

Robust Image Retrieval with Ring Partition and Multichannel Decoded Local Binary Patterns

Anjaly N S¹, Riyamol Sidhique²

P.G. Student, Department of Computer Science Engineering, Indira Gandhi Institute of Engineering and

Technology for Women, Nellikuzhy, Kerala, India¹

Associate Professor, Department of Computer Science Engineering, Indira Gandhi Institute of Engineering and

Technology for Women, Nellikuzhy, Kerala, India²

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. We incorporate ring partition to image retrieval for enhancing rotation robustness. As ring partition is unrelated to image rotation, the statistical features that are extracted from image rings in perceptually uniform color space, are rotation invariant and stable. 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. 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.

KEYWORDS: Image retrieval, Local patterns, Multichannel, LBP, Color, Texture.

I. 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. 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), Neighborhood Discriminant Hashing (NDH), Evolutionary Compact Embedding (ECE) and unsupervised bilinear local hashing (UBLH). These methods can be used with the high discriminative descriptors to improve the efficiency of image search. These approaches are introduced basically for

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

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. 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. Another interesting descriptor is GIST which is basically a holistic representation of features and has gained wider publicity due its high discriminative ability. In order to encode the region based descriptors into a single descriptor, a vector locally aggregated descriptors (VLAD) has been proposed in the literature .

Paper is organized as follows. Section II describes multichannel feature extraction approaches. The flow diagram represents the step of the algorithm. Section III describes the proposed system with the flowchart of the workflow. Section IV presents experimental results showing results of images tested. Finally, Section V presents conclusion.

II. RELATED WORK

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) , Neighborhood Discriminant Hashing (NDH) , Evolutionary Compact Embedding (ECE) and unsupervised bilinear local hashing (UBLH). These methods can be used with the high discriminative descriptors to improve the efficiency of image search.

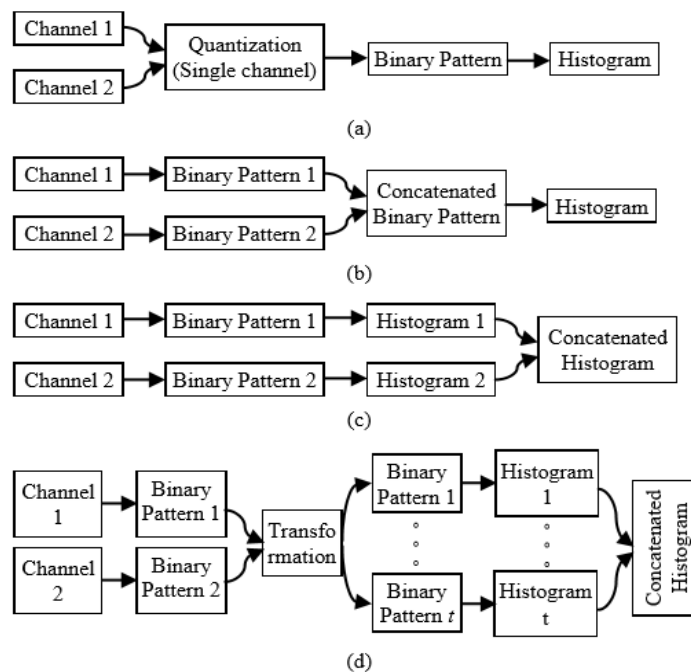


Fig. 1. Illustration of four types of the multichannel feature extraction technique using two input channels, (a) Each channel is quantized and merged to form a single channel and then descriptor is computed over it, (b) Binary patterns extracted over each channel are concatenated to form a single binary 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

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

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

quantizes each channel then merges each quantized channel to form a single channel and form the feature vector over it. 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.

Heng et al. 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. 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). . 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 4th 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 cooccurrence 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.

III. MULTICHANNEL DECODED LOCAL BINARY PATTERNS WITH RING PARTITION

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 (maLBP) 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^{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 R of any pixel $I_t(x, y)$ for $x \in [1, u]$ and $y \in [1, v]$ are defined as $I_t^n(x, y)$, where $n \in [1, N]$. Then, according to the definition of the Local Binary Pattern (LBP), a local binary pattern $LBP_t(x, y)$ for a particular pixel (x, y) in t^{th} channel is generated by computing a binary value $LBP_t^n(x, y)$ given by the following equation,

$$LBP_t(x, y) = \sum_{n=1}^N LBP_t^n(x, y) \times \mathcal{F}^n \quad (1)$$

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,

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

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

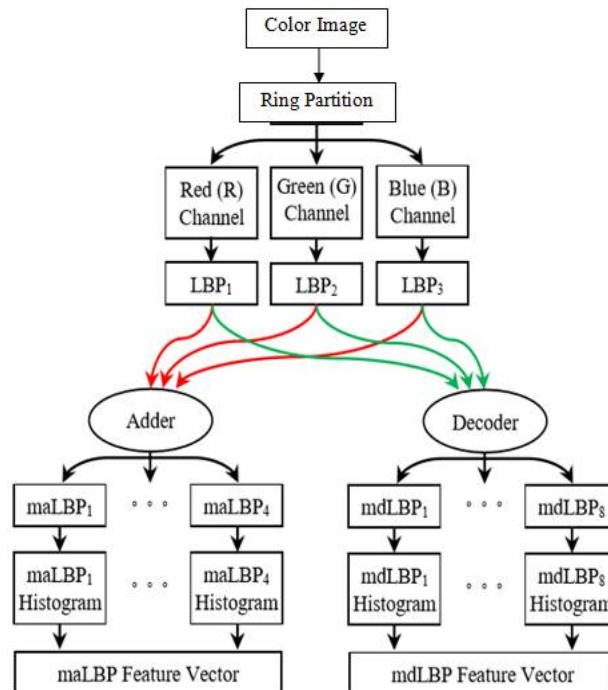


Fig. 2. The flowchart of computation of multichannel adder based local binary pattern feature vector and multichannel decoder based local binary pattern feature vector.

The process of computation of $maLBP$ and $mdLBP$ feature descriptor of an image is illustrated in Fig.2 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

We have set of N binary values $LBP(x,y)$ for a particular pixel (x,y) corresponding to each neighbor $I_t^n(x,y)$ of t^{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 $maLBP_{t1}^n(x,y)$ and multichannel decoder based local binary patterns $mdLBP_{t2}^n(x,y)$ are the outputs of the multichannel LBP adder and multichannel LBP decoder respectively, where $t1 \in [1,c+1]$ and $t2 \in [1,2c]$. Note that the values of $LBP_{tn}(x,y)$ are in the binary form (i.e. either 0 or 1). Thus, the values of $maLBP_{t1}^n(x,y)$ and $mdLBP_{t2}^n(x,y)$ 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 of pixel (x,y) .

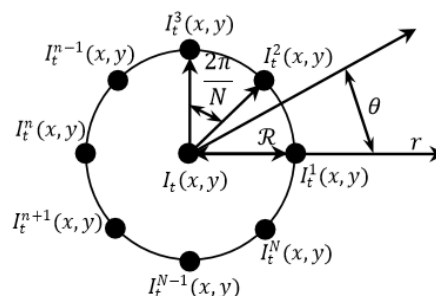


Fig. 3 The local neighbors of a center pixel

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

Mathematically, the $maM(x,y)$ and $mdM_n(x,y)$ are defined as,

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

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

We denote $LBP_t^n(x,y)$ for $\forall n \in [1,N]$ and $\forall t \in [1,c]$ by input patterns, $maLBP_{t1}^n(x,y)$ for $\forall n \in [1,N]$ and $\forall t1 \in [1,c+1]$ by adder patterns and $mdLBP_{t2}^n(x,y)$ for $\forall n \in [1,N]$ and $\forall t2 \in [1,2^c]$ by decoder patterns respectively. The multichannel adder based local binary pattern $maLBP_{t1}^n(x,y)$ for pixel x,y from multichannel adder map $maM_n(x,y)$ and $t1$ is defined as,

$$maLBP_{t1}^n(x,y) = \begin{cases} 1, & \text{if } maM^n(x,y) = (t1 - 1) \\ 0, & \text{Otherwise} \end{cases} \quad (6)$$

for $\forall t1 \in [1,c+1]$ and $\forall n \in [1,N]$

Similarly, the multichannel decoder based local binary pattern $mdLBP_{t2}^n(x,y)$ for pixel (x,y) from multichannel decoder map $mdM_n(x,y)$ and $t2$ can be computed as,

$$mdLBP_{t2}^n(x,y) = \begin{cases} 1, & \text{if } mdM^n(x,y) = (t2 - 1) \\ 0, & \text{Otherwise} \end{cases} \quad (7)$$

For $\forall t2 \in [1,2^c]$ and $\forall n \in [1,N]$.

The multichannel adder based local binary patterns ($maLBP_{t1}^n(x,y) \mid \forall t1 \in [1,c+1]$) for the center pixel (x,y) is computed using $maLBP_{t1}^n(x,y)$ in the following manner,

$$maLBP_{t1}(x,y) = \sum_{n=1}^N maLBP_{t1}^n(x,y) \times \mathcal{F}^n \quad (8)$$

and multichannel decoder based local binary patterns ($mdLBP_{t2}^n(x,y) \mid \forall t2 \in [1,2^c]$) is computed for the center pixel x,y from $mdLBP_{t2}^n(x,y)$ using the following equation,

$$mdLBP_{t2}(x,y) = \sum_{n=1}^N mdLBP_{t2}^n(x,y) \times \mathcal{F}^n \quad (9)$$

where $\mathcal{F}^n$ is a weighting function defined in (3).

The feature vector (i.e. histogram) of $t1^{th}$ output channel of the adder (i.e. $maLBP_{t1}$) is computed using the following equation,

$$\mathcal{H}^{maLBP_{t1}(\zeta)} = \frac{1}{(u-2R)(v-2R)} \sum_{x=R+1}^{u-R} \sum_{y=R+1}^{v-R} \xi(maLBP_{t1}(x,y), \zeta) \quad (10)$$

for $\forall \zeta \in [0,2^N - 1]$ and $\forall t1 \in [1,c+1]$, 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 $t2^{th}$ output channel of the decoder (i.e. $mdLBP_{t2}$) is computed as follows,

$$\mathcal{H}^{mdLBP_{t2}(\zeta)} = \frac{1}{(u-2R)(v-2R)} \sum_{x=R+1}^{u-R} \sum_{y=R+1}^{v-R} \xi(mdLBP_{t2}(x,y), \zeta) \quad (12)$$

for $\forall \zeta \in [0,2^N - 1]$ and $\forall t2 \in [1,2^c]$.

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

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

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

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

IV. EXPERIMENTAL RESULTS

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 are as follows: 1) Euclidean distance, 2) L1 or Manhattan distance, 3) Canberra distance, 4) Chi-square (Chisq) or χ^2 distance, 5) Cosine distance, and 6) D1 distance.

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,

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

Where 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,

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

Where 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,

$$Pr(k) = \frac{NS}{NR} \quad \& \quad Re(k) = \frac{NS}{ND} \quad \forall k \in [1, C_j] \quad (17)$$

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.

Table I
ARP(%) Using different distance measures

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

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 (χ^2), Cosine and D1 distance measures and shown the results in Table I for 10 numbers of retrieved images. 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 χ^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 χ^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 χ^2 distance measure will be used to find the dissimilarity between two descriptors.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

V. CONCLUSION

We have proposed a robust image retrieval with ring partition which has good rotation robustness and reaches a desirable discrimination. Since ring partition is unrelated to image rotation, our ring-based statistical features, extracted from the RGB color space, are rotation-invariant and stable. This leads to that our retrieval can resist image rotation with any angle.

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] L. Liu, M. Yu, and Ling Shao, "Unsupervised local feature hashing for image similarity search," IEEE Transactions on Cybernetics, 2015.
- [4] 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.
- [5] 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.
- [6] A. Hadid and G. Zhao, "Computer vision using local binary patterns," vol. 40. Springer, 2011.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.