Boosting Local Binary Pattern with Bag-of-Filters for Content Based Image Retrieval

Shiv Ram Dubey, *Student Member, IEEE*, Satish Kumar Singh, *Senior Member, IEEE*,
Rajat Kumar Singh, *Senior Member, IEEE*
Indian Institute of Information Technology, Allahabad, India
shivram1987@gmail.com, sk.singh@iiita.ac.in, rajatsingh@iiita.ac.in

Abstract—**Local descriptors are mostly extracted over the raw intensity values of the image and requires the analysis of patterns in local pixel neighborhoods. The raw images don't provide much local relationship information required for the local descriptors. To resolve this problem, a new method of local image feature description is proposed in this paper based on the patterns extracted from the multiple filtered images. The whole process is structured as follows: First, the images are filtered with the bag of filters (BoF) and then local binary pattern (LBP) is computed over each filtered image and finally concatenated to find out the single BoF-LBP descriptor. The local information enriches in the filtered images improves the discriminative ability of the final descriptor. Content based image retrieval (CBIR) experiments are performed to observe the effectiveness of the proposed approach and compared with the recent state-of-the-art methods. The experiments are conducted over four benchmark databases having natural and texture images such as Corel-1k, Corel-10k, MIT-VisTex and STex-512S. The experimental results confirm that the introduced approach is able to improve the retrieval performance of the CBIR system.**

Index Terms— **Image retrieval, Local patterns, Gradient filters, LBP, SLBP, Texture**

I. INTRODUCTION

CONTENT based image retrieval (CBIR) is demanding more and more attention due to its rapid growth in many places. The aim of CBIR is to extract the similar images of a given image from huge databases on the basis of the visual features such color, texture, etc. [1]. Image retrieval has several applications such as in biomedical, agriculture, object recognition, etc. Zhang et al. have surveyed the several CBIR systems in view of the high-level semantics [2].

The discriminative feature description is required to model the characteristics of the image for more accurate CBIR system. Various feature descriptors are proposed and used for image retrieval and gained wide publicity [3]-[9]. These methods have utilized color, texture and local information of the image to design the feature descriptor. The local descriptors such as local binary pattern (LBP) have shown very promising discriminative ability in several applications [10]. The LBP is widely adopted in the Computer Vision research community for its simplicity as well as effectively [11]. Various variants of LBP are available through the published literature which is inspired by the great success of LBP. Some typical examples are Local Ternary Pattern (LTP) [12], Local Derivative Pattern (LDP) [13], Interleaved Intensity Order Based Local Descriptor (IOLD) [26], and

Local Tetra Pattern (LTrP) [6]. These descriptors are mainly computed over the raw intensity values. In order to utilize the richer local information, many researchers performed some kind of preprocessing before the feature extraction. Some typical examples are Sobel Local Binary Pattern (SOBEL-LBP) [14], Local Edge Binary Pattern (LEBP) [4], Semi Structure Local Binary Pattern (SLBP) [15] and Spherical Symmetric 3D Local Ternary Pattern (SS-3D-LTP) [8]. James has compared the preprocessed images directly which is obtained by multiple filtering [16] for face recognition.

Inspired from the [4], [8], [14-16] we have proposed a Bag-of-Filters (BoF) as well as local information based image feature description for CBIR system. The idea behind this is to enrich the local information of the image by BoF which is further capitalized by the local descriptor such as LBP. In other words, the LBP encodes the local relationship of the image, whereas BoF highlights the local information such as edges and corners. So, these two operations are complementary to each other and will be more discriminating if combined together. It will be shown in the experiment that the performance of LBP is boosted significantly in conjunction with the BoF.

In the rest of this paper, we presented the idea of combining Bag-of-Filter and Local Binary Pattern in section II as well as experimental results in section III with concluding remarks in section IV.

II. PROPOSED CBIR SYSTEM

The schematic diagram of the proposed CBIR system is presented in the Fig. 1. First of all, each image is processed by Bag-of-Filters (BoF) to obtain the multiple gradient filtered images which are having different kinds of crucial information such as edges, corners, etc. Basically, these filters enhance the local information contained in the image. In order to encode such information locally in the descriptor form, Local Binary Pattern (LBP) operator is applied over each filtered image and finally all descriptors are concatenated to construct the final BoF-LBP feature descriptor. On the basis of the proposed descriptor the query image is matched with the database images by finding the distance between the BoF-LBP descriptor of query image and database images. The most relevant images of the query image are retrieved from the database on the basis of the shortest distances between the query image and database images.
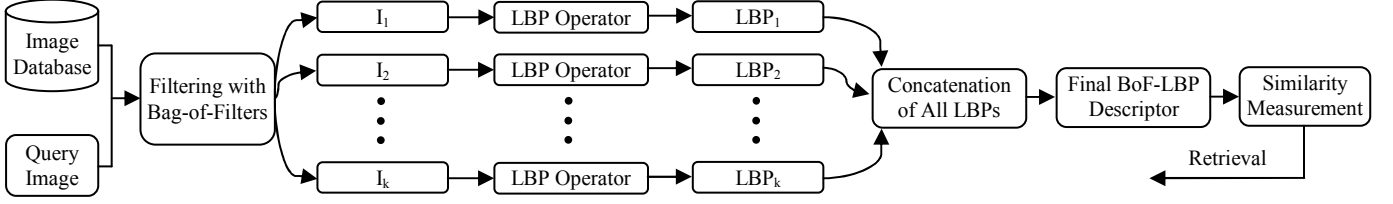
Fig. 1. The working framework of the proposed Content Based Image Retrieval (CBIR) system using Bag-of-Filters and Local Binary Pattern (LBP).

Let, I is a gray scaled image of dimension $m \times n$ (i.e. I is having m rows and n columns) over which we want to compute the BoF-LBP descriptor. Any particular pixel (x, y) of image I is denoted by $I^{x,y}$. Consider $F_i|_{i=1,2,\dots,k}$ is the mask for i^{th} filter of BoF. Let the image I_i is the filtered image obtained by applying the ith filter over the image I, and defined by the following formulae,

$$I_i = I * F_i \qquad (1)$$

where the operator '$*$' represents the convolution operation over image I by mask F_i. The image I_i basically highlights the important features of the image I which are having characteristics of the filter mask F_i. Note that if the dimensions of I and F_i are $m \times n$ and $m_i^F \times n_i^F$ respectively, then the dimension of I_i (i.e. $m_i^I \times n_i^I$) can be given as,

$$m_i^I = m - m_i^F + 1 \qquad (2)$$

$$n_i^I = n - n_i^F + 1 \qquad (3)$$

Now, in order to encode the local characteristics of the image I_i, the Local Binary Pattern (LBP) [10] operator is computed over it and the resulting image is represented by the LBP_i which is having dimension of $m_i^{LBP} \times n_i^{LBP}$. Let, the LBP_i value computed for any particular pixel (x, y) of I_i (i.e. for $I_i^{x,y}$) is represented by $LBP_i^{x,y}$. The $LBP_i^{x,y}$ is calculated in following manner,

$$LBP_i^{x,y} = \sum_{\omega=1}^{N} LBP_i^{x,y,\omega} \times (2)^{(\omega-1)} \qquad (4)$$

where N is the number of local neighbors of a pixel (x, y) at a circle of radius R from the (x, y) and spaced equally. If $I_i^{x_\omega, y_\omega}$ is the intensity value of the ωth neighbor of $I_i^{x,y}$ then its coordinates (x_ω, y_ω) are given as follows,

$$x_\omega = x - R \times sin\left((\omega-1) \times \frac{2\pi}{N}\right) \qquad (5)$$

$$y_\omega = y + R \times cos\left((\omega-1) \times \frac{2\pi}{N}\right) \qquad (6)$$

The $LBP_i^{x,y,\omega}$ is the binary value obtained by comparing the center pixel $I_i^{x,y}$ with its ωth neighbor $I_i^{x_\omega, y_\omega}$ and mathematically computed as,

$$LBP_i^{x,y,\omega} = \begin{cases} 1 & if\ I_i^{x_\omega,y_\omega} \geq I_i^{x,y} \\ 0 & Else \end{cases} \qquad (7)$$

The final BoF-LBP feature vector is calculated by concatenating the LBP feature vector computed over each filtered image and given as follows,

$$BoF\text{-}LBP = [LBP_{1,h}, LBP_{2,h}, \dots, LBP_{k,h}] \qquad (8)$$

where $LBP_{i,h}$ is the histogram of LBP_i (i.e. LBP map over the ith filtered image) and determined using the following formulae,

$$LBP_{i,h} = \sum_{x=R+1}^{m_i^I - R} \sum_{y=R+1}^{n_i^I - R} \xi\left(LBP_i^{x,y}, \gamma\right) \qquad (9)$$

for $\forall\ \gamma \in [0, 2^N - 1]$, where, $\xi(\alpha, \beta)$ is given as,

$$\xi(\alpha, \beta) = \begin{cases} 1, & if\ \alpha = \beta \\ 0, & Else \end{cases} \qquad (10)$$

Finally, the $BoF\text{-}LBP$ feature vector is normalized such that the range of its values become 0 to 1.

In this paper, we have used $k = 5$ different type of filter masks as the Bag-of-Filters which are capturing varying range of information. The five filters used here are Average, Horizontal-vertical difference, Diagonal difference, Sobel edge in vertical direction and Sobel edge in horizontal direction (Refer [16] for more detail about these filters). Fig. 2 presents the 3×3 mask for these filters.

Fig. 2. The five filters used in this paper as the Bag-of-Filters are, (a) Average filter, i.e. F_1, (b) Horizontal-vertical difference filter, i.e. F_2, (c) Diagonal filters, i.e. F_3, (d) Sobel edge in vertical direction, i.e. F_4, and (e) Sobel edge in horizontal direction, i.e. F_5.
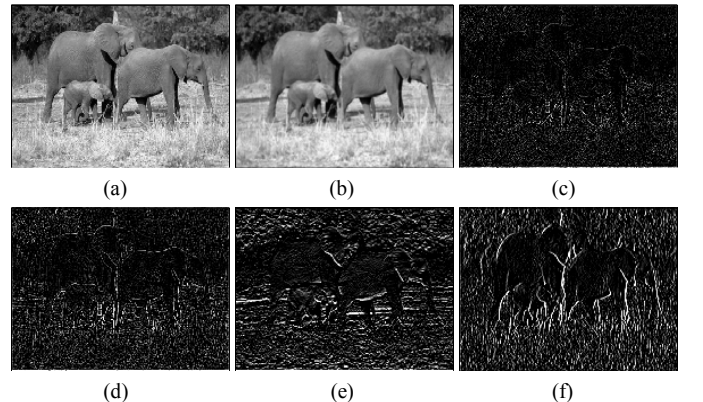

Fig. 3. (a) An example image, (b-f) the image obtained after applying the 5 filters with mask $F_i|_{i=1,2,3,4,5}$ respectively over the example image of (a).

The images obtained by convolving the five considered filter mask over an example image are depicted in Fig. 3. It can be observed that the Average filter (i.e. F_1) gives the low frequency information (i.e. smooth variations), whereas, the remaining filters (i.e. $F_i|_{i=2,3,4,5}$) provide the high frequency oriented information (i.e. edges in particular directions). The combination of both types of filters increase the discriminating ability of $BoF\text{-}LBP$ descriptor.

III. EXPERIMENTS AND RESULTS

In this section first we discuss the databases used for the experiments in this paper, then the distance measures and evaluation criteria are explained and finally the experimental results are presented with discussions.

Four benchmark and widely adopted databases, namely Corel-1k [17], Corel-10k [18], MITVis-Tex [19] and STex-512S [20] are used for the image retrieval experiments here. Corel-1k and Corel-10k are the natural databases, whereas, MITVis-Tex and STex-512S are the texture databases. We have converted all the images into the grayscale if it is a color image for the extraction of the descriptors. The number of categories in the database as well as the number of images in a category with a total number of images in the database is summarized in Table 1.

In order to find the distance (i.e. dissimilarity) between two feature vectors, Euclidean, Canberra, L1, D1 and Chi-square (χ^2) distance measures are used in this paper [6] [21-24, 27]. We evaluated the performance of descriptors for CBIR system over a particular database using Average Retrieval Precision (ARP) and Average Retrieval Rate (ARR) metrics in this paper. ARP and ARR are computed by finding the mean of Average Precisions and Average Recalls obtained over all categories of that database respectively. The Average Precision and Average Recall for a particular category of any database is determined by finding the mean of precisions and recalls respectively, for each image of that category (i.e. by turning each image of that category as the query image). The precision and recall for a query image is computed as follows:

$$Precision = \frac{\#CIR}{\#IR} \quad \& \quad Recall = \frac{\#CIR}{\#MID} \qquad (11)$$

where $\#CIR$ is the number of correct images retrieved, $\#IR$ is the number of images retrieved, and $\#MID$ is the number of matching images in the database.

In order to justify the improved performance of proposed BoF-LBP descriptor, we have compared the retrieval results with recent and state-of-the-art descriptors such as Local Binary Pattern (LBP) [11], Semi-structure Local Binary Pattern (SLBP) [15], Sobel Local Binary Pattern (SOBEL-LBP) [14], Local Ternary Pattern (LTP) [12], Local Derivative Pattern (LDP) [13], Local Tetra Pattern (LTrP) [6], and Spherical Symmetric 3-Dimensional Local Ternary Pattern (SS-3D-LTP) [8]. We have used the online available code of LBP descriptor [25] and implemented the rest of descriptors.

The effect of Euclidean, Canberra, L1, D1, and Chi-square distance measures over the performance of BoF-LBP descriptor is listed in Table II in terms of the ARP when the number of images retrieved ($\#IR$) is 10. The performance of

proposed descriptor over Corel-1k and STex-512S is better using Canberra and D1 distances respectively, whereas, its performance is better using Chi-square distance over other databases. The average performance using each distance measure over all databases is also reported in the last row of this table with best performance using Chi-square. On the basis of this result, we assumed the Chi-square distance measure to find the dissimilarity between two feature vectors in the rest of the results of this paper.

The performance comparison among each descriptor is made using the plots of ARP (%) vs #IR and ARR (%) vs #IR over natural databases in the Fig. 4-5 (i.e. results over Corel-1k and Corel-10k databases are plotted in the Fig. 4 and Fig. 5 respectively). It is clearly seen that the proposed BoF-LBP descriptor outperforms the LBP, SLBP, SOBEL-LBP, LTP, LDP, LTrP, and SS-3D-LBP descriptors over both the natural databases using both ARP and ARR evaluation metrics. We also reported the results over textural databases such as MITVis-Tex and STex-512S in terms of the ARP (%) as a function of the number of top matches (i.e. number of retrieved images) in Fig. 6. The performance of our method is clearly better than the other ones over both the textural databases. It is also noticed across the plots of Fig. 4-6 that the improvement caused by the proposed descriptor is more if the size of the database is large (see the database size and results over Corel-10k and STex-512S databases as compared to the remaining descriptors). From these results, we can make an assertion that the introduced feature descriptor outperforms recent and state-of-the-art feature descriptors over both natural and textural images.

TABLE I
IMAGE DATABASES SUMMARY

No.	Database Name	Image Size	#Class	# Images in Each Class	#Images (Total)
1.	Corel-1k [17]	384×256	10	100	1000
2.	Corel-10k [18]	120×80	80	Vary	10,800
3.	MIT-VisTex [19]	128×128	40	16	640
4.	STex-512S [20]	128×128	26	Vary	7616

TABLE II
ARP VALUES USING BoF-LBP WHEN NUMBER OF TOP MATCHES ARE 10

Database/Distance	Euclidean	Canberra	L1	D1	χ^2
Corel-1k	63.53	**73.84**	71.35	71.59	72.50
Corel-10k	33.07	31.25	38.35	38.65	**39.87**
MIT-VisTex	67.06	67.67	72.66	72.64	**73.25**
STex-512S	71.35	69.43	79.18	**79.32**	78.95
Average	58.75	60.55	65.39	65.55	**66.14**

TABLE III
PERFORMANCE COMPARISON OF $BoF\text{-}LBP$ WITH $LBP_{i,h}|_{i=1,2,3,4,5}$ IN TERMS OF THE ARP VALUES OVER EACH DATABASE WHEN $\#IR = 10$

Database	$LBP_{1,h}$	$LBP_{2,h}$	$LBP_{3,h}$	$LBP_{4,h}$	$LBP_{5,h}$	*BoF-LBP*
Corel-1k	68.47	68.45	70.29	66.37	70.22	**72.50**
Corel-10k	35.02	30.95	31.30	28.67	29.06	**39.87**
MIT-VisTex	67.47	63.92	64.56	65.70	66.58	**73.25**
Stex-512S	68.78	64.87	61.63	60.79	60.50	**78.95**

The ARP values in percentage when number of top matches are 10 over each database for $LBP_{i,h}|_{i=1,2,3,4,5}$ and BoF-LBP descriptors (i.e. BoF as well standalone Filtered LBPs) are demonstrated in Table III. It can be visualized that the performance of combined filtered LBP (i.e. BoF-LBP) is highly improved as compared to the standalone filtered LBP. Among standalone filtered LBPs, the descriptor over averaged filtered image (i.e. for i = 1) is having a more discriminative ability than the remaining ones in most of the cases. One possible explanation for this behavior is the low frequency information contained by the averaging filter. We have shown

the top 10 matching retrieved images for one query images from each database using LBP and filtering based descriptors such as LBP, SLBP, SOBEL-LBP, and BoF-LBP in Fig. 7. The incorrect matches are highlighted in 'Red color' rectangles. For the query image from Corel-1k database (i.e. first column of Fig. 7 (a)), the precision obtained by LBP, SLBP, SOBEL-LBP, and BoF-LBP descriptors are 90%, 80%, 90%, and 100% respectively (see the last 10 columns of 1st to 4th row of Fig. 7 (a) respectively). The number of incorrect matches out of 10 is 4, 6, 7, and 2 for LBP, SLBP, SOBEL-LBP, and BoF-LBP respectively for a query from Corel-10k.
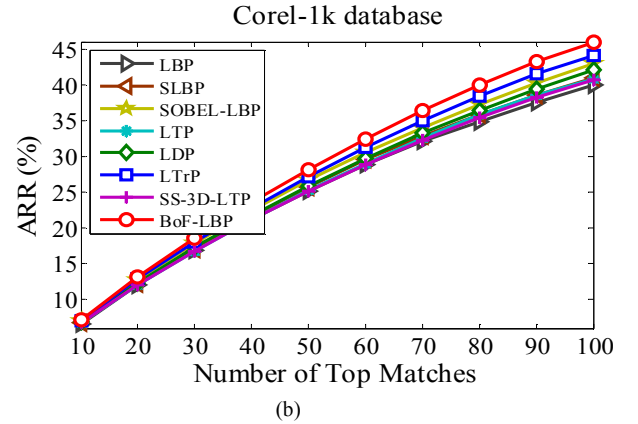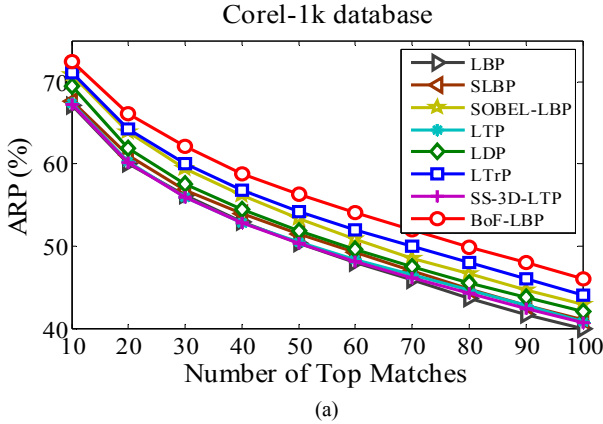

Fig. 4. The performance comparison BoF-LBP descriptor with LBP, SLBP, SOBEL-LBP, LTP, LDP, LTrP, and SS-3D-LTP descriptors over Corel-1k database using (a) ARP (%) and (b) ARR (%).
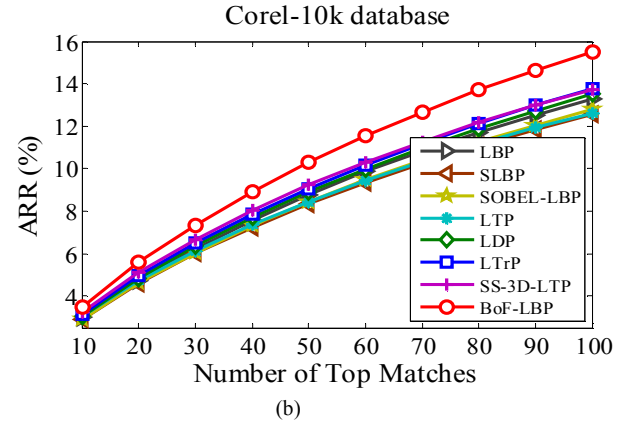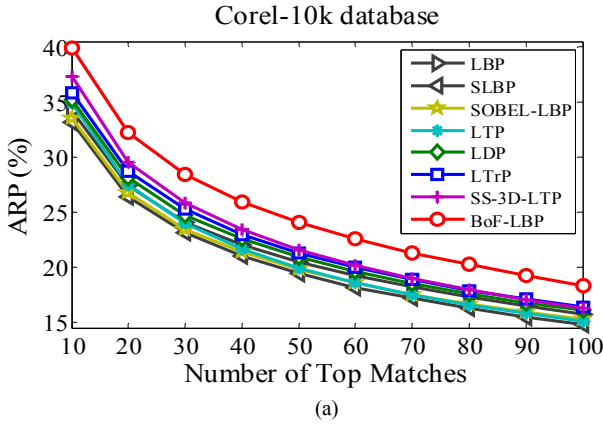

Fig. 5. The performance comparison proposed descriptor with other descriptors over Corel-10k database using (a) ARP (%) and (b) ARR (%).
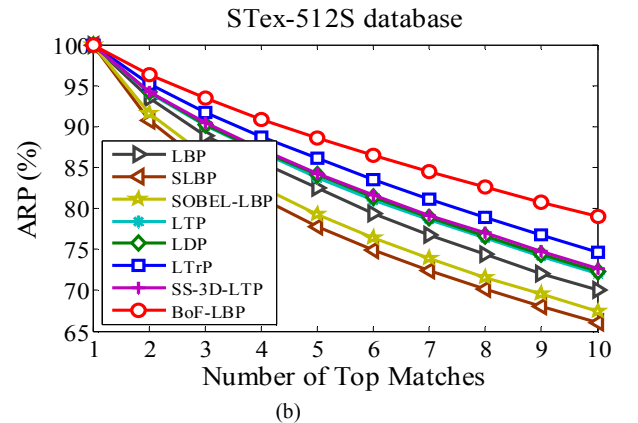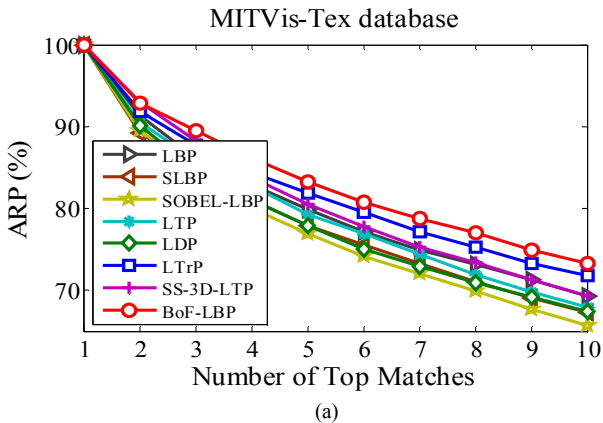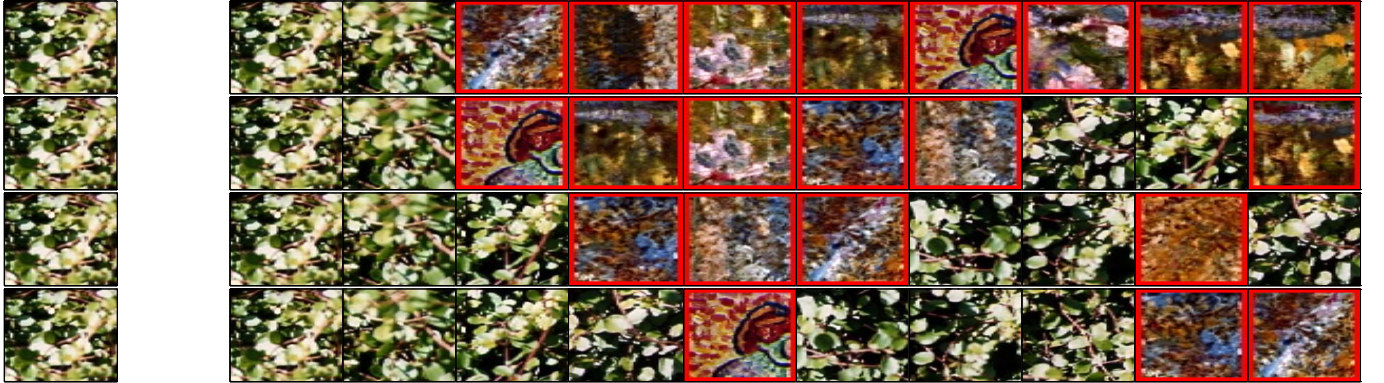

Fig. 6. The performance of each descriptor in terms of the ARP (%) over (a) MITVis-Tex and (b) STex-512S databases.

(a) Query image is taken from the Corel-1k database

(b) Query image is taken from the Corel-10k database

(d) Query image is taken from the MITVis-Tex database

(d) Query image is taken from the STex-512S database

Fig. 7.   The top 10 retrieved images (in the last 10 columns) for a query image (in the first column) from (a) Corel-1k, (b) Corel-10k, (c) MITVis-Tex, and (d) STex-512S database using LBP (in the 1st row), SLBP (in the 2nd row), SOBEL-LBP (in the 3rd row), and BoF-LBP (in the 4th row). Note that the incorrect retrieved images are enclosed in 'Red' rectangles for clear visualization. It should also be noted that the images are shown in color here, but the descriptors are computed over gray scale version of color images.

TABLE IV
THE DIMENSIONS OF EACH DESCRIPTOR

Descriptor	LBP	SLBP	SOBEL − LBP	LTP	LDP	LTrP	SS − 3D − LTP	BoF − LBP
Dimension	256	256	2×256	2×256	4×256	13×256	10×256	5×256

The performance of proposed descriptor is far better than other ones over natural database query images. The similar trend can also be visualized over the textural database query images as displayed in Fig. 7 (c-d). Moreover, it is also noticed here that the performance of BoF-LBP is more improved over the databases having more number of images.

The dimension of each descriptor is also summarized in the Table IV. It can be seen that the dimension of BoF-LBP is lower than LTrP and SS-3D-LTP, whereas the performance is better. Moreover, the dimension of BoF-LBP is higher than the remaining ones, whereas the performance is improved significantly.

IV. CONCLUSION

A bag-of-filter (BoF) based local binary pattern (LBP) coding scheme is presented in this paper for content based image retrieval. The BoF basically explores the local features of the image in several ways which is further captured by the LBP to generate a more discriminative BoF-LBP descriptor. The performance of introduced descriptor is tested using an image retrieval framework over four benchmark databases including two natural as well as two textural databases. The proposed one is also compared with the state-of-the-art descriptors. The experimental results point out the superiority of the BoF-LBP descriptor over the existing descriptors over each database. It is observed that the Chi-square distance measure is better suited with the BoF-LBP descriptor. It has also been experimented that the low frequency information such as smooth variations is having more importance as compared to the high frequency information such as corners, edges, etc. to model the more discriminative descriptor. Another conclusion of this research is that the proposed descriptor is better performing over the large databases. The future scope of this study is the consideration of the color images on modeling of the descriptor.

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 Trans. Pattern Anal. Mach. Intell.*, 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, pp. 262–282, 2007.

3. S.R. Dubey, S.K. Singh, and R.K. Singh, "Rotation and scale invariant hybrid image descriptor and retrieval," *Computers & Electrical Engineering*, 2015.

4. J. Sun, G. Fan and X. Wu, "New local edge binary patterns for image retrieval," In *Proceedings of the 20th IEEE International Conference on Image Processing*, pp. 4014-4018, 2013.

5. S.R. Dubey, S.K. Singh, and R.K. Singh, "Local neighbourhood-based robust colour occurrence descriptor for colour image retrieval," *IET Image Processing*, vol. 9, no. 7, pp. 578-586, 2015.

6. S. Murala, R.P. Maheshwari and R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2874-2886, 2012.

7. 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*, 2014.

8. S. Murala and QM J. Wu, "Spherical symmetric 3D local ternary patterns for natural, texture and biomedical image indexing and retrieval," *Neurocomputing*, vol. 149, pp. 1502-1514, 2015.

9. N. Singh, S.R. Dubey, P. Dixit, and J.P. Gupta, "Semantic image retrieval by combining color, texture and shape features," In *Proceedings of the International Conference on Computing Sciences*, pp. 116-120, 2012.

10. A. Hadid and G. Zhao, "*Computer vision using local binary patterns*," Vol. 40. Springer, 2011.

11. T. Ojala, M. Pietikainen and T. Maenpaa, "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.

12. 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.

13. 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.

14. S. Zhao, Y. Gao and B. Zhang, "SOBEL-LBP," In *Proceedings of the 15th IEEE International Conference on Image Processing*, pp. 2144-2147, 2008.

15. K. Jeong, J. Choi, and G. Jang, "Semi-Local Structure Patterns for Robust Face Detection," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1400-1403, 2015.

16. A.P. James, "One-sample face recognition with local similarity decisions," *International Journal of Applied Pattern Recognition*, vol. 1, no. 1, pp. 61-80, 2013.

17. Corel Photo Collection Color Image Database, online available on http://wang.ist.psu.edu/docs/related.shtml.

18. The COREL Database for Content based Image Retrieval, online available on https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval.

19. MIT Vision and Modeling Group, Cambridge, 'Vision texture', http://vismod.media.mit.edu/pub/.

20. Salzburg Texture Image Database, http://www.wavelab.at/sources/STex/.

21. "www.cs.columbia.edu/~mmerler/project/code/pdist2.m".

22. 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.

23. 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.

24. 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.

25. "http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab".

26. 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.

27. 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*, 2015.