

# Multichannel Decoded Local Binary Patterns for Content Based Image Retrieval

Shiv Ram Dubey, *Student Member, IEEE*, Satish Kumar Singh, *Senior Member, IEEE*,  
and Rajat Kumar Singh, *Senior Member, IEEE*

**Abstract**— Local binary pattern (LBP) is widely adopted for efficient image feature description and simplicity. To describe the color images, it is required to combine the LBPs from each channel of the image. The traditional way of binary combination is to simply concatenate the LBPs from each channel, but it increases the dimensionality of the pattern. In order to cope with this problem, this paper proposes a novel method for image description with multichannel decoded local binary patterns. We introduce adder and decoder based two schemas for the combination of the LBPs from more than one channel. Image retrieval experiments are performed to observe the effectiveness of the proposed approaches and compared with the existing ways of multichannel techniques. The experiments are performed over twelve benchmark natural scene and color texture image databases such as Corel-1k, MIT-VisTex, USPTex, Colored Brodatz, etc. It is observed that the introduced multichannel adder and decoder based local binary patterns significantly improves the retrieval performance over each database and outperforms the other multichannel based approaches in terms of the average retrieval precision and average retrieval rate.

**Index Terms**— Image retrieval, Local patterns, Multichannel, LBP, Color, Texture.

## 1 INTRODUCTION

IMAGE indexing and retrieval is demanding more and more attention due to its rapid growth in many places. Image retrieval has several applications such as in object recognition, biomedical, agriculture, etc. [1]. The aim of Content Based Image Retrieval (CBIR) is to extract the similar images of a given image from huge databases by matching a given query image with the images of the database. Matching of two images is facilitated by the matching of actually its feature descriptors (i.e. image signatures). It means the performance of any image retrieval system heavily depends upon the image feature descriptors being matched [2]. Color, texture, shape, gradient, etc. are the basic type of features to describe the image [2, 3, 4]. Texture based image feature description is very common in the research community. Recently, local pattern based descriptors have been used for the purpose of image feature description. Local binary pattern (LBP) [5, 6]

has extensively gained the popularity due to its simplicity and effectiveness in several applications [7, 8, 9, 10]. Inspired from the recognition of LBP, several other LBP variants are proposed in the literature [11, 12, 13, 14, 15, 16, 17, 36, 37]. These approaches are introduced basically for gray images, in other words only for one channel and performed well but most of the times in real cases the natural color images are required to be characterize which are having multiple channel.

A performance evaluating of color descriptors such as color SIFT (we have termed mSIFT for color SIFT in this paper), Opponent SIFT, etc. are made for object and scene Recognition in [39]. These descriptors first find the regions in the image using region detectors, then compute the descriptor over each region and finally the descriptor is formed by using bag-of-words (BoW) model. Researchers are also working to upgrade the BoW model [45]. Another interesting descriptor is GIST which is basically a holistic representation of features and has gained wider publicity due its high discriminative ability [40, 41, 42]. In order to encode the region based descriptors into a single descriptor, a vector locally aggregated descriptors (VLAD) has been proposed in the literature [43]. Recently, it is used with deep networks for image retrieval [44]. Fisher kernels are also used with deep learning for the classification [46, 47]. Very recently, a hybrid classification approach is designed by combining the fisher vectors with the neural networks [49]. Some other recent developments are deep convolutional neural networks for imagenet classification [48], super vector coding [50], discriminative sparse neighbor coding [51], fast coding with neighbor-to-neighbor Search [52], projected transfer sparse coding [53] and implicitly transferred codebooks based visual representation [54]. These methods generally better for the classification problem, whereas we designed the descriptors in this paper for image retrieval. Our methods do not require any training information in the descriptor construction process. Still, we compared the results with SIFT and GIST for image retrieval.

A recent trend of CBIR has been efficient search and retrieval for large-scale datasets using hashing and binary coding techniques. Various methods proposed recently for the large scale image hashing for efficient image search such as Multiview Alignment Hashing (MAH) [61], Neighborhood Discriminant Hashing (NDH) [62], Evolutionary Compact Embedding (ECE) [63] and unsupervised bilinear local hashing (UBLH) [64]. These methods can be used with the high discriminative descriptors to improve the efficiency of image search.

Manuscript received November 07, 2015; revised February 13, 2016; accepted May 30, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ling Shao.

The Authors are with the Indian Institute of Information Technology, Allahabad, India (shivram1987@gmail.com, sk.singh@iitaa.ac.in, rajatsingh@iitaa.ac.in).

The final paper DOI: <http://dx.doi.org/10.1109/TIP.2016.2577887>

Copyright (c) 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

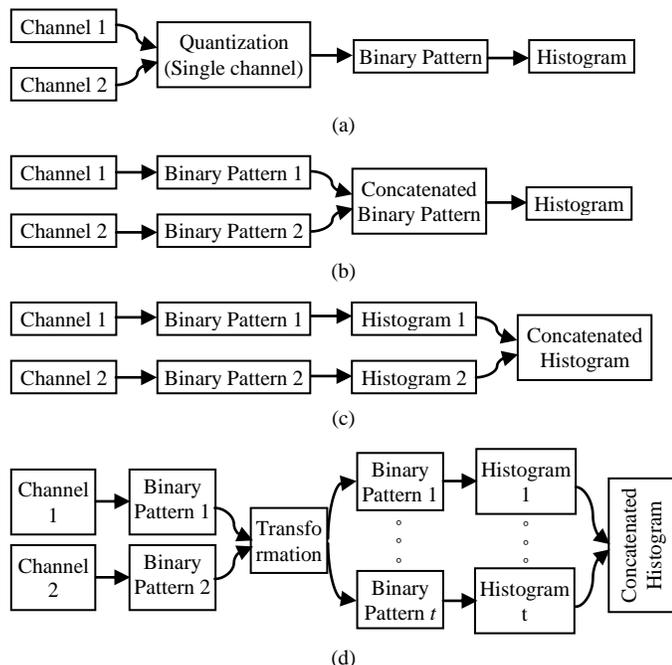


Fig. 1. Illustration of four types of the multichannel feature extraction technique using two input channels, (a) Each channel is quantized and merged to form a single channel and then descriptor is computed over it, (b) Binary patterns extracted over each channel are concatenated to form a single binary pattern and then histogram is computed over it, obviously this mechanism results the high dimensional feature vector, (c) Histograms of binary patterns extracted over each channel are concatenated to form the final feature vector, obviously the mutual information among each is not utilized, and (d) Binary patterns extracted over each channel are converted into other binary patterns using some processing and finally histogram of generated binary patterns are concatenated to form the final feature vector (generalized versions are proposed in this paper).

To describe the color images using local patterns, several researchers adopted the multichannel feature extraction approaches. These techniques can be classified in five categories. The first category as shown in Fig. 1(a) first quantizes each channel then merges each quantized channel to form a single channel and form the feature vector over it. Some typical example of this category is Local Color Occurrence Descriptor (LCOD) [18], Rotation and Scale Invariant Hybrid Descriptor (RSHD) [35], Color Difference Histogram (CDH) [38] and Color CENTRIST [19]. LCOD basically quantized the Red, Green and Blue channels of the image and formed a single image by pooling the quantized images and finally computed the occurrences of each quantized color locally to form the feature descriptor [18]. Similarly, RSHD computed the occurrences of textural patterns [35] and CDH used the color quantization in its construction process [38]. Chu et al. [19] have quantized the H, S and V channels of the HSV color image into 2, 4 and 32 values respectively and represented by 1, 2 and 5 binary bits respectively. They concatenated the 1, 2 and 5 binary bits of quantized H, S and V channels and converted back into the decimal to find the single channel image and finally the features are computed over this image. The major drawback of this category is the loss of information in the process of quantization. The second category simply concatenates the

binary patterns of each channel into the single one as depicted in the Fig. 1(b). The dimension of the final descriptor is very high and not suited for the real time computer vision applications. In the third category (see Fig. 1(c)), the histograms are computed for each channel independently and finally aggregated to form the feature descriptor, for example, [20, 21, 22, 23, 24, 25].

Heng et al. [20] computed the multiple types of LBP patterns over multiple channels of the image such as Cr, Cb, Gray, Low pass and High pass channels and concatenated the histograms of all LBPs to form the single feature descriptor. To reduce the dimension of the feature descriptor, they selected some features from the histograms of LBPs using shrink boost method. Choi et al. [21] computed the LBP histograms over each channel of a YIQ color image and finally concatenated to form the final features. Zhu et al. [22] have extracted the multi-scale LBPs by varying the number of local neighbors and radius of local neighborhood over each channel of the image and concatenated all LBPs to construct the single descriptor. They also concatenated multiple LBPs extracted from each channel of RGB color image [23]. The histograms of multi-scale LBPs are also aggregated in [24] but over each channel of multiple color spaces such as RGB, HSV, YCbCr, etc. To reduce the dimension of the descriptor, Principle Component Analysis is employed in [25]. A local color vector binary pattern is defined by Lee et al. for face recognition [25]. They computed the histogram of color norm pattern (i.e. LBP of color norm values) using Y, I and Q channels as well as the histogram of color angular pattern (i.e. LBP of color angle values) using Y and I channels and finally concatenated these histograms to form the descriptor. The main problem with these approaches is that the discriminative ability is not much improved because these methods have not utilized the inter channel information of the images very efficiently. In order to overcome the drawback of the third category, the fourth category comes into the picture where some of bits of the binary patterns of two channels are transformed and then the rest of the histogram computation and concatenation takes place over the transformed binary patterns as portrayed in the Fig. 1(d). The mCENTRIST [26] is an example of this category where Xiao et al. [26] have used at most two channels at a time for the transformation. In this method, the problem arises when more than two channels are required to model, then the author suggested to apply the same mechanism over each combination of two channels which in turn increases the computational cost of the descriptor.

In furtherance of solving the above mentioned problems of multichannel based feature descriptors, we generalized the 4<sup>th</sup> category of multichannel based descriptors where any number of channels can be used simultaneously for the transformation. In this scheme a transformation function is used to encode the relationship among the local binary patterns of channels. We proposed two new approaches of this category in this paper, where transformation is done on the basis of adder and decoder concepts. The Local Binary Pattern [6] is used in conjunction with our methods as the feature description over each Red, Green and Blue channel of the image. Consider the

case of color LBP, where simply the LBP histogram over each channel are just concatenated, there is no cross-channel co-occurrence information, whereas, if we want to preserve the cross-channel co-occurrence information then the dimension of the final descriptor will be too high. So, in order to capture the cross-channel co-occurrence information to some extent, we proposed the adder and decoder based method with lower dimensions. Moreover, the joint information of each channel is captured in each of the output channels of adder and decoder before the computation of the histogram. We validated the proposed approach against the image retrieval experiments over ten benchmark databases including natural scenes and color textures.

The rest of the paper is organized in following manner; Section II introduces the multichannel decoded Local Binary Patterns; Section III discusses the distance measures and evaluation criteria. Image retrieval experiments using proposed methods are performed in section IV with results discussion; and finally section V concludes the paper.

## 2 MULTICHANNEL DECODED LOCAL BINARY PATTERNS

In this section, we proposed two multichannel decoded local binary pattern approaches namely multichannel adder based local binary pattern (*maLBP*) and multichannel decoder based local binary pattern (*mdLBP*) to utilize the local binary pattern information of multiple channels in efficient manners. Total  $c + 1$  and  $2^c$  number of output channels are generated by using multichannel adder and decoder respectively from  $c$  number of input channels for  $c \geq 2$ .

Let  $I_t$  is the  $t^{\text{th}}$  channel of any image  $I$  of size  $u \times v \times c$ , where  $t \in [1, c]$  and  $c$  is the total number of channels. If the  $N$  neighbors equally-spaced at radius  $\mathcal{R}$  of any pixel  $I_t(x, y)$  for  $x \in [1, u]$  and  $y \in [1, v]$  are defined as  $I_t^n(x, y)$  also depicted in Fig. 2, where  $n \in [1, N]$ . Then, according to the definition of the Local Binary Pattern (LBP) [6], a local binary pattern  $LBP_t^n(x, y)$  for a particular pixel  $(x, y)$  in  $t^{\text{th}}$  channel is generated by computing a binary value  $LBP_t^n(x, y)$  given by the following equation,

$$LBP_t^n(x, y) = \sum_{n=1}^N LBP_t^n(x, y) \times f^n, \quad \forall t \in [1, c] \quad (1)$$

where,

$$LBP_t^n(x, y) = \begin{cases} 1, & I_t^n(x, y) \geq I_t(x, y) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

and  $f^n$  is a weighting function defined by the following equation,

$$f^n = (2)^{(n-1)}, \quad \forall n \in [1, N] \quad (3)$$

We have set of  $N$  binary values  $LBP_t^n(x, y)$  for a particular pixel  $(x, y)$  corresponding to each neighbor  $I_t^n(x, y)$  of  $t^{\text{th}}$  channel. Now we apply the proposed concept of multichannel LBP adder and multichannel LBP decoder by considering  $LBP_t^n(x, y) | \forall t \in [1, c]$  as the  $c$  input channels.

Let, the multichannel adder based local binary patterns  $maLBP_{t_1}^n(x, y)$  and multichannel decoder based local binary patterns  $mdLBP_{t_2}^n(x, y)$  are the outputs of the multichannel LBP adder and multichannel LBP decoder respectively, where  $t_1 \in [1, c + 1]$  and  $t_2 \in [1, 2^c]$ . Note that the values of

$LBP_t^n(x, y)$  are in the binary form (i.e. either 0 or 1). Thus, the values of  $maLBP_{t_1}^n(x, y)$  and  $mdLBP_{t_2}^n(x, y)$  are also in the binary form generated from the multichannel adder map  $maM^n(x, y)$  and multichannel decoder map  $mdM^n(x, y)$  respectively corresponding to the each neighbor  $n$  of pixel  $(x, y)$ .

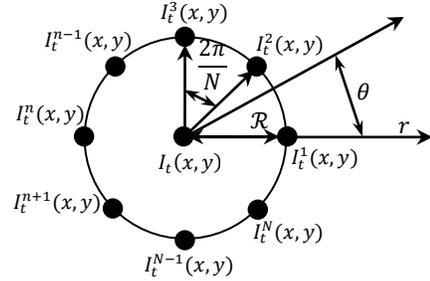


Fig. 2. The local neighbors  $I_t^n(x, y)$  of a center pixel  $I_t(x, y)$  in  $t^{\text{th}}$  channel in polar coordinate system for  $n \in [1, N]$  and  $t \in [1, c]$ .

TABLE I

TRUTH TABLE OF ADDER AND DECODER MAP WITH 3 INPUT CHANNELS

$LBP_1^n(x, y)$	$LBP_2^n(x, y)$	$LBP_3^n(x, y)$	$maM^n(x, y)$	$mdM^n(x, y)$
0	0	0	0	0
0	0	1	1	1
0	1	0	1	2
0	1	1	2	3
1	0	0	1	4
1	0	1	2	5
1	1	0	2	6
1	1	1	3	7

The truth map of  $maM^n(x, y)$  and  $mdM^n(x, y)$  for  $c = 3$  are shown in Table 1 are having 4 and 8 distinct values respectively. Mathematically, the  $maM^n(x, y)$  and  $mdM^n(x, y)$  are defined as,

$$maM^n(x, y) = \sum_{t=1}^c LBP_t^n(x, y) \quad (4)$$

$$mdM^n(x, y) = \sum_{t=1}^c 2^{(c-t)} \times LBP_t^n(x, y) \quad (5)$$

We denote  $LBP_t^n(x, y)$  for  $\forall n \in [1, N]$  and  $\forall t \in [1, c]$  by input patterns,  $maLBP_{t_1}^n(x, y)$  for  $\forall n \in [1, N]$  and  $\forall t_1 \in [1, c + 1]$  by adder patterns and  $mdLBP_{t_2}^n(x, y)$  for  $\forall n \in [1, N]$  and  $\forall t_2 \in [1, 2^c]$  by decoder patterns respectively.

The multichannel adder based local binary pattern  $maLBP_{t_1}^n(x, y)$  for pixel  $(x, y)$  from multichannel adder map  $maM^n(x, y)$  and  $t_1$  is defined as,

$$maLBP_{t_1}^n(x, y) = \begin{cases} 1, & \text{if } maM^n(x, y) = (t_1 - 1) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

for  $\forall t_1 \in [1, c + 1]$  and  $\forall n \in [1, N]$ .

Similarly, the multichannel decoder based local binary pattern  $mdLBP_{t_2}^n(x, y)$  for pixel  $(x, y)$  from multichannel decoder map  $mdM^n(x, y)$  and  $t_2$  can be computed as,

$$mdLBP_{t_2}^n(x, y) = \begin{cases} 1, & \text{if } mdM^n(x, y) = (t_2 - 1) \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

for  $\forall t_2 \in [1, 2^c]$  and  $\forall n \in [1, N]$ .

The computation of  $maLBP_{t_1}^n(x, y)$  for  $\forall t_1 \in [1, c + 1]$  and  $mdLBP_{t_2}^n(x, y)$  for  $\forall t_2 \in [1, 2^c]$  from input  $LBP_t^n(x, y)$  using an example is illustrated in Fig. 3 for  $c = 3$  and  $N = 8$ .

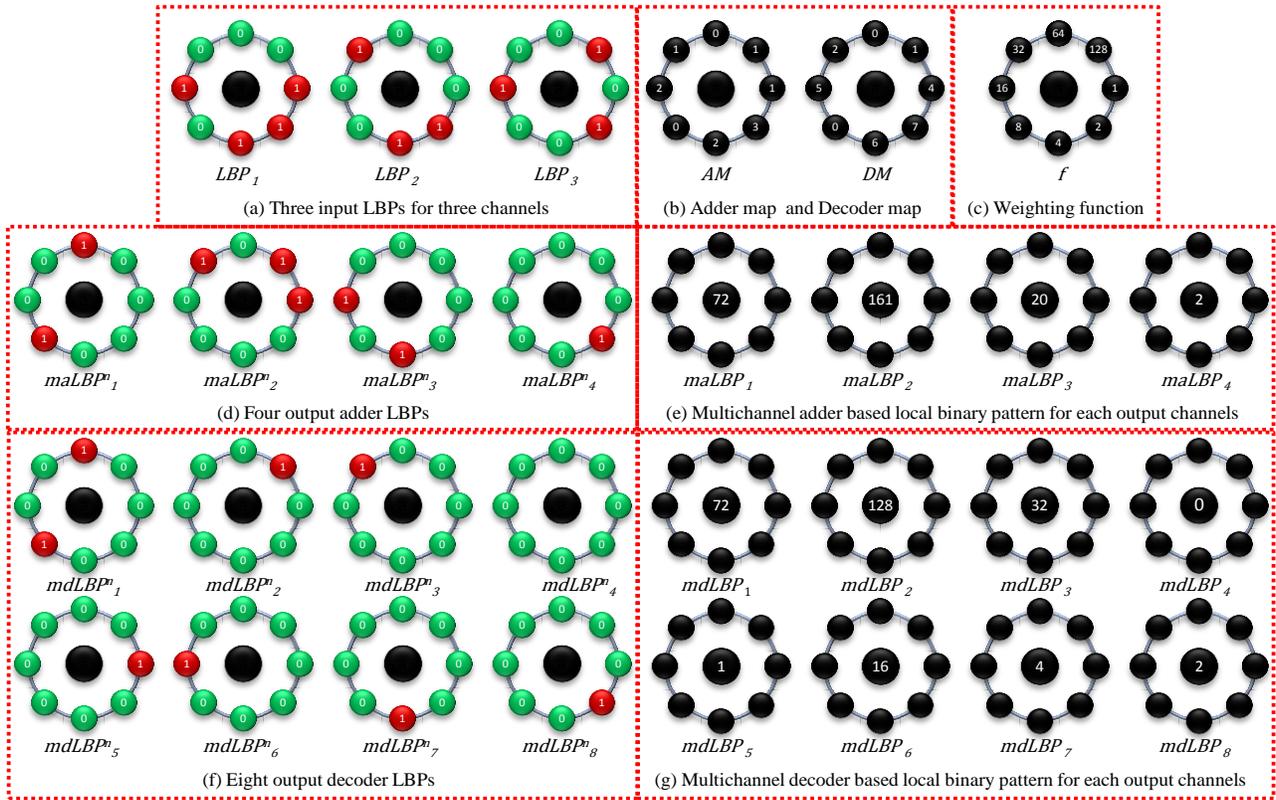


Fig. 3. An illustration of the computation of the adder/decoder binary patterns, and adder/decoder decimal values for  $c = 3$  and  $N = 8$ .

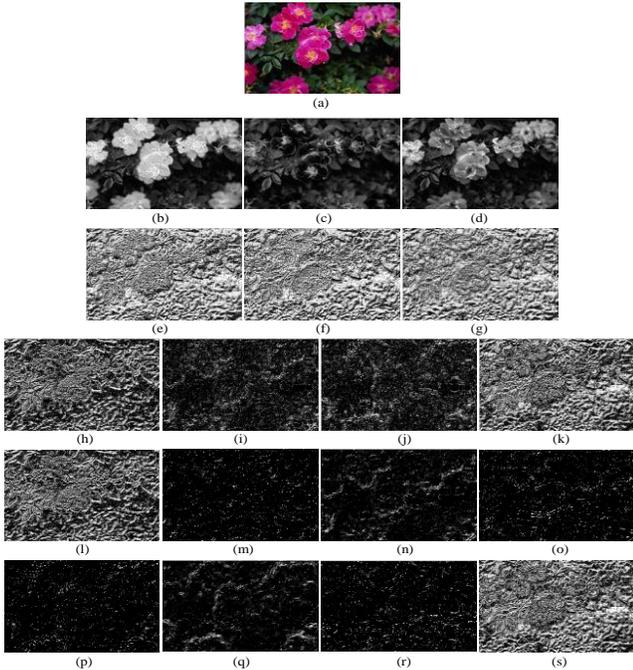


Fig. 4. (a) RGB image, (b) R channel, (c) G channel, (d) B channel, (e) LBP map over R channel, (f) LBP map over G channel, (g) LBP map over B channel, (h-k) 4 output channels of the adder, and (l-s) 8 output channels of the decoder using 3 input LBP maps of R, G and B channels.

The input patterns containing  $LBP_t^n(x, y)$  for  $\forall t \in [1, 3]$  are shown in Fig. 4(a). The multichannel adder map  $maM^n(x, y)$  and multichannel decoder map  $mdM^n(x, y)$  generated using input patterns are depicted in Fig. 3(b). The value of  $c$  is 3 in the example of Fig. 3 so the range of  $t_1$  and  $t_2$  is  $[1, 4]$  and

$[1, 8]$  respectively. The adder patterns containing  $maLBP_{t_1}^n(x, y)$  for  $t_1 \in [1, 4]$  and  $n \in [1, 8]$  generated from  $maM^n(x, y)$  of Fig. 3(b) is depicted in Fig. 3(d) and the decoder patterns containing  $mdLBP_{t_2}^n(x, y)$  for  $t_2 \in [1, 8]$  and  $n \in [1, 8]$  generated from  $mdM^n(x, y)$  of Fig. 3(b) is illustrated in the Fig. 3(f). Note that, the values of  $maLBP_{t_1}^n(x, y)$  will be 1 only if the value of  $maM^n(x, y)$  is  $(t_1 - 1)$  and the values of  $mdLBP_{t_2}^n(x, y)$  will be 1 only if the value of  $mdM^n(x, y)$  is  $(t_2 - 1)$ .

Consider the RGB image, if the value of  $mdLBP_{t_2}^n$  is 1 for  $t_2 = 1$ , it means that  $n^{th}$  neighbor in red, green and blue channels are smaller than center values in respective channels. If the value of  $mdLBP_{t_2}^n$  is 1 for  $t_2 = 4$ , it means that the  $n^{th}$  neighbor in red channel is smaller than the center value in that channel and the  $n^{th}$  neighbor in green and blue channels are greater than the center values in respective channels. In other words,  $mdLBP_{t_2}^n$  represents a unique combination of red, green and blue channels i.e. encoding of the cross channel information.

The multichannel adder based local binary patterns ( $maLBP_{t_1}^n(x, y) | \forall t_1 \in [1, c + 1]$ ) for the center pixel  $(x, y)$  is computed using  $maLBP_{t_1}^n(x, y)$  in the following manner,

$$maLBP_{t_1}(x, y) = \sum_{n=1}^N maLBP_{t_1}^n(x, y) \times f^n \quad (8)$$

and multichannel decoder based local binary patterns ( $mdLBP_{t_2}^n(x, y) | \forall t_2 \in [1, 2^c]$ ) is computed for the center pixel  $(x, y)$  from  $mdLBP_{t_2}^n(x, y)$  using the following equation,

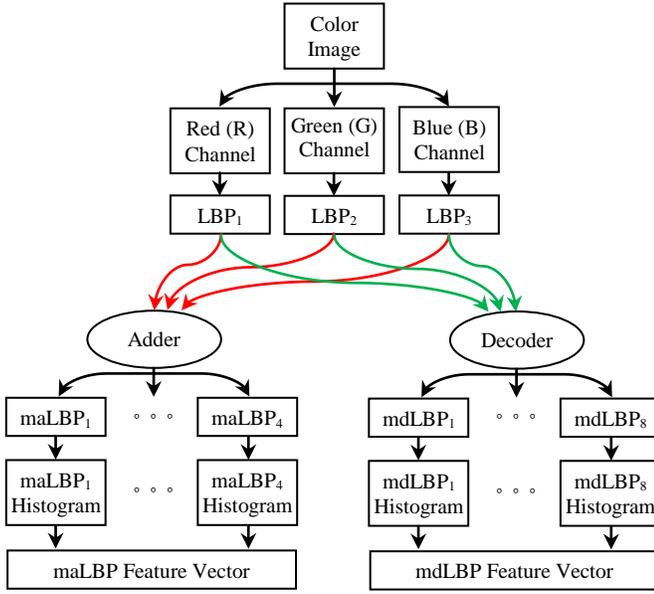


Fig. 5. The flowchart of computation of multichannel adder based local binary pattern feature vector (i.e.  $maLBP$ ) and multichannel decoder based local binary pattern feature vector (i.e.  $mdLBP$ ) of an image from its Red (R), Green (G) and Blue (B) channels.

$$mdLBP_{t_2}(x, y) = \sum_{n=1}^N mdLBP_{t_2}^n(x, y) \times f^n \quad (9)$$

where  $f^n$  is a weighting function defined in (3).

The four multichannel adder local binary patterns (i.e.  $maLBP_{t_1}$  for  $t_1 \in [1,4]$ ) and eight multichannel decoder local binary patterns (i.e.  $mdLBP_{t_2}$  for  $t_2 \in [1,8]$ ) of example binary patterns (i.e.  $LBP_t^n(x, y)$  for  $t \in [1,3]$  and  $n \in [1,8]$ ) of Fig. 3(a) are mentioned in the Fig. 3(e) and Fig. 3(g) respectively. An illustration of the adder output channels and decoder output channels are presented in the Fig. 4 for an example image. An input image in RGB color space (i.e.  $c = 3$ ) is shown in Fig. 4(a). The corresponding Red (R), Green (G) and Blue (B) channels are extracted in the Fig. 4(b-d) respectively. Three LBPs corresponding to the Fig. 4(b-d) are portrayed in the Fig. 4(e-g) for R, G and B channels respectively. The four output channels of the adder and eight output channels of the decoder are displayed in Fig. 4(h-k) and Fig. 4(l-s) respectively. It can be perceived from the Fig. 4 that the decoder channels are having a better texture differentiation as compared to the adder channels and input channels while adder channels are better differentiated than input channels. In other words, we can say that by applying the adder and decoder transformation the inter channel de-correlated information among the adder and decoder channels increases as compared to the same among the input channels.

The feature vector (i.e. histogram) of  $t_1^{th}$  output channel of the adder (i.e.  $maLBP_{t_1}$ ) is computed using the following equation,

$$\mathcal{H}^{maLBP_{t_1}}(\zeta) = \frac{1}{(u-2\mathcal{R})(v-2\mathcal{R})} \sum_{x=\mathcal{R}+1}^{u-\mathcal{R}} \sum_{y=\mathcal{R}+1}^{v-\mathcal{R}} \xi(maLBP_{t_1}(x, y), \zeta) \quad (10)$$

for  $\forall \zeta \in [0, 2^N - 1]$  and  $\forall t_1 \in [1, c + 1]$ , where  $u \times v$  is the dimension of the input image  $I$  (i.e. total number of pixels) and  $\xi(\delta_1, \delta_2)$  is a function and given as follows,

$$\xi(\delta_1, \delta_2) = \begin{cases} 1, & \text{if } \delta_1 = \delta_2 \\ 0, & \text{Else} \end{cases} \quad (11)$$

Similarly, the feature vector of  $t_2^{th}$  output channel of the decoder (i.e.  $mdLBP_{t_2}$ ) is computed as follows,

$$\mathcal{H}^{mdLBP_{t_2}}(\zeta) = \frac{1}{(u-2\mathcal{R})(v-2\mathcal{R})} \sum_{x=\mathcal{R}+1}^{u-\mathcal{R}} \sum_{y=\mathcal{R}+1}^{v-\mathcal{R}} \xi(mdLBP_{t_2}(x, y), \zeta) \quad (12)$$

for  $\forall \zeta \in [0, 2^N - 1]$  and  $\forall t_2 \in [1, 2^c]$ .

The final feature vector of multichannel adder based LBP and multichannel decoder based LBP are given by concatenating the histograms of  $maLBP$ s and  $mdLBP$ s over each output channel respectively and given as,

$$maLBP = \frac{1}{c+1} [\mathcal{H}^{maLBP_1}, \mathcal{H}^{maLBP_2}, \dots, \mathcal{H}^{maLBP_{c+1}}] \quad (13)$$

$$mdLBP = \frac{1}{2^c} [\mathcal{H}^{mdLBP_1}, \mathcal{H}^{mdLBP_2}, \dots, \mathcal{H}^{mdLBP_{2^c}}] \quad (14)$$

The process of computation of  $maLBP$  and  $mdLBP$  feature descriptor of an image is illustrated in Fig. 5 with the help of a schematic diagram. In this diagram, Red, Green and Blue channels of the image are considered as the three input channels. Thus, four and eight output channels are produced by the adder and decoder respectively.

### 3 DISTANCE MEASURES AND EVALUATION CRITERIA

In this section, we discuss the several distance measures and evaluation criteria to confirm the improved performance of proposed feature descriptors for image retrieval.

#### 3.1 Distance measure

The basic aim of distance measures is to find out the similarity between the feature vectors of two images. Six types of distances used in this paper [55-56] are as follows: 1) Euclidean distance, 2)  $L_1$  or Manhattan distance, 3) Canberra distance, 4) Chi-square (Chisq) or  $\chi^2$  distance, 5) Cosine distance, and 6)  $D_1$  distance.

#### 3.2 Evaluation criteria

In content based image retrieval, the main task is to find most similar images of a query image in the whole database. We used each image of any database as a query image and retrieved  $NR$  most similar images. We used *Precision* and *Recall* curves to represent the effectiveness of proposed descriptor. For a particular database, the average retrieval precision (ARP) and average retrieval rate (ARR) are given as follows,

$$ARP = \frac{\sum_{i=1}^C AP(i)}{C} \quad \& \quad ARR = \frac{\sum_{i=1}^C AR(i)}{C} \quad (15)$$

TABLE II  
IMAGE DATABASES SUMMARY

Database Name	Image Size	# Class	# Images in Each Class	# Images (Total)
Corel-1k [27]	384×256	10	100	1000
MIT-VisTex [28]	128×128	40	16	640
STex-512S [29]	128×128	26	Vary	7616
USPTex [30]	128×128	191	12	2292
FTVL [31]	154×116	15	Vary	2612
KTH-TIPS [32]	200×200	10	81	810
KTH-TIPS2a [32]	200×200	11	396 or 432	4608
Corel-TeX [27]	120×80	6	100	600
ALOT [33]	192×128	250	16	4000
ZuBuD [34]	640×480	1001	5	1005
Colored Brodatz [57]	40×40	112	256	28672
ALOT-Complete [33]	192×128	250	100	25000

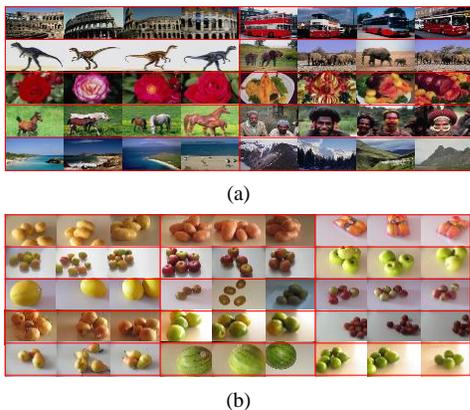


Fig. 6. Example images of the (a) Corel-1k [27], (2) FTVL database [30].

Where  $\mathbb{C}$  is the total number of categories in that database,  $AP$  and  $AR$  are the average precision and average recall respectively for a particular category of that database and defined as follows for  $i^{th}$  category,

$$AP(i) = \frac{\sum_{j=1}^{\mathbb{C}_i} Pr(j)}{\mathbb{C}_i} \quad \& \quad AR(i) = \frac{\sum_{j=1}^{\mathbb{C}_i} Re(j)}{\mathbb{C}_i} \quad \forall i \in [1, \mathbb{C}] \quad (16)$$

Where  $\mathbb{C}_i$  is the number of images in the  $i^{th}$  category of that database,  $Pr$  and  $Re$  are the precision and recall for a query image and defined by the following equation,

$$Pr(j) = \frac{NS}{NR} \quad \& \quad Re(j) = \frac{NS}{ND} \quad \forall j \in [1, \mathbb{C}_i] \quad (17)$$

where  $NS$  is the number of retrieved similar images,  $NR$  is the number of retrieved images, and  $ND$  is the number of similar images in the whole database.

#### 4 EXPERIMENTS AND RESULTS

We conducted extensive CBIR experiments over twelve databases containing the color images of natural scenes, textures, etc. The number of images, number of categories, number of images in each category and image resolutions used in the experiments are mentioned in Table II. The images of a category of a particular database are semantically similar. For example, the Corel-1k database consists of 1000 images from 10 categories namely ‘Buildings’, ‘Buses’, ‘Dinosaurs’,

‘Elephants’, ‘Flowers’, ‘Food’, ‘Horses’, ‘Africans’, ‘Beaches’ and ‘Mountains’ having 100 images each. FTVL database is having 2612 images of fruits and vegetables of 15 types such as ‘Agata Potato’, ‘Asterix Potato’, ‘Cashew’, ‘Diamond Peach’, ‘Fuji Apple’, ‘Granny Smith Apple’, ‘Honeydew Melon’, ‘Kiwi’, ‘Nectarine’, ‘Onion’, ‘Orange’, ‘Plum’, ‘Spanish Pear’, ‘Watermelon’ and ‘Taiti Lime’. Fig. 6(a-b) depicts some images of each category of the Corel-1k and FTVL databases respectively. In the experiments, each image of the database is turned as the query image. For each query image, the system retrieves top matching images from the database on the basis of the shortest similarity score measured using different distances between the query image and database images. If the returned image is from the category of the query image, then we say that the system has appropriately retrieved the target image, else, the system has failed to retrieve the target image.

The performances of different descriptors are investigated using average precision, average recall, ARP and ARR. To demonstrate the effectiveness of the proposed approach, we compared our results of Multichannel Adder and Decoder Local Binary Pattern (i.e. maLBP & mdLBP) with existing methods such as Local Binary Pattern (LBP) [6], Color Local Binary Pattern (cLBP) [21], Multi-Scale Color Local Binary Pattern (mscLBP) [22], and mCENTRIST [26] over each database. We also considered a uniform pattern ( $u2$ ) and rotation invariant uniform pattern ( $riu2$ ) [14] of each descriptor and compared its performances. We have considered  $N = 8$  and  $\mathcal{R} = 1$  in all experiments.

TABLE III  
ARP (%) USING DIFFERENT DISTANCE MEASURES ON COREL-1K DATABASE

Category	ARP (%); (Top matches = 10)					
	Euclidean	L1	Canberra	$\chi^2$	Cosine	D1
LBP	59.41	66.57	<b>68.94</b>	67.08	60.47	67.22
cLBP	61.81	68.93	<b>70.91</b>	69.60	63.22	69.09
mscLBP	65.06	71.15	<b>73.36</b>	71.73	66.30	71.16
mCENTRIST	63.23	69.70	<b>73.61</b>	70.96	64.24	69.81
maLBP	62.12	71.86	70.57	<b>73.43</b>	65.77	72.31
mdLBP	64.55	73.63	68.28	<b>74.93</b>	64.98	74.05

TABLE IV  
ARP (%) USING DIFFERENT DISTANCE MEASURES ON MIT-VISITEX DATABASE

Category	ARP (%); (Top matches = 10)					
	Euclidean	L1	Canberra	$\chi^2$	Cosine	D1
LBP	62.87	68.57	53.28	<b>69.36</b>	63.44	68.95
cLBP	63.59	71.03	61.52	<b>71.94</b>	64.98	71.08
mscLBP	66.97	73.19	71.75	<b>73.52</b>	68.14	73.22
mCENTRIST	66.12	71.17	69.70	<b>72.64</b>	66.52	71.28
maLBP	64.62	73.05	66.63	<b>74.86</b>	67.37	73.66
mdLBP	73.72	79.67	67.25	<b>80.08</b>	73.05	79.89

TABLE V  
ARP (%) USING DIFFERENT DISTANCE MEASURES ON USPTEX DATABASE

Category	ARP (%); (Top matches = 10)					
	Euclidean	L1	Canberra	$\chi^2$	Cosine	D1
LBP	62.95	66.83	50.97	<b>68.44</b>	63.98	66.98
cLBP	68.80	72.27	63.67	<b>73.69</b>	70.29	72.37
mscLBP	71.73	74.66	73.52	<b>75.66</b>	73.48	74.69
mCENTRIST	63.84	68.89	68.33	<b>70.74</b>	64.25	68.97
maLBP	58.41	71.17	61.79	<b>75.67</b>	66.51	72.20
mdLBP	70.82	80.13	64.59	<b>82.50</b>	74.24	80.56

#### 4.1 Experiments with different distance measures

We performed an experiment to investigate that which distance measure is better suited for the proposed scheme. We compared the performance of image retrieval in terms of the average precision rate (ARP) in percentage using Euclidean, L1, Canberra, Chi-square ( $\chi^2$ ), Cosine and D1 distance measures and shown the results in Table III-V over Corel-1k, MIT-VisTex and USPTex databases respectively for 10 numbers of retrieved images. In Table III-V, all distance measures are evaluated for the introduced methods maLBP and mdLBP as well as existing methods LBP, cLBP, mscLBP, and mCENTRIST. Note that three color channels Red (R), Green (G) and Blue (B) are used to find the proposed descriptors for this experiment. The ARP of each LBP, cLBP, mscLBP, and mCENTRIST is better with  $\chi^2$  distance over two databases, whereas it is better over each database for maLBP and mdLBP. From the results of this experiment, it is drawn that  $\chi^2$  distance measure is better suited with the introduced concept of multichannel decoded patterns for image retrieval. It is also noticed that this distance is better suited for remaining descriptors also in most of the cases. In the rest of the results of this paper  $\chi^2$  distance measure will be used to find the dissimilarity between two descriptors.

#### 4.2 Experiments with varying number of input channels

The behavior of proposed multichannel based descriptors is also observed by conducting an experiment with varying number of input channels ( $c$ ). For  $c = 3$ , all three channels Red (R), Green (G) and Blue (B) are used whereas for  $c = 2$ , three combinations are considered; 1) (R, G), 2) (R, B) and 3) (G, B). The image retrieval experiments are performed over Corel-1k, MIT-VisTex, Corel-Tex and ALOT databases as depicted in Fig. 7(a-d) respectively. In this experiment, 10 images are retrieved using maLBP and mdLBP with (R, G), (R, B), (G, B) and (R, G, B) input channels and ARP (%) is computed to demonstrate the performance. For  $c = 2$ , the ARP of maLBP is generally better than the ARP of mdLBP except over ALOT database. Whereas, for  $c = 3$ , the ARP of mdLBP is far better than the ARP of maLBP. One possible explanation of this behavior is that in case of decoder the  $c$  inputs are being transformed into  $2^c$  outputs (i.e. no information loss), whereas, in case of adder the  $c$  inputs are being transformed into  $c + 1$  outputs (i.e. some information is being lost). Moreover, the information loss becomes more for higher values of  $c$  (i.e. the loss is more for  $c = 3$  as compared to  $c = 2$  because the number of outputs in {adder, decoder} is {3, 4} and {4, 8} for  $c = 2$  and 3 respectively). One more important assertion observed in this experiment is that the performance of both maLBP and mdLBP is better for  $c = 3$  (i.e. using three channels) as compared to the  $c = 2$  (i.e. using only two channels) except the case of maLBP over ALOT database. In order to know the reason for this, we computed the average standard deviation (ASD) for all the images of each database over each channel. The ASD over {Red, Green, Blue} channels is {59.44, 54.87, 55.45}, {52.26, 48.25,

43.83}, {55.03, 53.78, 51.23}, and {36.98, 35.40, 34.89} for Corel-1k, MIT-VisTex, Corel-Tex, and ALOT database respectively. It can be observed that the ASD for ALOT database is very low which means that the intensity values are very close to each other. This is the fact that maLBP for  $c = 3$  is not performing better than maLBP for  $c = 2$  over ALOT database because the actual combinations of input are less. It is also pointed out that the degree of performance improvement for  $c = 3$  is much better for mdLBP as compared to maLBP descriptor. From this experiment, it is determined that all three channels play a crucial role and by removing any one channel the performance degrades drastically.

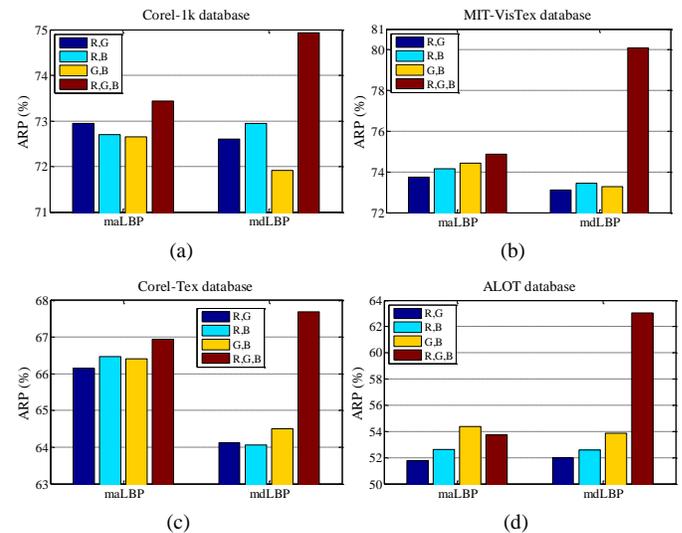


Fig. 7. ARP (%) for different combinations of channels of RGB color space using maLBP and mdLBP descriptors over (a) Corel-1k, (b) MIT-VisTex, (c) Corel-Tex, and (d) ALOT database when 10 similar images are retrieved.

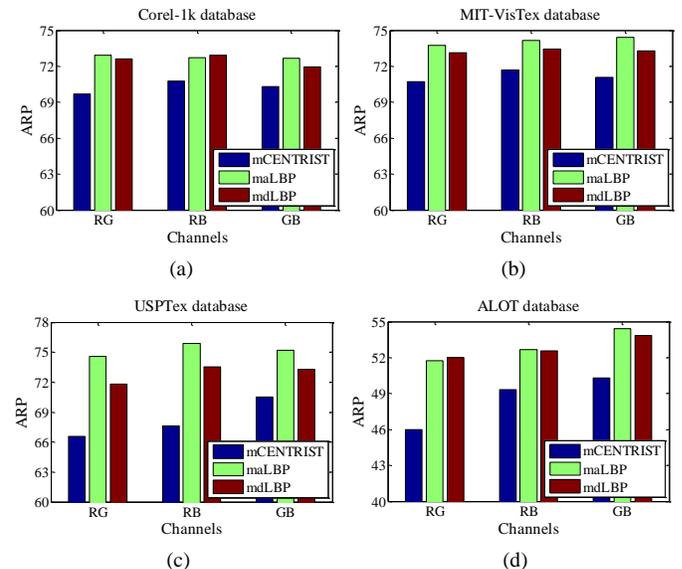


Fig. 8. ARP (%) for different combinations of two channels of RGB color space using mCENTRIST, maLBP and mdLBP descriptors over (a) Corel-1k, (b) MIT-VisTex, (c) USPTex, and (d) ALOT database when 10 similar images are retrieved.

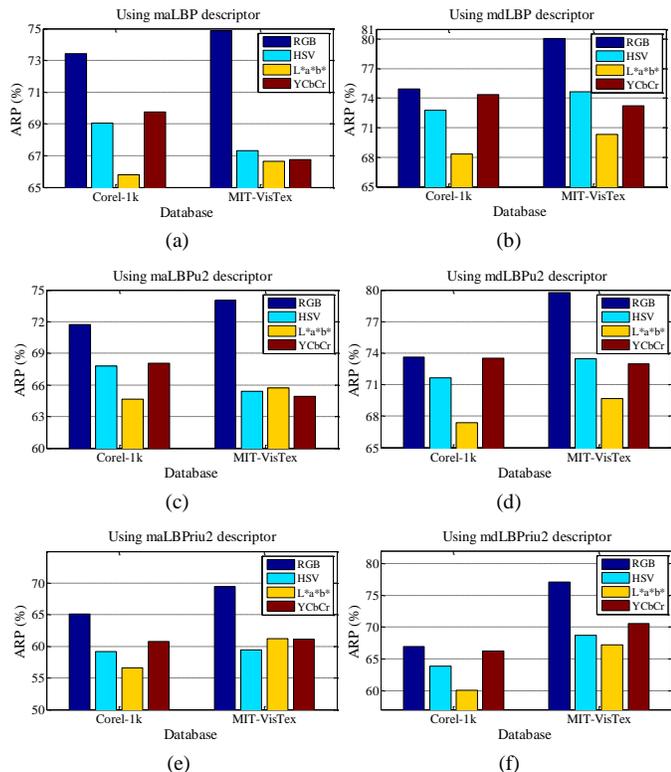


Fig. 9. The comparison of performance among RGB, HSV,  $L^*a^*b^*$  and YCbCr color spaces in terms of the ARP (%) using (a) maLBP, (b) mdLBP, (c) maLBPu2, (d) mdLBPu2, (e) maLBPriu2, and (f) mdLBPriu2 descriptors over Corel-1k and MIT-VisTex databases when 10 numbers of top similar images are retrieved.

In the rest of the paper, the value of  $c$  will be considered as 3 until or otherwise stated. We also compared the ARP obtained using maLBP and mdLBP with mCENTRIST over four databases for only combinations of two channels in Fig. 8 and found that the performance of proposed descriptors are always better as compared to the mCENTRIST for  $c = 2$ .

### 4.3 Experiments with different color spaces

In order to find out the preferable color space for our schemes, we performed the image retrieval experiments over Corel-1k and MIT-VisTex databases in four different color spaces namely RGB, HSV,  $L^*a^*b^*$  and YCbCr. The results of these experiments are represented in terms of the ARP in Fig. 9(a-f) using maLBP, mdLBP, maLBPu2, mdLBPu2, maLBPriu2, and mdLBPriu2 descriptors respectively for 10 numbers of retrieved images. The performance of each descriptor in each database is better in RGB color space because the channels of RGB color space are highly correlated as compared to the HSV,  $L^*a^*b^*$  and YCbCr color spaces. The performance of each descriptor is poor in  $L^*a^*b^*$  color space over natural scene database (i.e. Corel-1k) and also in most of the experiments over color texture database (i.e. MIT-VisTex). The ARP of each descriptor in YCbCr color space is better as compared to the same in HSV color space over Corel-1k database. Whereas, it is reversed over MIT-VisTex database except using mdLBPriu2 descriptor. In the rest of the paper, the RGB color space will be used until or otherwise stated.

### 4.4 Comparison with existing approaches

We have performed extensive image retrieval experiments over ten databases of varying number of categories as well as varying number of images per category to report the improved performance of proposed multichannel decoded local binary patterns. We have reported the results using average retrieval precision (ARP), average retrieval rate (ARR), average precision per category (AP) and average recall per category (AR) as the function of number of retrieved images (NR). Fig. 10 shows the ARP vs NR plots for LBP, cLBP, mscLBP, mCENTRIST, maLBP and mdLBP descriptors over Corel-1k, MIT-VisTex, STeX, USPTex, FTVL, KTH-TIPS, KTH-TIPS2a, and Corel-1k databases. The ARP values for each  $NR \in [1, 10]$  using decoder based mdLBP descriptor are higher than other descriptors over each database (see Fig. 10). Moreover, the performance of mdLBP is much better over MIT-VisTex, STeX-512S, USPTex, and FTVL databases. The performance of adder based maLBP descriptor is also better than other descriptors such as LBP, cLBP, mscLBP and mCENTRIST over Corel-1k, MIT-VisTex, and USPTex databases as depicted in Fig. 10. The performance of each descriptor over each database is also compared using uniform (i.e.  $u2$ ) and rotation invariant uniform (i.e.  $riu2$ ) patterns in Fig. 11-12 respectively using ARP vs NR curves. It can be observed that the ARP values using mdLBPu2 and mdLBPriu2 are far better than the ARP values using remaining descriptors under both  $u2$  and  $riu2$  transformation (see Fig. 11-12). As far as adder based maLBP is concerned, its performance is more improved under  $u2$  as compared to the  $riu2$ . The performance of mCENTRIST is drastically degraded under  $u2$  and  $riu2$  conditions. One possible explanation for this behavior of mCENTRIST is that it is computed by interchanging the half of the binary pattern of two channels which causes the loss in rotation invariance property. While other hand, the performance of mscLBP is improved under  $u2$  and  $riu2$  conditions which exhibit its rotation invariant property. The image retrieval experimental results over remaining databases are demonstrated in terms of the ARR vs NR plots using each descriptor in Fig. 13. The performance of each descriptor is compared over ALOT and ZuBuD databases without any transformation (see Fig. 13(a-b)), with  $u2$  transformation (see Fig. 13(c-d)) and with  $riu2$  transformation (see Fig. 13(e-f)). The ARR values using proposed maLBP are higher than existing approaches in most of the results of Fig. 13, while the ARR values using proposed mdLBP are higher than other approaches in each result of the Fig. 13, moreover, the degree of improvement in the performance of mdLBP is higher over the ALOT database as compared to the ZuBuD database. We also explored the categorical performance of the descriptors over Corel-1k, MIT-VisTex, FTVL and STeX-512S databases in Fig. 14. Average precision is measured over Corel-1k (using descriptors without transformation) and MIT-VisTex (using descriptors with  $u2$  transformation) databases, while, average recall is measured over FTVL (using descriptors without transformation) and STeX-512S (using descriptors with  $riu2$  transformation) databases.

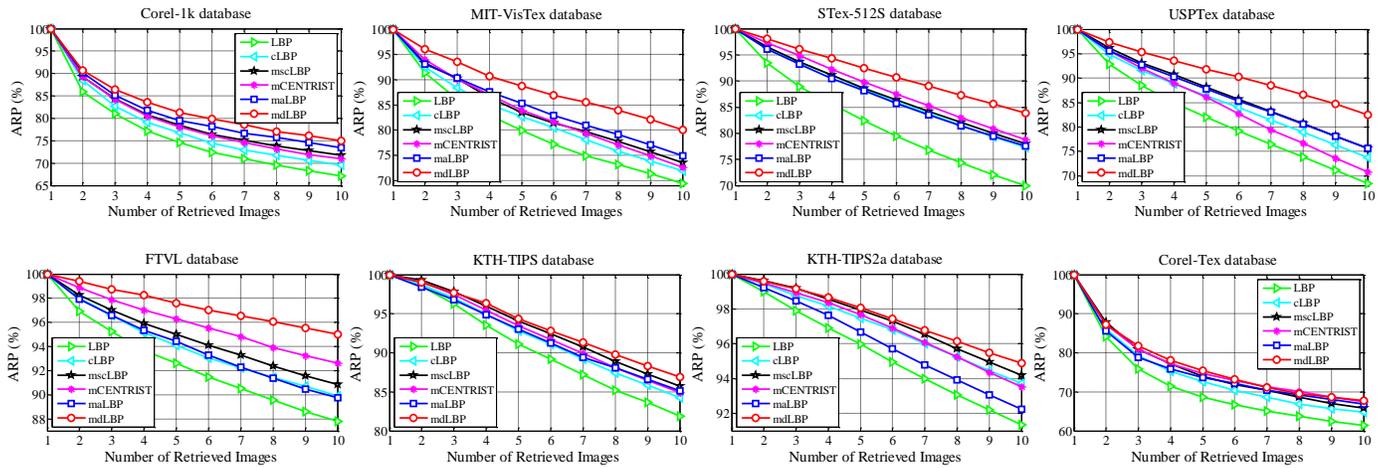


Fig. 10. The performance comparison of proposed maLBP and mdLBP descriptor with existing approaches such as LBP, cLBP, mscLBP, and mCENTRIST descriptors using ARP vs number of retrieved images over Corel-1k, MIT-VisTex, STex, USPTex, FTVL, KTH-TIPS, KTH-TIPS2a, and Corel-Tex databases.

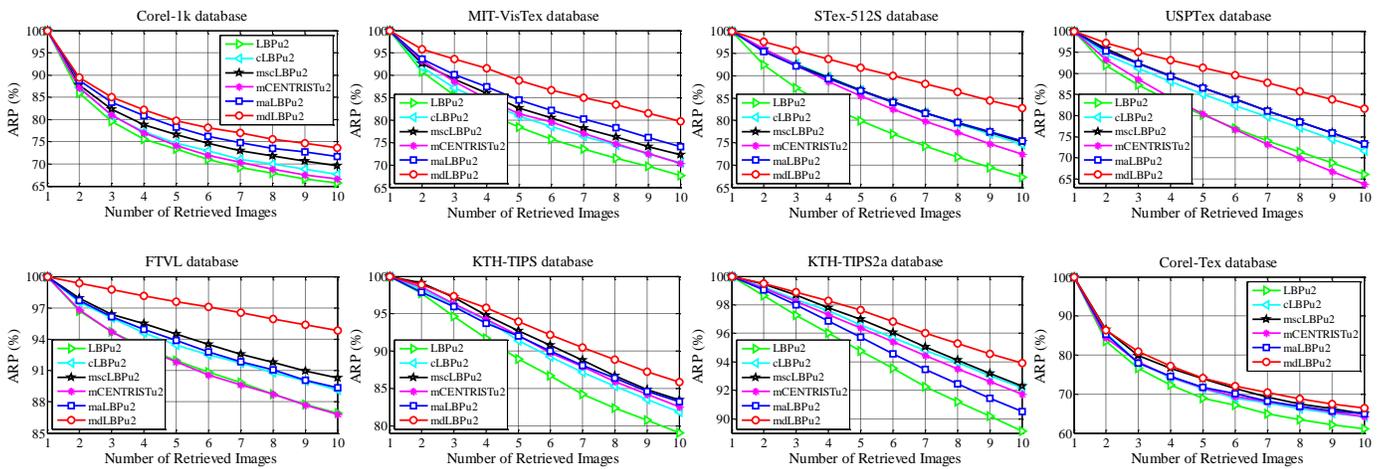


Fig. 11. The performance comparison of proposed maLBP and mdLBP descriptor with existing approaches such as LBP, cLBP, mscLBP, and mCENTRIST descriptors under uniform transformation (u2) using ARP vs number of retrieved images plot over Corel-1k, MIT-VisTex, STex, USPTex, FTVL, KTH-TIPS, KTH-TIPS2a, and Corel-Tex databases.

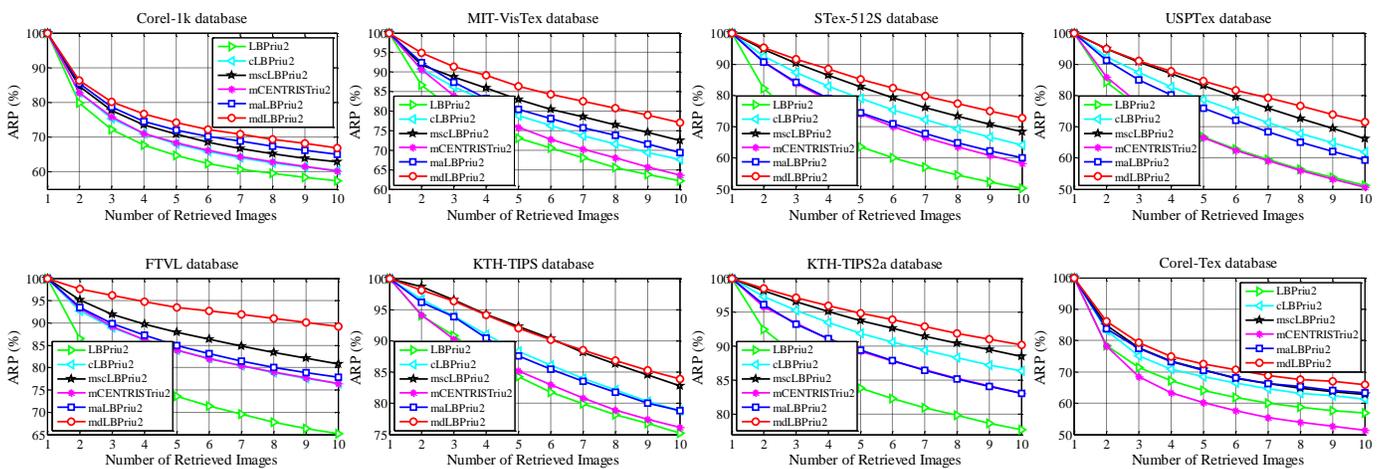


Fig. 12. The performance comparison of proposed multichannel decoded local binary patterns with existing approaches under rotation invariant uniform transformation (riu2) using ARP vs number of retrieved images curve over Corel-1k, MIT-VisTex, STex, USPTex, FTVL, KTH-TIPS, KTH-TIPS2a, and Corel-Tex databases.

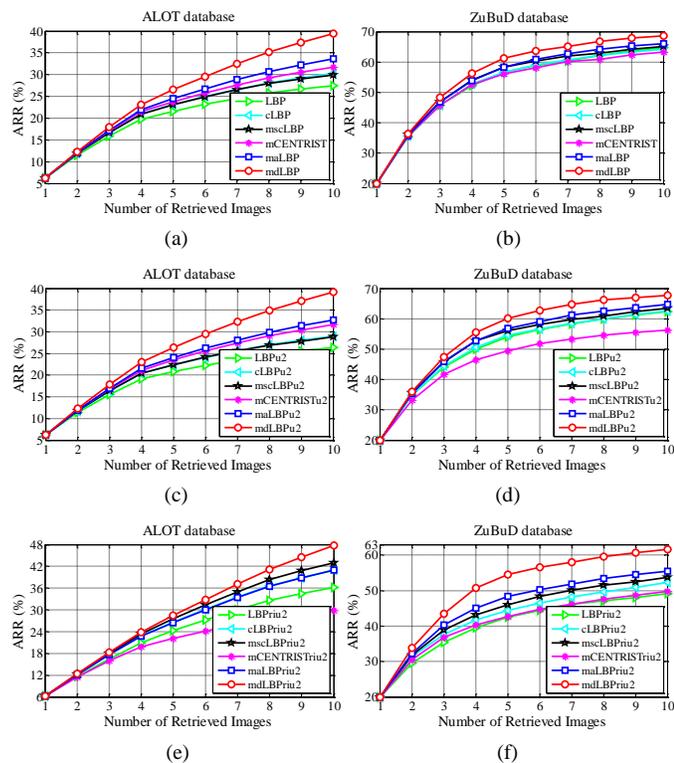


Fig. 13. The performance evaluation of proposed methods under (a-b) without transformation, (c-d)  $u2$  transformation, and (e-f)  $riu2$  transformation in terms of the ARR vs number of retrieved images over ALOT and ZuBuD databases.

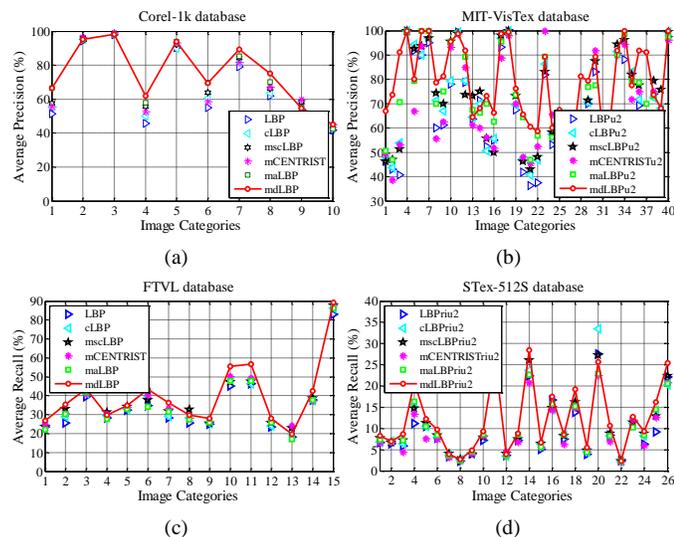


Fig. 14. Experimental results using different descriptors in terms of the average precision (%) over (a) Corel-1k, and (b) MIT-VisTex, and average recall (%) over (c) FTVL, and (d) STex-512S databases for each category of the database.

It is noticed across the plots of the Fig. 14 that the performance of mdLBP is better and consistent in most of the categories of each database while at the other end the performance of maLBP is also better in most of the categories as compared to the existing multichannel based approaches. We have drawn the following assertions from Fig. 10-14:

1. The proposed decoder based mdLBP descriptor outperformed the other multichannel based descriptors

in terms of the ARP over Corel-1k, MIT-VisTex, STex, USPTex, FTVL, KTH-TIPS, KTH-TIPS2a, and Corel-1k databases.

2. The mdLBP descriptor outperformed the other multichannel based descriptors in terms of the ARR also over ALOT and ZuBuD databases.
3. The mdLBPu2 and mdLBPriu2 are also outperformed the remaining descriptors under  $u2$  and  $riu2$  conditions respectively.
4. The performance of proposed maLBP descriptor is not as improved as it is improved by mdLBP descriptor in terms of the ARP and ARR values.
5. The categorical performance of the proposed methods is also better as compared to the existing methods including  $u2$  and  $riu2$  scenarios.

The top 10 similar retrieved images for the 9 query images using each descriptor from the Corel-1k database is displayed in Fig. 15. The Corel-1k database is having 10 categories with 100 images in each category. Note that the rows of each subfigure of Fig. 15 represent the retrieved images using different descriptors in following manner: 1<sup>st</sup> row using LBP, 2<sup>nd</sup> row using cLBP, 3<sup>rd</sup> row using mscLBP, 4<sup>th</sup> row using mCENTRIST, 5<sup>th</sup> row using maLBP and 6<sup>th</sup> row using mdLBP; the 10 columns in each subfigure corresponds to the 10 retrieved images in decreasing order of similarity (i.e. the images in the 1<sup>st</sup> column are the top most similar images while it is the query images also). Fig. 15(a) has shown the retrieved images for a query image from the ‘Building’ category. The {precision, recall} obtained by LBP, cLBP, mscLBP, mCENTRIST, maLBP and mdLBP for this example are {30%, 3%}, {30%, 3%}, {30%, 3%}, {40%, 4%}, {60%, 6%}, and {60%, 6%} respectively. A query image from the ‘Bus’ category is considered in the Fig. 15(b). In this example, only proposed maLBP and mdLBP descriptors are able to retrieve the all images from the ‘Bus’ category (i.e. 100% precision). All the descriptors have gained 100% precision for an example query image from ‘Dinosaurs’ category as depicted in the Fig. 15(c). Whereas, the retrieved images using mscLBP and mdLBP are more semantically similar with the query image because the orientation of the ‘Dinosaur’ is same in 3<sup>rd</sup> and 6<sup>th</sup> rows of the Fig. 15(c). The precision achieved using LBP, cLBP, mscLBP, mCENTRIST, maLBP and mdLBP descriptors for an example query image of ‘Elephant’ are 60%, 60%, 70%, 60%, 60%, and 90% (see the Fig. 15(d)). In Fig. 15(e), more semantic images are retrieved by the mdLBP for a query image from ‘Flower’ as all the retrieved are having similar appearance as illustrated in the 6<sup>th</sup> row of the Fig. 15(e). The retrieval precision for a query image of type ‘Food’ is very high for the proposed approaches, whereas, it is very low for the existing approaches as portrayed in the Fig. 15(f). The number of correct images retrieved using LBP, cLBP, mscLBP, mCENTRIST, maLBP and mdLBP descriptors for a query image from ‘Horse’ category are 7, 7, 6, 6, 9, and 10 (see the Fig. 15(g)). The retrieval precision gained by proposed descriptors are also high as compared to the existing descriptors for the query images from categories ‘Africans’ and ‘Beaches’ as demonstrated in the Fig. 15(h-i) respectively.

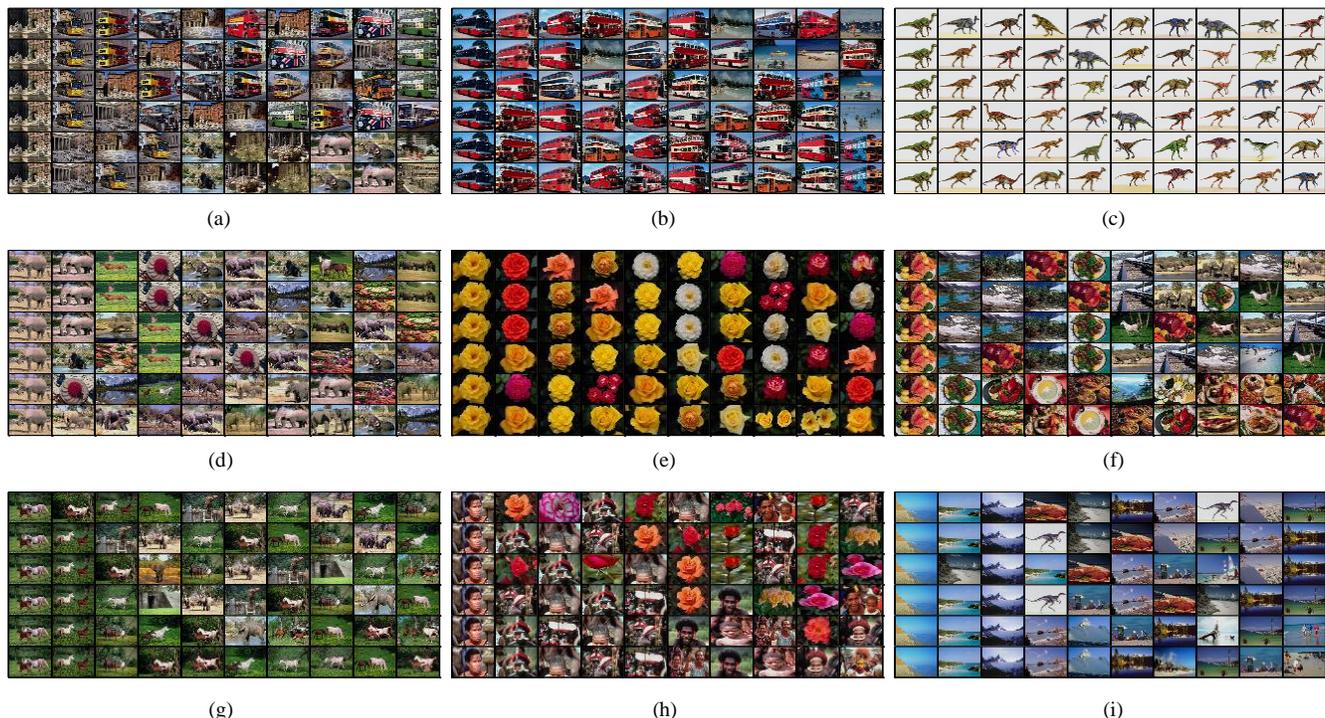


Fig. 15. Top 10 retrieved images using each descriptor from Corel-1k database by considering the query image from (a) ‘Building’, (b) ‘Bus’, (c) ‘Dinosaurs’, (d) ‘Elephant’, (e) ‘Flower’, (f) ‘Food’, (g) ‘Horse’, (h) ‘Africans’, and (i) ‘Beaches’ categories. Note that 6 rows in each subfigure corresponds to the different descriptor such as LBP (1<sup>st</sup> row), cLBP (2<sup>nd</sup> row), mscLBP (3<sup>rd</sup> row), mCENTRIST (4<sup>th</sup> row), maLBP (5<sup>th</sup> row) and mdLBP (6<sup>th</sup> row) and 10 columns in each subfigure corresponds to the 10 retrieved images in decreasing similarity order; images in the 1<sup>st</sup> column are query images as well as the top most similar images.

TABLE VI  
FEATURE DIMENSIONS VS TIME COMPLEXITY IN SECONDS

Method	Feature Dimension	Feature Extraction Time (Corel-1k database)	Total Retrieval Time (Corel-1k database)	Feature Extraction Time (MIT-VisTex database)	Total Retrieval Time (MIT-VisTex database)
LBP	256	31.00	2.73	4.67	1.19
cLBP	3×256	89.52	11.75	6.44	3.38
mscLBP	9×256	310.95	56.40	13.03	10.07
mCENTRIST	6×256	128.01	33.31	9.72	6.7
maLBP	4×256	102.45	16.49	9.24	4.39
mdLBP	8×256	228.97	47.93	12.48	7.8
LBPu2	59	32.00	0.74	5.11	0.41
cLBPu2	3×59	88.78	1.85	7.16	0.88
mscLBPu2	9×59	316.34	5.50	19.50	2.43
mCENTRISTu2	6×59	132.20	3.72	10.46	1.64
maLBPu2	4×59	105.49	2.41	9.41	1.11
mdLBPu2	8×59	233.93	4.71	13.63	2.03
LBPrui2	10	31.73	0.28	5.27	0.23
cLBPrui2	3×10	88.64	0.46	6.95	0.31
mscLBPrui2	9×10	316.11	1.02	19.13	0.53
mCENTRISTriu2	6×10	130.94	0.74	10.08	0.42
maLBPrui2	4×10	105.09	0.55	9.32	0.35
mdLBPrui2	8×10	233.10	0.91	13.23	0.48

It is deduced from the retrieval results that the precision and recall using proposed multichannel based maLBP and mdLBP descriptor is high as compared to the same using LBP and existing multichannel based approaches such as cLBP, mscLBP and mCENTRIST descriptors. It is also observed that

the performance of mdLBP is better than the maLBP. It is shown by the experiments that proposed mdLBP method outperforms other methods because mdLBP encodes each combination of the red, green and blue channels locally from its LBP binary values. The color in images is depicted by three values but most of methods process these values separately which loss the cross channel information. Whereas, mdLBP takes all the combinations of LBP binary value computed over each channel using a decoder based methodology.

#### 4.5 Analysis over feature extraction and retrieval time

We analyzed the feature extraction time as well as the retrieval time for each descriptor over Corel-1k and MIT-VisTex databases. Both the feature extraction and retrieval time are computed in seconds using a personal computer having Intel(R) Core(TM) i5 CPU 650@3.20 GHz processor, 4 GB RAM, and 32-bit Windows 7 Ultimate operating system with 4-cores active. The feature dimension of each descriptor is mentioned in the Table VI with feature extraction and retrieval times over Corel-1k and MIT-VisTex databases. The feature extraction time of mdLBP is {2.55, 1.93}, {0.73, 0.96} and {1.78, 1.28} times slower than the feature extraction time of cLBP, mscLBP and mCENTRIST respectively over {Corel-1k, MIT-VisTex} databases. While at the other end, the feature extraction time of maLBP is nearly {-13%, -30%}, {204%, 41%} and {25%, 5%} faster than the feature extraction time of cLBP, mscLBP and mCENTRIST respectively over {Corel-1k, MIT-VisTex} databases. The retrieval time using mdLBP is {4, 2.31}, {0.85, 0.77}, {1.44, 1.16}, and {2.9, 1.78} times slower than the retrieval time

using cLBP, mscLBP and mCENTRIST, and maLBP descriptors over {Corel-1k, MIT-VisTex} databases. The feature extraction time of each descriptor is nearly equal with  $u2$  and  $riu2$  transformations also. The retrieval time using mdLBPu2 and mdLBPriu2 is nearly 10 and 50 times better respectively than the retrieval time using mdLBP over Corel-1k database. The feature extraction time and retrieval time using each descriptor is more over Corel-1k database because this database is having more number of images with large resolution as compared to the MIT-VisTex database images. From the Table VI, it is explored that the maLBP is more time efficient whereas mdLBP is less time efficient as the dimension of mdLBP is higher than others except mscLBP.

#### 4.6 Comparison with non-LBP based descriptors

In order to depict the suitability of the proposed descriptors, we also compared with state-of-the-art non-LBP based descriptors such as multichannel SIFT [39], multichannel GIST [39], and color difference histogram (CDH) [38]. We have concatenated the SIFT and GIST descriptors computed over the Red, Green and Blue channels to form the mSIFT and mGIST descriptors. In order to compute the OpponentSIFT descriptor, we concatenated the SIFT descriptor computed over the opponent Red, opponent Green and opponent Blue channels. We have used the online available code of SIFT [58] used by Wang et al. [59] and GIST [60] released by Torralba [40], whereas implemented the CDH descriptor. The ARP values for 10 numbers of retrieved images using mSIFT, OpponentSIFT, mGIST, CDH, maLBP and mdLBP descriptors are listed in Table VII. It is investigated that mdLBP outperforms other descriptors over nearly each database except KTH-TIPS. The presence of images with different scales in KTH-TIPS database is the reason behind this performance. It is our belief that the scaling problem in our descriptors can be overcome by adopting the multi-scale scenario of LBP in the proposed architecture. It is also observed that the performance of maLBP is better than non-LBP based descriptors in most of the cases. Crucial information produced by this result is that the performance of mSIFT and OpponentSIFT descriptors are drastically down over textural databases as compared to the natural databases. One possible reason is that the local regions are not being detected very accurately over the textural databases.

TABLE VII  
THE ARP VALUES FOR NR=10 USING NON-LBP AND PROPOSED DESCRIPTORS

Database	Descriptors					
	mSIFT	OpponentSIFT	mGIST	CDH	maLBP	mdLBP
Corel-1k	45.14	49.93	62.83	<b>74.08</b>	73.43	<b>74.93</b>
MIT-VisTex	19.63	19.47	43.98	62.94	<b>74.86</b>	<b>80.08</b>
STex-512S	22.55	21.41	46.86	59.08	<b>77.54</b>	<b>83.89</b>
USPTex	16.80	17.67	43.93	55.89	<b>75.67</b>	<b>82.50</b>
FTVL	64.99	77.05	61.70	85.04	<b>89.74</b>	<b>95.02</b>
KTH-TIPS	41.07	43.19	<b>89.56</b>	85.23	85.12	<b>86.91</b>
KTH-TIPS2a	44.72	40.24	89.50	89.73	<b>92.23</b>	<b>94.87</b>
Corel-Tex	32.80	31.57	52.83	66.47	<b>66.93</b>	<b>67.68</b>
ALOT	23.52	26.38	42.44	44.35	<b>53.74</b>	<b>63.04</b>
ZuBuD	25.15	29.03	<b>33.49</b>	32.67	33.03	<b>34.32</b>

Copyright (c) 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

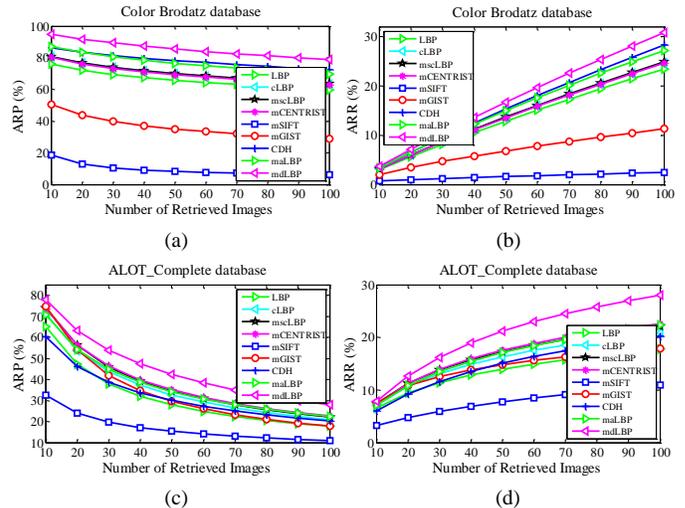


Fig. 16. Comparison of proposed descriptors maLBP and mdLBP with LBP, cLBP, mscLBP, mCENTRIST, mSIFT, mGIST and CDH over large databases such as (a-b) Colored Brodatz and (c-d) ALOT-Complete in terms of the ARP and ARR.

TABLE VIII  
PERFORMANCE ANALYSIS OF PROPOSED IDEA WITH CLBP IN TERMS OF THE ARP WHEN THE NUMBER OF RETRIEVED IMAGES IS 10

Method	Database					
	Corel-1k	KTH-TIPS	ALOT	ZuBuD	Corel-Tex	USPTex
maLBPu2	71.73	83.21	52.37	32.40	65.02	73.34
maCLBPu2	73.11	87.00	59.21	33.96	65.23	76.16
mdLBPu2	73.60	85.84	62.69	33.88	66.45	81.59
mdCLBPu2	75.29	88.69	68.73	35.15	67.75	81.89

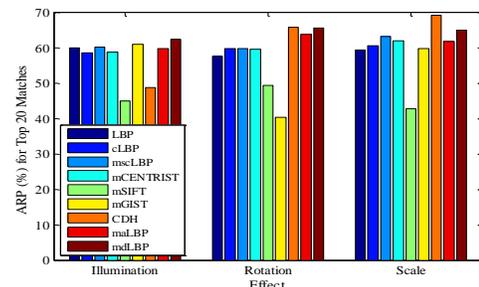


Fig. 17. Comparison of proposed descriptors maLBP and mdLBP with LBP, cLBP, mscLBP, mCENTRIST, mSIFT, mGIST and CDH over the databases having uniform illumination, rotation and image scale changes.

#### 4.7 Results over large databases

We also tested the proposed descriptors over large color texture databases such as Colored Brodatz [57] and ALOT-Complete [33]. The Colored Brodatz database has 112 color texture images of dimension  $640 \times 640$ . Each image is partitioned into 256 non-overlapping images of dimension  $40 \times 40$  which represent one category of the database. Thus, Colored Brodatz database consists of the 28672 images from 112 categories. The ALOT-Complete database is having total 25000 images from 250 categories with 100 images per category. We have turned the first 10 images of each category as the query for this experiment instead of random images to ensure the reproducibility of the results. The ARP and ARR values over these databases are depicted in Fig. 16 for LBP,

cLBP, mscLBP, mCENTRIST, mSIFT, mGIST, CDH, maLBP and mdLBP descriptors. As expected, mdLBP outperforms all the other descriptors over both color texture databases, whereas maLBP is not that good. This analysis shows that decoder based multichannel LBP descriptor can also be used for large-scale color texture image retrieval.

#### 4.8 Analyzing the CLBP in proposed architecture

In order to exhibit the generalized properties of the proposed idea, we computed the CLBP (i.e. sign and magnitude both) with adder and decoder mechanism under u2 transformation and termed as maCLBPu2 and mdCLBPu2 respectively. We also compared maCLBPu2 and mdCLBPu2 with maLBPu2 and mdLBPu2 in Table VIII over Corel-1k, KTH-TIPS, ALOT, ZuBuD, Corel-TeX and USPTex databases. It is obvious that the retrieval performance for CLBP based descriptors is generally better as it utilizes both sign and magnitude information of local differences. This experiment also explores the adoptability nature of proposed adder and decoder for LBP based descriptors.

#### 4.9 Analyzing the robustness of proposed descriptors

In order to emphasize the performance of proposed descriptors under uniform illumination, rotation and image scale differences, we synthesized the *Illumination*, *Rotation*, and *Scale* databases. The *Illumination* database is obtained by adding -60, -30, 0, 30, and 60 in each channel (i.e. Red, Green and Blue) of the first 20 images of each category of *Corel-1k* database. The *Rotate* database is obtained by rotating the first 25 images of each category of *Corel-1k* with angle 0, 90, 180, and 270 degrees. The *Scale* database is obtained by scaling the first 20 images of each category of *Corel-1k* at the scales of 0.5, 0.75, 1, 1.25, and 1.5. Thus, each database consists of the 1000 images with 100 images per category.

The retrieval results over these databases are displayed in Fig. 17 in terms of ARP (%) when 20 images are retrieved using each descriptor. The mdLBP outperforms the remaining descriptors in case of uniform intensity change. The performance of mdLBP is also better in rotation and scaling conditions except CDH. The performance of CDH is very well in rotation and scaling conditions because a) it considers the number of occurrences in local neighborhood and b) it considers the larger neighborhood respectively. The robustness of maLBP is also comparable with the similar kind of descriptors.

## 5 CONCLUSION

In this paper, two multichannel decoded local binary patterns are introduced namely multichannel adder local binary pattern (maLBP) and multichannel decoder local binary pattern (mdLBP). Basically both maLBP and mdLBP have utilized the local information of multiple channels on the basis of the adder and decoder concepts. The proposed methods are evaluated using image retrieval experiments over ten databases having images of natural scene and color textures.

The results are computed in terms of the average precision rate and average retrieval rate and improved performance is observed when compared with the results of the existing multichannel based approaches over each database. From the experimental results, it is concluded that the maLBP descriptor is not showing the best performance in most of the cases while mdLBP descriptor outperforms the existing state-of-the-art multichannel based descriptors. It is also deduced that Chi-square distance measure is better suited with the proposed image descriptors. The performance of the proposed descriptors is much improved for three input channels and also in the RGB color space. The performance of mdLBP is also superior to non-LBP descriptors. It is also pointed out that mdLBP outperforms the state-of-the-art descriptors over large databases. Experiments also suggested that the introduced approach is generalized and can be applied over any LBP based descriptor. The increased dimension of the decoder based descriptor slows down the retrieval time which is the future direction of this research. One future aspect of this research is to make the descriptors noise robust which can be achieved by using the noise robust binary patterns over each channel as the input to the adder/decoder.

## REFERENCES

1. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, 2000.
2. Y. Liu, D. Zhang, G. Lu, and W.Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.
3. S.R. Dubey, S.K. Singh and R.K. Singh, "Rotation and Illumination Invariant Interleaved Intensity Order Based Local Descriptor," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5323-5333, 2014.
4. S.R. Dubey, S.K. Singh and R.K. Singh, "A Multi-Channel based Illumination Compensation Mechanism for Brightness Invariant Image Retrieval," *Multimedia Tools and Applications*, vol. 74, no. 24, pp. 11223-11253, 2015.
5. T. Ojala, M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, 1996.
6. T. Ojala, M. Pietikäinen and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.
7. A. Hadid and G. Zhao, "Computer vision using local binary patterns," vol. 40. Springer, 2011.
8. T. Ahonen, A. Hadid and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.
9. D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 41, no. 6, pp. 765-781, 2011.
10. C. Shan, S. Gong and P.W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.
11. S.R. Dubey, S.K. Singh and R.K. Singh, "Local Diagonal Extrema Pattern: A new and Efficient Feature Descriptor for CT Image Retrieval," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1215-1219, 2015.

12. M. Heikkilä, M. Pietikäinen and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425-436, 2009.
13. S. Liao, M.W.K. Law, and A.C.S. Chung, "Dominant local binary patterns for texture classification," *IEEE Transactions on Image Processing*, vol. 18, no. 5, pp. 1107-1118, 2009.
14. Z. Guo, L. Zhang and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," *Pattern recognition*, vol. 43, no. 3, pp. 706-719, 2010.
15. Z. Guo, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, 2010.
16. X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635-1650, 2010.
17. B. Zhang, Y. Gao, S. Zhao and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 533-544, 2010.
18. S.R. Dubey, S.K. Singh and R.K. Singh, "Local Neighborhood Based Robust Colour Occurrence Descriptor for Colour Image Retrieval," *IET Image Processing*, vol. 9, no. 7, pp. 578-586, 2015.
19. W.T. Chu, C.H. Chen and H.N. Hsu, "Color CENTRIST: Embedding color information in scene categorization," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 840-854, 2014.
20. C.K. Heng, S. Yokomitsu, Y. Matsumoto and H. Tamura, "Shrink boost for selecting multi-lbp histogram features in object detection," In *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 3250-3257, 2012.
21. J.Y. Choi, K.N. Plataniotis and Y.M. Ro, "Using colour local binary pattern features for face recognition," In *17th IEEE International Conference on Image Processing (ICIP)*, pp. 4541-4544, 2010.
22. C. Zhu, C.E. Bichot and L. Chen, "Multi-scale Color Local Binary Patterns for Visual Object Classes Recognition" In *Proceedings of the IEEE International Conference on Pattern Recognition*, pp. 3065-3068, 2010.
23. C. Zhu, C.E. Bichot and L. Chen, "Image region description using orthogonal combination of local binary patterns enhanced with color information," *Pattern Recognition*, vol. 46, no. 7, pp. 1949-1963, 2013.
24. S. Banerji, A. Verma and C. Liu, "Novel color LBP descriptors for scene and image texture classification," In *Proceedings of the 15th International Conference on Image Processing, Computer Vision, and Pattern Recognition*, Las Vegas, Nevada, pp. 537-543, 2011.
25. S.H. Lee, J.Y. Choi, Y.M. Ro and K.N. Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2347-2353, 2012.
26. Y. Xiao, J. Wu and J. Yuan, "mCENTRIST: A Multi-Channel Feature Generation Mechanism for Scene Categorization," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 823-836, 2014.
27. Corel Photo Collection Color Image Database, online available on <http://wang.ist.psu.edu/docs/realtd/>.
28. MIT Vision and Modeling Group, Cambridge, 'Vision texture', <http://vismod.media.mit.edu/pub/>.
29. Salzburg Texture Image Database, <http://www.wavelab.at/sources/STex/>.
30. A. R. Backes, D. Casanova, and O. M. Bruno, "Color texture analysis based on fractal descriptors," *Pattern Recognition*, vol. 45, no. 5, pp. 1984-1992, 2012.
31. <http://www.ic.unicamp.br/~rocha/pub/downloads/tropical-fruits-DB-1024x768.tar.gz/>.
32. KTH-TIPS texture image database, online available on <http://www.nada.kth.se/cvap/databases/kth-tips/index.html>.
33. G. J. Burghouts and J. M. Geusebroek, "Material-specific adaptation of color invariant features", *Pattern Recognition Letters*, vol. 30, pp. 306-313, 2009.
34. Zurich Buildings Database (ZuBuD), <http://www.vision.ee.ethz.ch/datasets/index.en.html>.
35. S.R. Dubey, S.K. Singh and R.K. Singh, "Rotation and scale invariant hybrid image descriptor and retrieval," *Computers & Electrical Engineering*, vol. 46, pp. 288-302, 2015.
36. S.R. Dubey, S.K. Singh and R.K. Singh, "Local Bit-plane Decoded Pattern: A Novel Feature Descriptor for Biomedical Image Retrieval," *IEEE Journal of Biomedical and Health Informatics*, 2015. (In Press)
37. S.R. Dubey, S.K. Singh and R.K. Singh, "Local Wavelet Pattern: A New Feature Descriptor for Image Retrieval in Medical CT Databases," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5892-5903, 2015.
38. G.H. Liu and J.Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognition*, vol. 46, no. 1, pp. 188-198, 2013.
39. Koen E.A. Van De Sande, T. Gevers and Cees G.M. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1582-1596, 2010.
40. A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International Journal of computer vision*, vol. 42, no. 3, pp. 145-175, 2001.
41. M. Brown and S. Süssstrunk, "Multi-spectral SIFT for scene category recognition," In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 177-184, 2011.
42. M. Douze, H. Jégou, H. Sandhawalia, L. Amsaleg and C. Schmid, "Evaluation of gist descriptors for web-scale image search," In *ACM International Conference on Image and Video Retrieval*, pp. 19-27, 2009.
43. H. Jégou, M. Douze, C. Schmid and P. Pérez, "Aggregating local descriptors into a compact image representation," In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3304-3311, 2010.
44. J. Yue-Hei Ng, F. Yang and L.S. Davis, "Exploiting Local Features from Deep Networks for Image Retrieval," In *IEEE International Conference on Computer Vision and Pattern Recognition, DeepVision Workshop (CVPRW)*, pp. 53-61, 2015.
45. H. Jégou, M. Douze and C. Schmid, "Improving bag-of-features for large scale image search," *International Journal of Computer Vision*, vol. 87, no. 3, pp. 316-336, 2010.
46. V. Sydorov, M. Sakurada and C.H. Lampert, "Deep Fisher Kernels--End to End Learning of the Fisher Kernel GMM Parameters," In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1402-1409, 2014.
47. K. Simonyan, A. Vedaldi and A. Zisserman, "Deep fisher networks for large-scale image classification," In *Advances in neural information processing systems*, pp. 163-171, 2013.
48. A. Krizhevsky, I. Sutskever and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," In *Advances in neural information processing systems*, pp. 1097-1105, 2012.
49. F. Perronnin and D. Larlus, "Fisher Vectors Meet Neural Networks: A Hybrid Classification Architecture," In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3743-3752, 2015.
50. X. Zhou, K. Yu, T. Zhang and T.S. Huang, "Image classification using super-vector coding of local image descriptors," In *European Conference of Computer Vision*, pp. 141-154, 2010.
51. X. Bai, C. Yan, P. Ren, L. Bai and J. Zhou, "Discriminative sparse neighbor coding," *Multimedia Tools and Applications*, 2015. (In Press)
52. N. Inoue and K. Shinoda, "Fast Coding of Feature Vectors using Neighbor-To-Neighbor Search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015. (In Press)
53. X. Li, M. Fang and J.J. Zhang, "Projected transfer sparse coding for cross domain image representation," *Journal of Visual Communication and Image Representation*, vol. 33, pp. 265-272, 2015.
54. C. Zhang, J. Cheng, J. Liu, J. Pang, Q. Huang and Q. Tian, "Beyond explicit codebook generation: Visual Representation Using Implicitly Transferred codebooks," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5777 - 5788, 2015.
55. S. Murala and QM Jonathan Wu, "Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval," *Neurocomputing*, vol. 119, pp. 399-412, 2013.
56. [www.cs.columbia.edu/~mmler/project/code/pdist2.m](http://www.cs.columbia.edu/~mmler/project/code/pdist2.m).

57. Colored Brodatz Database, [http://multibandtexture.recherche.usherbrooke.ca/colored%20\\_brodatz.html](http://multibandtexture.recherche.usherbrooke.ca/colored%20_brodatz.html).
58. <http://www.ifp.illinois.edu/~jyang29/LLC.htm>.
59. J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality constrained linear coding for image classification," In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3360–3367, 2010.
60. <http://people.csail.mit.edu/torr/alba/code/spatialenvelope/>.
61. L. Liu, M. Yu, and L. Shao, "Multiview alignment hashing for efficient image search," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 956-966, 2015.
62. J. Tang, Z. Li, M. Wang, and R. Zhao, "Neighborhood discriminant hashing for large-scale image retrieval," *IEEE Transactions on Image Processing*, vol. 24, no. 9, pp. 2827-2840, 2015.
63. L. Liu and L. Shao, "Sequential Compact Code Learning for Unsupervised Image Hashing," *IEEE Transactions on Neural Networks and Learning Systems*, 2015. (In Press)
64. L. Liu, M. Yu, and Ling Shao, "Unsupervised local feature hashing for image similarity search," *IEEE Transactions on Cybernetics*, 2015. (In Press)