar{g}_{P-1} - g_c) - S(\bar{g}_0 - g_c)| + \sum_{n=0}^{P-1} |S(\bar{g}_n - g_c) - S(\bar{g}_{n-1} - g_c)| \tag{3}$$

where *R* and *P* determine the size of neighborhood. The distance between points on the radial is indicated by *d* and the odd value of *m* shows the number of radial points. The average value of them is considered as a neighbor point set of each center pixel. The visualization of this method is shown in fig 1. In this operator, sign and magnitude are used and riu2 mapping are applied to make a histogram of RMLBP code as features. This method can be embedded in any versions of LBP such as CLBP and LTP.

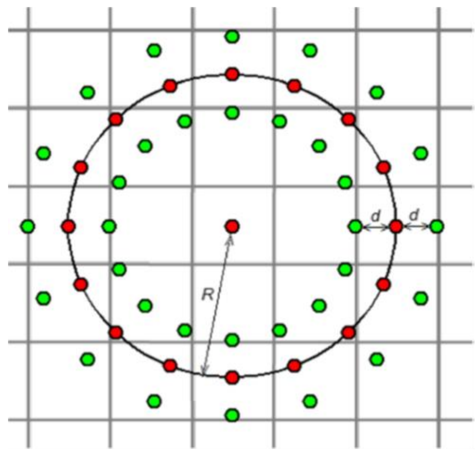

Fig 1. Radial Mean LBP(*R*=2, *P*=16, *d*=1, *m*=5) (Shakoor & Boostani, 2018)

2.2 Color Radial Mean Completed Local Binary pattern (CRMCLBP)

Here, we describe the proposed CRMCLBP (color version of Radial Mean Completed Local Binary Pattern) to extract color texture features of color images. At first, the channels of the color image are separated (R, G, B). Next, the RMLBP operator which has been described in equations (1) to (3) is independently applied on each channel. After that, the histogram of each channel is computed and concatenated together to form the feature vector.

Based on the LBP version, the sign, magnitude and combination of them could be used for each channel. In this paper, we choose Completed Local Binary pattern (CLBP) (Guo et al., 2010) to develop its Color Radial Mean version. CLBP generates six type of feature vectors which should be obtained for each channel. Some of these features are constructed with different combinations of three components: sign difference (RMCLBP_S), magnitude difference (RMCLBP_M) and the threshold of the central gray values of the patterns (RMCLBP_C). In this research, RGB space is chosen as a color space model. After computing three main components by using equations (1) to (3), the histogram of each RMCLBP components are calculated separately as follows:

$$RMCLBP\_SH^{Col} = h((CRMCLBP\_S_{R,P,d,m}^{riu2})^{Col}, num) \tag{4}$$

In above equation, the superscript *Col* is used to denote Red, Green or Blue components of the RGB space model, *h* is histogram function and *num* is maximum LBP pattern whose value depend on the selected mapping (riu2). The final CRMCLBP_SH descriptor is simply obtained as:

$$CRMCLBP\_SH \quad = [RMCLBP\_SH^{Red}, RMCLBP\_SH^{Green}, RMCLBP\_SH^{Blue}] \tag{5}$$

The CRMCLBP_MH feature is calculated similar to equations (1-5):

$$CRMCLBP\_MH \quad = [RMCLBP\_MH^{Red}, RMCLBP\_MH^{Green}, RMCLBP\_MH^{Blue}] \tag{6}$$

The CRMCLBP_C is computed by comparing the center pixel of each pattern with the average gray level of whole image. Equation (7) shows the calculation of this feature.

$$\begin{aligned} &RMCLBP\_C^{Red} = S\left(g_c^{Red}, \bar{g}_{image}\right) \\ &RMCLBP\_C^{Green} = S\left(g_c^{Green}, \bar{g}_{image}\right) \\ &RMCLBP\_C^{Blue} = S\left(g_c^{Blue}, \bar{g}_{image}\right) \\ &CRMCLBP\_C = [RMCLBP\_C^{Red}, RMCLBP\_C^{Green}, RMCLBP\_C^{Blue}] \end{aligned} \tag{7}$$

Similar to equations 5, 6 and 7, other descriptors including *M/CH*, *S_M/CH*, *SMH* and *S/M/CH* could be constructed using histogram of combination of the three RMCLBP operators. For example, the *S/M/CH* descriptor is presented in following equation.

$$\begin{aligned} &CRMCLBP\_S/M/CH \\ &\qquad = [RMCLBP\_S/M/CH^{Red}, RMCLBP\_S/M/CH^{Green}, RMCLBP\_S/M/CH^{Blue}] \end{aligned} \tag{8}$$

2.3 Adaptive Feature Weighting for Color Radial Mean Local Binary pattern (WCRMCLBP)

Consider the image retrieval as a learning problem. In the previous section, we extracted a set of features from available images. To distinguish discriminative from irrelevant features, we can use a feature weighting algorithm. In this paper, Particle Swarm Optimization algorithm (Kennedy & Eberhart, 1995; Shi & Eberhart, 1998) is used to generate the optimum weight for each feature. PSO algorithm has three steps: parameters adjustment, population initialization and search procedure.

In this research, the $CRMCLBP\_S/M/CH$ descriptor of equation 8 is used to extract feature vector from each color texture image as follows:

$$Data(i)_{H\times L} = CRMCLBP\_S/M/C \quad (image_i)\ \forall\ i = 1, 2, \dots\ H. \tag{9}$$

where *H* and *L* shows the number of instances (i.e. image) and the size of the feature vector respectively. The Data matrix is given to the PSO algorithm to find the best and optimum weight for each feature. The performance of the image retrieval system is used as the fitness function and the decision variables are weight vectors for the features. It should be noted that the positon of each particle is a weight vector of same size of LBP feature vector.

The personal best (pbest) values are the best solution for each particle which have been found so far. The global best (gbest) value is the best solution has been achieved by the entire population that is utilized to simulate the communication between population members.

In the first step of the PSO, the parameters are initialized by constriction coefficients theory (Clerc & Kennedy, 2002). These are inertia weight (*w*), personal learning coefficient (c_1) and global learning coefficient (c_2). Other parameters including fitness or cost function, maximum number of iterations, particle definition and size of population and decision variable are also initialized in this step.

At the begin, the positions of particles are randomly initialized by uniform distribution function (U (0,1)). In the search procedure, after calculating the fitness value for each particle, the pbest, gbest, velocity and position values are updated. The update formula for the velocity v and position x variables are expressed as:

$$v_i(t+1) = wv_i(t) + r_1c_1(p_i(t) - x_i(t)) + r_2c_2(g(t) - x_i(t)) \quad (10)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (11)$$

where w is inertia coefficient, r_1 and r_2 are random numbers (U(0,1)). The c_1 and c_2 parameters are the acceleration coefficient. The terms $wv_i(t)$, $r_1c_1(p_i(t) - x_i(t))$ and $r_2c_2(g(t) - x_i(t))$ are known as inertia term, cognitive and social component respectively. These three components are combined together according to equation (10) to create new velocity vector. Then this new velocity translates the current position to the new position in the search space.

This search procedure continues for a certain number of iterations or until an expected condition is reached. At the end of the algorithm, PSO returns the best weight vector for color RMCLBP features. This way, the weighted color descriptor, WCRMCLBP, is obtained.

2.4. Prototype Data Model (PDM)

In this section, the second proposed method is explained in detail. The method, entitled PDM, incorporates the novel idea of selecting a set of few pixels from each image as prototypes which can be used for image classification or retrieval based on Nearest Neighbor rule.

In this method, first, the CRMCLBP operator is applied on each color image for extracting color texture features. This operator has three outputs $CRMCLBP\_S$, $CRMCLBP\_M$ and $CRMCLBP\_C$ with same size as $M \times N \times 3$. It means that, each output has three color components with M-by-N size. Each pixel is considered as an instance including 3×3 features (i.e., 3 outputs $\times$ 3 color component). To do this, the output matrices are reshaped to MN-by-3 matrices whose rows are the pixel vector with three dimensions. Then, the reshaped matrices are concatenated to construct a MN-by-9 matrix, called $CRMCLBP$. The rows and columns are the pixel vectors and LBP features respectively.

Then, to select *k*-best prototypes, the *k*-means algorithm is used. The $CRMCLBP$ matrix is divided into *k* clusters whose centers are selected as *k* prototypes. Therefore, the size of matrix is reduced and change to a k-by-9 matrix. It is noteworthy that the entire image can be representative with a few number of prototypes. The Prototype Data Model (PDM) is summarized in fig 2.

$$[CRMCLBP\_S, CRMCLBP\_M, CRMCLBP\_C] = CRMCLBP(image)$$

$$CRMCLBP\_SR = reshap(CRMCLBP\_S, [M \times N, 3])$$
$$CRMCLBP\_MR = reshap(CRMCLBP\_M, [M \times N, 3])$$
$$CRMCLBP\_CR = reshap(CRMCLBP\_C, [M \times N, 3])$$

$$CRMCLBP = [CRMCLB\_SR\ , CRMCLBP\_MR\ \ , CRMCLBP\_CR\ \ ]$$

$$PDM = Kmeans(CRMCLBP, K)$$

Fig 2. Pseudocode of the Prototype Data Model

This procedure is applied on all images of the dataset and the PDM of each image is extracted. Afterwards, the similarity of each query image and all images in the dataset should be computed. The proposed similarity measure is explained in the next section. Obviously, for improving the performance of the proposed method, the PSO algorithm could be applied on the PDM data to find the best k-by-9 weight matrix. This weighted version is named Weighted Prototype Data Model (WPDM).

2.4.1 Proposed Similarity Measure for comparing images based on PDM features

To compare PDMs, a similarity measure is needed to compare each pair of PDMs. Suppose, two PDM_i and PDM_j are extracted from two color images. For exemplification, assume a two-dimensional feature space (f_1 and f_2) and each PDM_i includes two prototypes PDM_{i1} and PDM_{i2}. As shown in fig 3, employing a distance measure (such as Euclidian, Chi-square, etc.), four distances could be calculated, which are labeled as *a*, *b*, *c* and *d*. For example, the distance between PDM_{i1} and PDM_{j1} determines the value of *a*, and we can use well-known distance measures for this purpose (will be explained in section 4.2). In this research, we introduce min(*a*+*c*, *b*+*d*) as the measure of distance between PDM_i and PDM_j, which is called PdmDist.

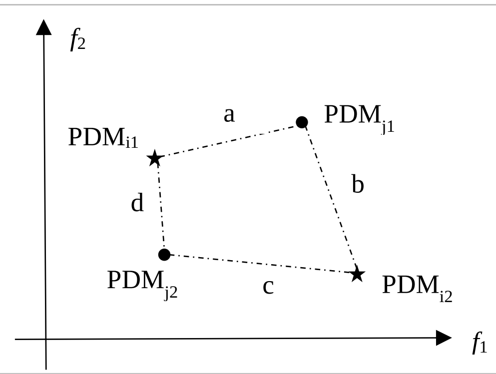

Fig 3. Visualization of the distance measure between two PDMs

3. Result and discussion

This section represents the experimental results of the proposed method. Statistical analysis is performed to compare performance of the proposed method with closely related existing methods. It should be noted that major parameters of these algorithms are adjusted according to the paper presented by Singh et al. (Singh et al., 2018). The performance is measured in term of mean average precision (mAP) and computation time.

3.1 Datasets

In this paper, four datasets Wang, Holidays, Corel- 5K and Corel- 10K are used for analyzing proposed method. These datasets are briefly explained in the following paragraphs.

Wang (J. Z. Wang, Li, & Wiederhold, 2001)**:** This color image dataset has 1000 images in 10 classes. The class labels are African people, beach, building, bus, dinosaur, elephant, flower, horse, glacier, and food. Each class has 100 images with 256×384 or 384×256 pixel resolution.

Holidays (Jegou, Douze, & Schmid, 2008): This dataset includes very high resolution images (2448×3204) with large variety of scene types such as nature, man-made artifact, water effects, fire effects, etc. The dataset contains 1491 images in 500 groups, each of which includes a single query image. The remaining 991 images are used as the training set. The number of images in the groups is variable and their contents are about personal holiday. To reach comparable results for evaluating the proposed method on this dataset, the size of all images are converted to 128×128 by using bicubic interpolation of MATLAB library (Singh et al., 2018).

Corel-5K (G. H. Liu & Yang, 2013; G. H. Liu, Yang, & Li, 2015): 5000 images from diverse contents such as tiger, mountain, mushroom, fort, ocean, car, ticket, etc are collected in this dataset. These images are grouped in 50 categories of 100 images with size of 192×128 or 128×192 in JPEG format

Corel-10K (G. H. Liu & Yang, 2013; G. H. Liu et al., 2015): This dataset consist of 10000 images of various objects such as cat, rose, sunset, duck, train, musical instrument, fish, eagle, judo-karate, etc. These images are grouped in 100 categories, each of them contains 100 JPEG images with size of 192×128 or 128×192.

3.2 Similarity measures and evaluation metrics

Several similarity measures have been proposed in the literature for image retrieval systems. In this paper, four similarity measures are utilized and explained in the following. The performance metrics for evaluating the proposed methods are also presented in this section.

Applying suitable similarity measures is a key factor for in the retrieval systems. In the experiments, we have used four well-known measures, namely: Chi-square, Canberra, Extended-Canberra, and Square-Chord. These measures are commonly used for comparing histogram-based feature vectors. They are suitable for the CRMCLBP method. In the case of second proposed method, PDM, these four measures are embedded in the proposed similarity measure, introduced in section 2.4.1.

Suppose that F_i^q is i^{th} feature of query image q and F_i^o indicates the i^{th} feature of the image o from a database of images. The letter l shows the size of feature vector. The formula of the distance measures are as follows:

Canberra distance: $$Dist_{CD}(q,o) = \sum_{i=0}^{l-1} \left|F_i^q - F_i^o\right| \ / \ (F_i^q + F_i^o) \tag{12}$$

Extended-Canberra distance: $$Dist_{ECD}(q,o) = \sum_{i=0}^{l-1} \left|F_i^q - F_i^o\right| \ / \ ((F_i^q + \overline{F^o}) + (F_i^d + \overline{F^o})) \tag{13}$$

Chi-Square distance: $$Dist_{Chi} = \sum_{i=0}^{l-1} \left(F_i^q - F_i^o\right)^2 \ / \ (F_i^q + F_i^o) \tag{14}$$

Square-Chord distance: $$Dist_{SC} = \sum_{i=0}^{s-1} (\sqrt{F_i^q} - \sqrt{F_i^o})^2 \tag{15}$$

where $\overline{F^q}$ and $\overline{F^o}$ are calculated as bellow:

$$\overline{F^q} = {}^1/_l \sum_{i=0}^{l-1} F_i^q \tag{16}$$

$$\overline{F^o} = {}^1/_l \sum_{i=0}^{l-1} F_i^o \tag{17}$$

Precision and recall are two most well-known performance measures which are commonly utilized for image retrieval systems. For evaluating the methods based on Liu et al. (G. H. Liu & Yang, 2013), first, for each query image, a list of top *N* images retrieved by the method is selected. Then, the number of actually relevant images in this list and in the overall database is determined. The precision *P(N)* and recall *R(N)* are defined as follows:

$$\begin{aligned} P(N) &= {}^{N_r}/_N \\ R(N) &= {}^{N_r}/_M \end{aligned} \tag{18}$$

where N_r indicates the number of relevant images retrieved among top *N* ranked images and *M* is the total number of images in the database that are relevant (i.e., having similar label) to the query image.

The mean of all precision values *P*(*n*) for *n*=1, 2, 3, . . ., *N* is called the average precision of a single query image $\bar{P}(q)$. The formula is shown as bellow:

$$\bar{P}(q) = {}^1/_N \sum_{n=1}^{N} P(n) \tag{19}$$

The mean of the all average precisions for all queries Q is called mean average precision (mAP). It is considered as the main performance measure in our experiments. The mAP is calculated as:

$$mAP = {}^{1}/_{Q} \sum_{q=1}^{Q} \bar{P}(q) \tag{20}$$

The mAP is not the best performance measure for imbalanced datasets. Therefore, we use the bull's eye performance (BEP) (S. Li, Lee, & Pun, 2009) instead of mAP for these type of datasets (for example Holidays dataset). The formula of this measure for a query image *q* is:

$$BEP(q) = {}^{N_q}/_{M} \tag{21}$$

In equation (44), *M* demonstrates the total number of images in the database that are relevant to the query image q. N_q is the number of relevant images among the top *2M* retrievals. The average value of this measure for all query images Q is used to evaluate the image retrieval methods. This metric is calculated as follows:

$$mBEP = {}^{1}/_{Q} \sum_{q=1}^{Q} BEP(q) \tag{22}$$

3.3 Experimental results

The values of parameters for each proposed descriptor are experimentally selected. To evaluate the performance of proposed methods, the results of them are compared with the closely related existing methods.

3.3.1 Parameters setting

The parameters *R*, *P*, *d* and *m* should be adjusted for the CRMCLBP method according to equation 26. The parameters *d* and *m* are set to 1 and 5 respectively. The process of obtaining these values have been explained in (Shakoor & Boostani, 2018). The number of neighbor pixels *P* and neighborhood radius *R* are tested with several values such as (*R*=1, *P*=8), (*R*=1.5, *P*=12) and (*R*=2, *P*=16). Then the pair of (*R*,*P*) providing higher performance with lower number of features is selected. Moreover, we use RMCLBP_S/M/C with rotation invariant uniform pattern (riu2) mapping for developing the proposed descriptors.

For the PDM method an additional parameter, *k*, should be set. It is used in the *k*-means algorithm and indicates the number of prototypes. For obtaining the best value of *k*, the performance of PDM has been tested for *k*=1, *k*=2, *k*=3 and *k*=4 with 10000 iterations and finally the *k*=2 has been selected according to its performance.

The PSO algorithm has several parameters. The value of Inertia weight (*w*), personal (c_1) and global (c_2) learning coefficients are set based on constriction coefficient theory (Clerc & Kennedy, 2002). The parameters are initialized as:

$$\varphi_1 = 2.05, \varphi_2 = 2.05$$
$$\varphi = \varphi_1 + \varphi_2$$
$$Chi = 2 \Big/ \left(\varphi - 2 + \sqrt{\varphi^2 - 4\varphi}\right) \tag{23}$$
$$w = Chi,\ c_1 = Chi\ \times \varphi_1, c_2 = Chi\ \times \varphi_2$$

In the experiments, the population of size 1000 is used for the PSO algorithm and the population members are considered as weight vectors with same size of feature vectors. The weight vector of the CRMCLBP has the size 1×600 (equal to the feature vector), and in the case of the PDM, it is 2×9. As mentioned before, the mAP measure is considered as the fitness function of the PSO algorithm.

3.3.2 Results on the Wang dataset

The retrieval performance of the proposed descriptors with different pair values for (*R*, *P*) are presented in table 1. In this table, the effectiveness of the color version of RMLBP descriptor, similarity measures and feature weighting are shown. In the experiment, four different distance metrics are used and the mAP results are reported before and after applying the feature weighting algorithm.

Obviously, the proposed color texture descriptor outperforms its gray scale version (RMLBP), according to the results that are achieved for all experimental settings (i.e., similarity measures and other parameters). This improvement is also evidenced on Corel-5k, Corel-10k and Holidays datasets in table 2 and table 3.

Without the feature weighting, the top 3 performance for the CRMCLBP (mAP=63.93, 63.77 and 63.08) are related to extended Canberra distance metric with (*R*, *P*) = (2, 16), (1.5, 12) and (1, 8) respectively. For these top 3 results, the length of the feature vectors are 1944, 1176 and 600, respectively. Therefore, the CRMCLBP with mAP 63.08 can be selected as the best method, since it has the lowest feature size.

After feature weighting by means of the PSO algorithm, the mAP increased more than 4 percent. The improvement indicates the effectiveness of applying the feature weighting in our methods. The mAP measure after feature weighting with the feature length of 600 and 1944 are respectively 67.57 and 68.11. It means that, the mAP result is improved slightly by using 3 times more features. Hence the mAP 67.57 using 600 features is selected for comparison of methods in the next section.

According to these experiments, extended Canberra distance is selected as the best distance metric and *R*=1 and *P*=8 are chosen as the acceptable values for neighborhood radius and number of neighbors.

The size of feature vector for the PDM method does not depend on *R* and *P* parameters and is fixed for all experiments but the PDM performance changes with different parameter settings (i.e., different values of *P*, *R*, *k* and different distance measures). The highest mAP value for PDM is 46.56. The result is remarkable regarding small size of the feature vector. The PDM has a low mAP but instead has the lowest number of features in table 1. The mAP of the PDM method is decreased by increasing the value of *P* and *R* parameters.

In the weighted version of the PDM (WPDM), the maximum value of mAP is 49.87. It is obtained by *P*=8, *R*=1 and using the extended Canberra distance.

Table 1. The image retrieval accuracy (mAP) of top 100 images for the proposed methods on Wang dataset (*d*=1, *m*=5 and *k*=2).

(R, P)	Methods	Number of features	Squared Chord	Canberra	Extended Canberra	K-Square
(1,8)	RMLBP (Shakoor & Boostani, 2018)	200	50.63	52.92	53.46	50.62
	CRMCLBP	600	59.16	58.61	63.08	59.62
	WCRMCLBP	600	62.91	61.55	**67.57**	63.10
	PDM	9	44.23	44.79	46.56	44.30
	WPDM	9	47.55	47.88	49.87	47.63
(1.5,12)	RMLBP (Shakoor & Boostani, 2018)	392	50.67	54.41	54.75	50.60
	CRMCLBP	1176	58.43	60.17	63.77	58.98
	WCRMCLBP	1176	61.88	63.37	**67.92**	62.07
	PDM	9	43.40	44.38	45.58	43.39
	WPDM	9	46.77	47.19	48.91	46.61
(2,16)	RMLBP (Shakoor & Boostani, 2018)	648	50.63	54.08	54.92	50.56
	CRMCLBP	1944	58.75	60.63	63.93	59.22
	WCRMCLBP	1944	62.02	63.78	**68.11**	63.05
	PDM	9	43.02	44.05	45.23	43.00
	WPDM	9	46.51	47.13	48.72	46.48

3.3.3 Results on the Corel-5k, Corel-10k and Holidays datasets

The retrieval results (mAP) of the proposed methods on three other datasets, Corel-5k, Corel-10k and Holidays are illustrated in table 2. In the reported experiments, the neighborhood pixels P and radius R are respectively set to 8 and 1 which result in the appropriate feature vector size of 600 (i.e., our objective is to avoid large length feature vectors). Other parameters including d, m and k are the same as the previous experiments (d=1, m=5 and k=2). Again, the proposed methods are tested with four distance metrics in two situations: with and without feature weighting.

Similar to previous examinations, extended Canberra distance provides the best results, the CRMCLBP method outperform the PDM method and the performance of both methods enhance significantly by feature weighting. As seen in table 2, the mAP values for both methods on Corel-10k are lower than the obtained values for Wang and Corel-5k because this dataset has 100 image categories and hence more complex than the other datasets. In spite of this fact, the feature weighting approach improves the performance more than 3 percent.

The number of image in each class in Holidays dataset is variable. Therefore, the mAP is not the best measure for evaluating our methods on this dataset. The average bull's eye performance (mBEP) is used instead of mean average precision (mAP).

3.4 Comparison of the methods based on accuracy

To show the efficiency and performance of our methods, we compared our results with the results of the state-of-the-art methods. The result of the proposed methods, CRMCLBP and PDM, and their weighted versions, WCRMCLBP and WPDM, on four datasets (Wang, Corel-5k, Corel-10k and Holidays) are shown in table 3, along with thirteen outstanding methods.

Table 2. The image retrieval accuracy (mAP) for top 100 images for the proposed methods on Corel-5k dataset for *R*=1, *P*=8, *d*=1 and *m*=5, *k*=2.

Datasets	Methods	Number of features	Squared Chord	Canberra	Extended Canberra	K-Square
Corel-5k	RMLBP (Shakoor & Boostani, 2018)	200	28.50	30.08	31.31	28.33
	CRMCLBP	600	38.43	39.31	42.95	38.90
	WCRMCLBP	600	41.27	43.01	**46.13**	41.33
	PDM	9	21.24	21.59	22.10	21.20
	WPDM	9	24.36	24.56	25.81	24.52
Corel-10k	RMLBP (Shakoor & Boostani, 2018)	200	21.54	22.79	23.93	21.42
	CRMCLBP	600	29.22	30.57	33.86	29.36
	WCRMCLBP	600	32.56	33.07	**37.11**	32.71
	PDM	9	15.99	16.08	16.89	15.88
	WPDM	9	19.23	19.95	20.10	18.98
Holidays	RMLBP (Shakoor & Boostani, 2018)	200	50.03	52.26	53.05	50.13
	CRMCLBP	600	58.23	58.11	61.33	58.67
	WCRMCLBP	600	61.56	60.99	**65.71**	62.91
	PDM	9	43.37	44.58	45.99	44.38
	WPDM	9	47.05	47.29	48.91	47.13

Mean average precision (mAP) and Bull's eye performance (mBEP) are used as image retrieval evaluation metrics (since the number of images in categories of Holidays dataset is variable, the mBEP is used instead of mAP measure).

This measure is computed for top one hundred images, N=100, for each dataset. These results have been achieved by extended Canberra distance metric. We point out the top five methods by labels (*a*) to (*e*). The rank of the methods is exactly the same for all of the datasets.

The highest retrieval performance belongs to the weighted proposed method (WCRMCLBP) with the average value of 54.13. The runner-up method is the combination of LBPC, LBPH and CH with 52.30 average retrieval metric. The third place 50.30 is achieved by our proposed method, CRMCLBP. MDLBP is the next best approach with average accuracy of 49.26. Next best average accuracy is obtained by MSLBP with value of 48.54. The second proposed method, PDM (and its weighted version WPDM) has obtained the acceptable results considering the fact that it uses only 9 features which is drastically lower than the number of features used by the top 5 methods.

It should be noted that, the less the number of features is, the better the computation time of similarity measure would be. For this reason, these descriptors can be good candidates to construct real-time approaches in image processing and computer vision applications such as real-time image retrieval systems. Therefore, the number of features is a key factor for comparing the performance of retrieval systems, so we have to select a method, which has both high accuracy and lower number of features. Features of the runner-up method 542 is the lowest number of features among top 5, which is slightly less than 600 features of the CRMCLBP method. The fourth and fifth best methods, MDLBP and MSLBP have respectively 2048 and 2034 features which are significantly higher than the dimension of the top 3 methods. Our proposed method and second best method (LBPC+LBPH+CH) with lower features provide the highest mAP values. It can be concluded that these methods generate more discriminative features than MDLBP and MSLBP.

PDM and WPDM have lowest number of features among all methods. These methods obtain good results with 9 features and it indicate that this features have high discriminative power to use them in retrieving and compressing images. According to the results of table 3, our proposed method, WCRMCLBP, outperforms the other methods for all datasets.

Table 3. Comparison of various methods in terms of mean average precession (mAP) and bull's eye performance (mBEP)

Method	No. of features	Dataset				
		Wang	Corel-5k	Corel-10k	Holidays	Average
LBP (Timo Ojala et al., 2002)	$3 \times 256 = 768$	56.93	35.75	29.33	61.63	45.91
ULBP (Lowe, 2004)	$3 \times 59 = 177$	54.19	34.26	28.08	59.19	43.93
MSLBP (Maenpaa et al., 2002)	$9 \times 256 = 2304$	**60.62 (*e*)**	**39.95 (*e*)**	**31.69 (*e*)**	**61.92 (*e*)**	**48.54 (*e*)**
LCVBP (S. H. Lee et al., 2012)	$4 \times 59 = 236$	56.83	37.95	29.72	60.17	46.16
RGB-OC-LBP (Zhu et al., 2013)	$3 \times 32 = 96$	49.39	28.12	23.33	54.09	38.73
QLRBP (Lan et al., 2016)	$3 \times 256 = 768$	56.03	36.54	27.64	54.80	43.75
CLSP (J. Li et al., 2016)	$10 \times 256 = 2560$	45.84	29.67	22.53	57.86	38.97
Gabor (Han & Ma, 2007)	96	59.53	36.95	29.68	54.74	45.22
MDLBP (Dubey et al., 2016)	$8 \times 256 = 2048$	**60.82 (*d*)**	**39.99 (*d*)**	**33.79 (*d*)**	**62.46 (*d*)**	**49.26 (*d*)**
LBPC (Singh et al., 2018)	256	58.05	34.08	27.25	60.48	44.96
LBPH (Singh et al., 2018)	256	50.72	28.23	21.98	44.23	36.29
CH (Singh et al., 2018)	30	48.37	25.91	19.18	51.90	36.34
LBPC+LBPH+CH (Singh et al., 2018)	$2 \times 256 + 30 = 542$	**65.16 (*b*)**	**43.81 (*b*)**	**36.99 (*b*)**	**63.25 (*b*)**	**52.30 (*b*)**
RMLBP (Shakoor & Boostani, 2018)	200	53.46	31.31	23.93	53.05	40.44
CRMCLBP	$3 \times 200 = 600$	**63.08 (*c*)**	**42.95 (*c*)**	**33.86 (*c*)**	**61.33 (*c*)**	**50.30 (*c*)**
WCRMCLBP	$3 \times 200 = 600$	**67.57 (*a*)**	**46.13 (*a*)**	**37.11 (*a*)**	**65.71 (*a*)**	**54.13 (*a*)**
PDM	9	46.56	22.10	16.08	45.99	32.68
WPDM	9	49.87	25.81	20.10	48.91	36.17

3.5 Comparison of computation time

For comparing the methods in term of efficiency, total computation time is also considered. In the experiment, total time is computed which includes the time of feature extraction, the time of features matching (computing the distance between query image and images from dataset) and the time of sorting database images based on distance measure. The computation time for various methods on Corel-5k and Corel-10k datasets are shown in table 4. The top 5 fastest methods are shown by numbers from 1 to 5.

The time complexity of distance calculation is $O(l)$, where l is the length of the feature vector. It indicates that the time intensively depends on the length of the feature vector. The time complexity of retrieving algorithm is $O(NM)$, where N is the number of retrieved images from a dataset that includes M images. In table 4, the retrieval time is the sum of times taken by distance metric and sorting algorithm. The feature extraction time depends on the size of image. Since the size of images in two datasets of table 4 is the same, for each method, the feature extraction time is not reported separately for two datasets.

Corel-5k: The lowest time for feature extraction belongs to CH with 30 features. Moreover, it has the lowest time for retrieving images from Corel-5k dataset. The total time for this method is 0.168 seconds which ranks it as the fastest method on this dataset. RGB-OC-LBP has the second best time for feature extraction and image retrieval with 96 features which is the same as the Gabor filtering. Total time for this method is extremely low which makes it the second-placed. The third and fourth fastest methods for feature extraction are LBPC and LBPH sequentially. The size of feature vector for the both methods is 256. The retrieval and total time of two methods dose not place in the five fastest methods. The fifth rank belong to our method, CRMCLBP for feature extraction with 0.044 (s). The proposed approach does not rank in term of retrieval time and total time. PDM with 9 features put in the third place in term of retrieval and total times. In general, CH is the fastest method among the all approaches in this dataset.

Corel-10k: The time of feature extraction is the same for two datasets and the investigation of this step is not necessary. Prototype Data Model (PDM) with 9 features has the highest speed for computing the retrieval step among the other methods. It also acquires the lowest total time with 0.524 (s). In respect of retrieval and total times, CH is ranked second quickest method with 0.550 (s) and 0.564 (s) respectively. RGB-OC-LBP attains the third fastest method in two time factors retrieval and total computation times. The forth rank belongs to ULBP with retrieval time 0.796 (s), total time 0.842 (s) and 177 features. LBPC is placed in fifth rank with total time 1.015 and 256 features. The fifth best method in term of retrieval time is LCVBP with 236 features and time 0.967 (s).

Table 4. Comparison of various methods for the computation time for top 100 images on Corel- 5K and Corel-10K datasets

Methods	# of features	Feature extraction time (s)	Retrieval time (s)		Total computation time (s)		
			Corel-5k	Corel-10k	Corel-5k	Corel-10k	Average of Total time
LBP (Timo Ojala et al., 2002)	768	0.099	0.668	1.885	0.767	1.984	1.375
ULBP (Lowe, 2004)	177	0.046	**0.293 (4)**	**0.796 (4)**	**0.339 (4)**	**0.842 (4)**	**0.590 (4)**
MSLBP (Maenpaa et al., 2002)	2304	0.731	1.757	4.074	2.488	4.805	3.646
LCVBP (S. H. Lee et al., 2012)	236	0.051	**0.312 (5)**	**0.967 (5)**	**0.363 (5)**	1.018	**0.690 (5)**
RGB-OC-LBP (Zhu et al., 2013)	96	**0.027 (2)**	**0.203 (2)**	**0.639 (3)**	**0.230 (2)**	**0.666 (3)**	**0.448 (3)**
QLRBP (Lan et al., 2016)	768	0.092	0.668	1.885	0.760	1.977	1.368
CLSP (J. Li et al., 2016)	2560	0.529	2.273	5.092	2.802	5.621	4.211
Gabor (Han & Ma, 2007)	96	0.765	**0.203 (2)**	**0.639 (3)**	0.968	1.404	1.186
MDLBP (Dubey et al., 2016)	2048	0.212	1.556	2.126	1.768	2.338	2.053
LBPC (Singh et al., 2018)	256	**0.038 (3)**	0.330	0.977	0.368	**1.015 (5)**	0.691
LBPH (Singh et al., 2018)	256	**0.040 (4)**	0.330	0.977	0.370	1.017	0.693
CH (Singh et al., 2018)	30	**0.014 (1)**	**0.154 (1)**	**0.550 (2)**	**0.168 (1)**	**0.564 (2)**	**0.366 (1)**
LBPC+LBPH+CH (Singh et al., 2018)	542	0.092	0.596	1.663	0.688	1.755	1.221
CRMCLBP	600	**0.044 (5)**	0.640	1.703	0.684	1.747	1.193
PDM	9	0.046	**0.241 (3)**	**0.478 (1)**	**0.287 (3)**	**0.524 (1)**	**0.405 (2)**

According to the results of Table 3 and 4, CH and proposed PDM are the fastest methods for retrieving images but their retrieval accuracy (mAP) is relatively low. The top 5 best methods in terms of mAP performance are WCRMCLBP with 600 features, LBPC+LBPH+CH with 542 feature, CRMCLBP with 600 features, MDLBP with 2048 and MSLBP with 2304 features. In this list, WCRMCLPB has the highest average performance with value of 54.13. It also has the acceptable average of total time with 1.193 (s) which is better than the average time of second best method (LBPC+LBPH+CH).

There is a direct relation between the number of features and retrieval time. In fact, the retrieval time for Corel-10k datasets are directly proportional to the number of features. The PDM method has the lowest retrieval time on Corel-10K while its time on Corel-5k is higher than CH, RGB-OC-LBP and Gabor. According to average time, the proposed method WCRMCLBP is the fastest methods among top 5 accurate methods (i.e., highest mAP remarked in table 3). The WCRMCLBP as a promising method provides an admissible tradeoff between retrieval time and accuracy.

4 Conclusion

Extracting efficient features and choosing the suitable similarity measures are two most significant factors for improving performance of a retrieval system. In this paper, two adaptive color descriptors are proposed for color image retrieval problems, named WCRMCLBP and WPDM.

The WCRMCLBP is the color version of RMCLBP, which is constructed by concatenation of three color components of RMCLBP outputs. The CRMCLBP_S/M/C operator is selected as the most powerful combination of CRMCLBP outputs for feature extraction. We use a circular window of size R=1 and p=8 which observed that provides very good retrieval accuracy with appropriate number of features. For making a histogram of features, the rotation invariant uniform pattern mapping (riu2) is employed, because it provides discriminative features. For finding similarity between features, according to the experiments on all datasets, extended-Canberra is selected as the best distance metric.

The WPDM method is based on CRMCLBP and a clustering method such as k-means algorithm for prototype selection. In the WPDM method, a small number of k local binary patterns are selected as representative prototypes, which can be used for image retrieval or classification. The set of prototypes is considered as the best representative patterns for all the pixels of an image. The best value of k is 2 for all datasets. Moreover, a similarity measure, PDMdist, is proposed to compare two sets of prototypes on behalf of comparing two images. It should be noted that, the PDM has a few number of discriminative features. The WPDM with just 9 features has not only an acceptable accuracy but also excellent speed. This method can be improved and utilized in many applications including image classification and image compression.

Both proposed methods are enhanced using an adaptive feature weighting algorithm based on Particle Swarm Optimization. We conclude that the use of feature weighting is very effective in enhancement of the retrieval accuracy.

Detailed experimental analysis for retrieval performance and computation time reveals that the WCRMCLBP method has appropriate speed and the highest retrieval accuracy among well-known and state-of-the-art methods.

Acknowledgement

The authors of this paper express their deepest gratitude to late Dr. Farshad Tajeripour, who paved the road for Mahmood Sotoodeh by unwavering support at the early stage of his Ph.D. thesis. We wish our deep condolence to his family and pray God to leave his soul at a heavenly place.

References

Aggarwal, A., Sharma, S., Singh, K., Singh, H., & Kumar, S. (2019). A new approach for effective retrieval and indexing of medical images. *Biomedical Signal Processing and Control, 50*, 10-34.

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding, 110*(3), 346-359.

Burghouts, G. J., & Geusebroek, J. M. (2009). Performance evaluation of local colour invariants. *Computer vision and image understanding, 113*(1), 48-62.

Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., & Fua, P. (2012). BRIEF: Computing a local binary descriptor very fast. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(7), 1281-1298.

Chatzichristofis, S. A., & Boutalis, Y. S. (2008). *CEDD: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval.* Paper presented at the International Conference on Computer Vision Systems.

Chen, L., Lu, G., & Zhang, D. (2004). *Effects of different Gabor filters parameters on image retrieval by texture.* Paper presented at the Multimedia Modelling Conference, 2004. Proceedings. 10th International, Brisbane, Queensland, Australia, Australia.

Choi, J. Y., Plataniotis, K. N., & Ro, Y. M. (2010). *Using colour local binary pattern features for face recognition.* Paper presented at the Image Processing (ICIP), 2010 17th IEEE International Conference on, Hong Kong, China.

Clerc, M., & Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE transactions on Evolutionary Computation, 6*(1), 58-73.

Dalal, N., & Triggs, B. (2005). *Histograms of oriented gradients for human detection.* Paper presented at the Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, San Diego, CA, USA, USA.

Dubey, S. R., Singh, S. K., & Singh, R. K. (2016). Multichannel decoded local binary patterns for content-based image retrieval. *IEEE Transactions on Image Processing, 25*(9), 4018-4032.

Faugeras, O. D., & Pratt, W. K. (1980). Decorrelation methods of texture feature extraction. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 2*(4), 323-332.

Furht, B. (2008). *Encyclopedia of multimedia* (2th ed.): Springer Science & Business Media.

Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing, 19*(6), 1657-1663.

Hamilton, W. R. (1866). *Elements of quaternions*: Longmans, Green, & Company.

Han, J., & Ma, K. K. (2007). Rotation-invariant and scale-invariant Gabor features for texture image retrieval. *Image and vision computing, 25*(9), 1474-1481.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics, 3*(6), 610-621.

Hiwale, S. S., & Dhotre, D. (2015). *Content-based image retrieval: Concept and current practices.* Paper presented at the Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on, Visakhapatnam, India.

Jegou, H., Douze, M., & Schmid, C. (2008). *Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search*. Paper presented at the Proceedings of the 10th European Conference on Computer Vision: Part I, Marseille, France.

Kennedy, J., & Eberhart, R. (1995). *Particle swarm optimization.* Paper presented at the Neural Networks, 1995. Proceedings., IEEE International Conference on, Perth, WA, Australia, Australia.

Kwitt, R., Meerwald, P., & Uhl, A. (2011). Efficient texture image retrieval using copulas in a Bayesian framework. *IEEE Transactions on Image Processing, 20*(7), 2063-2077.

Lan, R., Zhou, Y., & Tang, Y. Y. (2016). Quaternionic local ranking binary pattern: a local descriptor of color images. *IEEE Transactions on Image Processing, 25*(2), 566-579.

Lee, S. H., Choi, J. Y., Ro, Y. M., & Plataniotis, K. N. (2012). Local color vector binary patterns from multichannel face images for face recognition. *IEEE Transactions on Image Processing, 21*(4), 2347-2353.

Lee, T. S. (1996). Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(10), 959-971.

Li, C., Huang, Y., & Zhu, L. (2017). Color texture image retrieval based on Gaussian copula models of Gabor wavelets. *Pattern recognition, 64*, 118-129.

Li, J., Sang, N., & Gao, C. (2016). Completed local similarity pattern for color image recognition. *Neurocomputing, 182*, 111-117.

Li, S., Lee, M. C., & Pun, C. M. (2009). Complex Zernike moments features for shape-based image retrieval. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 39*(1), 227-237.

Lin, C.-H., Chen, H.-Y., & Wu, Y.-S. (2014). Study of image retrieval and classification based on adaptive features using genetic algorithm feature selection. *Expert Systems with Applications, 41*(15), 6611-6621.

Liu, G. H., & Yang, J. Y. (2013). Content-based image retrieval using color difference histogram. *Pattern recognition, 46*(1), 188-198.

Liu, G. H., Yang, J. Y., & Li, Z. (2015). Content-based image retrieval using computational visual attention model. *Pattern recognition, 48*(8), 2554-2566.

Liu, L., Wang, L., & Liu, X. (2011). *In defense of soft-assignment coding.* Paper presented at the Computer Vision (ICCV), 2011 IEEE International Conference on, Barcelona, Spain.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision, 60*(2), 91-110.

Mäenpää, T., & Pietikäinen, M. (2004). Classification with color and texture: jointly or separately? *Pattern recognition, 37*(8), 1629-1640.

Maenpaa, T., Pietikainen, M., & Viertola, J. (2002). *Separating color and pattern information for color texture discrimination.* Paper presented at the Pattern Recognition, 2002. Proceedings. 16th International Conference on, Quebec City, Quebec, Canada, Canada.

Manjunath, B. S., & Ma, W. Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(8), 837-842.

Manjunath, B. S., Ohm, J. R., Vasudevan, V. V., & Yamada, A. (2001). Color and texture descriptors. *IEEE Transactions on circuits and systems for video technology, 11*(6), 703-715.

Moosavi, M., Jahromi, M. Z., Ghodratnama, S., Taheri, M., & Sadreddini, M. (2012). A Cost sensitive learning method to tune the nearest neighbour for intrusion detection. *Iranian Journal of Science and Technology. Transactions of Electrical Engineering, 36*(E2), 109.

Ojala, T. (1997). Nonparametric texture analysis using simple spatial operators, with applications in visual inspection. *Dissertation, Acta Univ Oul C 105, Department of Electrical Engineering, University of Oulu, Finland., 105*.

Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition, 29*(1), 51-59.

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(7), 971-987.

Park, D. K., Jeon, Y. S., & Won, C. S. (2000). *Efficient use of local edge histogram descriptor.* Paper presented at the Proceedings of the 2000 ACM workshops on Multimedia, Los Angeles, California, USA.

Shakoor, M. H., & Boostani, R. (2018). Radial mean local binary pattern for noisy texture classification. *Multimedia Tools Appl., 77*(16), 21481-21508. doi:10.1007/s11042-017-5440-0

Shi, Y., & Eberhart, R. (1998). *A modified particle swarm optimizer.* Paper presented at the Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, Anchorage, AK, USA, USA.

Singh, C., Walia, E., & Kaur, K. P. (2018). Color texture description with novel local binary patterns for effective image retrieval. *Pattern recognition, 76*, 50-68.

Stricker, M. A., & Orengo, M. (1995). *Similarity of color images.* Paper presented at the Storage and Retrieval for Image and Video Databases III.

Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International journal of computer vision, 7*(1), 11-32.

Tuceryan, M., & Jain, A. K. (1993). Texture analysis *Handbook of pattern recognition and computer vision* (pp. 235-276): World Scientific.

Vassou, S. A., Anagnostopoulos, N., Amanatiadis, A., Christodoulou, K., & Chatzichristofis, S. A. (2017). *CoMo: a compact composite moment-based descriptor for image retrieval.* Paper presented at the Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing.

Wang, J. Z., Li, J., & Wiederhold, G. (2001). SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(9), 947-963.

Wang, X. Y., Yu, Y. J., & Yang, H. Y. (2011). An effective image retrieval scheme using color, texture and shape features. *Computer Standards & Interfaces, 33*(1), 59-68.

Zhu, C., Bichot, C. E., & Chen, L. (2013). Image region description using orthogonal combination of local binary patterns enhanced with color information. *Pattern recognition, 46*(7), 1949-1963.