Content-Based Video Indexing and Retrieval using Block Based Local Binary Patterns and Pixel Change Ratio Map (BBLBPPCRM)

M.Ravinder [#1], Dr.T.Venugopal [*2]

[#] Research Scholar, CSE, JNTUK, Kakinada, AP, India

[1] ravinder.rsh@gmail.com

[*] Associate Professor, CSE, JNTUHCES, Sultanpur, Medak, Telangana, India

[2] t_vgopal@rediffmail.com

Abstract—This paper introduces a novel framework meant for content-based video indexing and retrieval using block-based local binary patterns, and pixel change ratio map (BBLBPPCRM). The proposed algorithm is tested on a video dataset of forty video, in which ten videos are air-plane type, ten videos are sailing boat type, ten videos are car type, and ten videos are war tank type. The performance of the algorithm is compared with existing methods, volume local binary patterns (VLBP), and pixel change ratio map (PCRM). The novel algorithm has shown reasonably good performance compares to existing methods.

Keyword-Block-based local binary patterns, Video, Indexing, Retrieval, Pixel change ratio map

I. INTRODUCTION

Nowadays, the availability of larger storage devices with low cost makes it possible to have huge video databases. A video contains huge information with raw data, and less or no predefined structure [1], which is difficult to index and retrieve videos from a huge video database. Manually indexing videos in a huge video database are a trivial task. In the past literature, to overcome the problem of manual indexing, a number of content-based video indexing and retrieval frameworks have been proposed

In a content-based video indexing and retrieval framework, usually the initial step is structural analysis of video, after that, the next step is extracting features, and representing the video with feature vector. The structural analysis divides a video in to a number of video shots. Then, a video shot can be represented with different techniques; one of them is key-frame based video shot representation.

In this paper, we have represented a video with spatial features of key frame, and pixel change ratio map based temporal features. A number of key frames based algorithms have been proposed in the prior literature. Some of the existing key frame based frameworks can be found in [1], [3-5], [7], and [8-9]. A new two steps based on the total content of the shot can be seen in [2]. A simple method proposed in [6], represents a video shot with the first frame as a key frame. In [7], the authors have suggested a method of selecting the starting, and ending frames of the video as key-frames. In [1], [3], and [4], the authors have introduced a threshold based key frame extraction methods.

The rest of the paper is organized as follows. Section 2 talks about related work. Section 3 projects the proposed framework. Section 4 discusses about experimental results, and section 5 concludes the paper.

II. RELATED WORK

This section talks about some of the existing frameworks, they are local binary patterns (LBP), block based local binary patterns, volume local binary patterns (VLBP), and pixel change ration map (PCRM) based motion histogram methods.

A. *Local Binary Patterns (LBP)*

Ojala et al. in [10], have proposed the popular and successful method of local binary patterns (LBP), which has been applied in finger print recognition, texture extraction, and face detection tasks. The local binary patterns method mainly relies on extracting pattern value for the centre pixel in a 3x3 patch of an image, based on neighbourhood pixel values as described in the equation (1).

$$\mathrm{LBP(C)} = \sum_{p=1}^{8} 2^{(p-1)} \times f(I(C) - I(N_p)) \quad (1)$$

Where $f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & otherwise \end{cases}$

An example computation of local binary pattern value is as shown in figure 1.

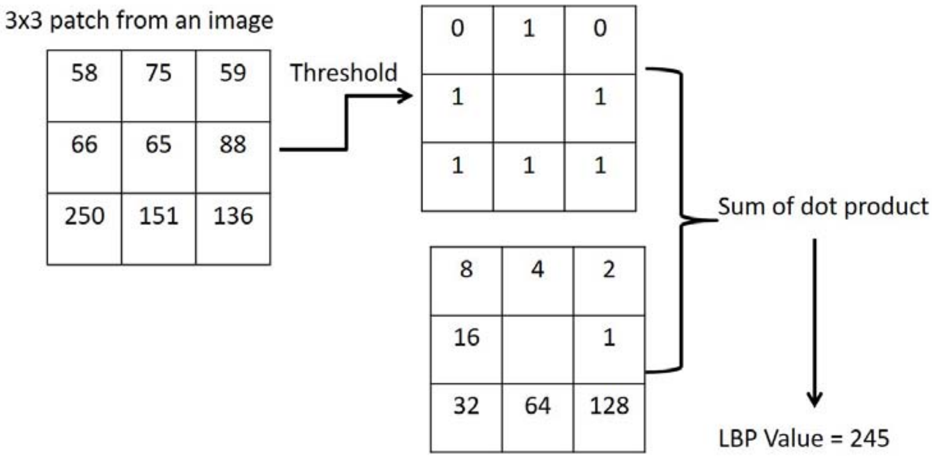


Figure.1. An example computation of local binary pattern value

B. Block-based Local Binary Patterns

Valtteri Takala et al. [11], have introduced block-based local binary patterns, in which instead of representing the whole image with the local binary patterns, the authors have divided the image in to blocks, after that, local binary patterns for individual blocks are calculated and summed as shown in below figure 2.

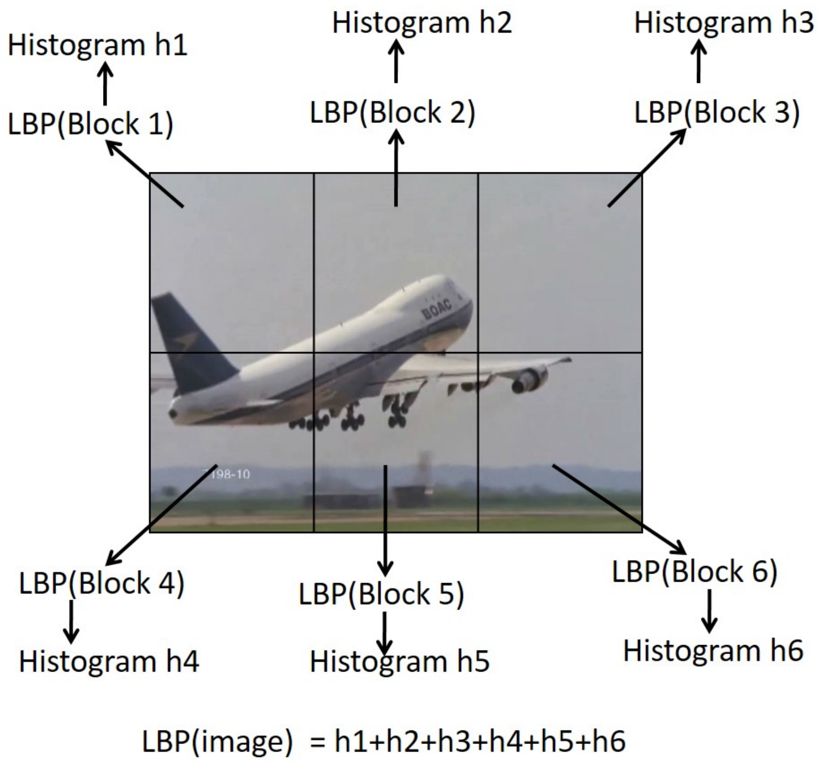


Figure.2. An example computation of block-based local binary patterns

In this paper, we have made a small modification to the original block based local binary patterns, instead of summing the local binary patterns of individual blocks as discussed above, we have concatenated the local binary patterns of the individual blocks.

C. Volume Local Binary Patterns (VLBP)

Guoying zhao et al. in [12], have extended the LBP method, and introduced Volume local binary patterns (VLBP) method, useful for extraction of temporal texture features. An example VLBP pattern value for a 3x3x3 cubic path of a group of three consecutive frames is as shown in figure 3.

The VLBP pattern values are computed by considering three consecutive frames from a video. Readers are redirected to [12] for the detailed information about how to extract VLBP from a video.

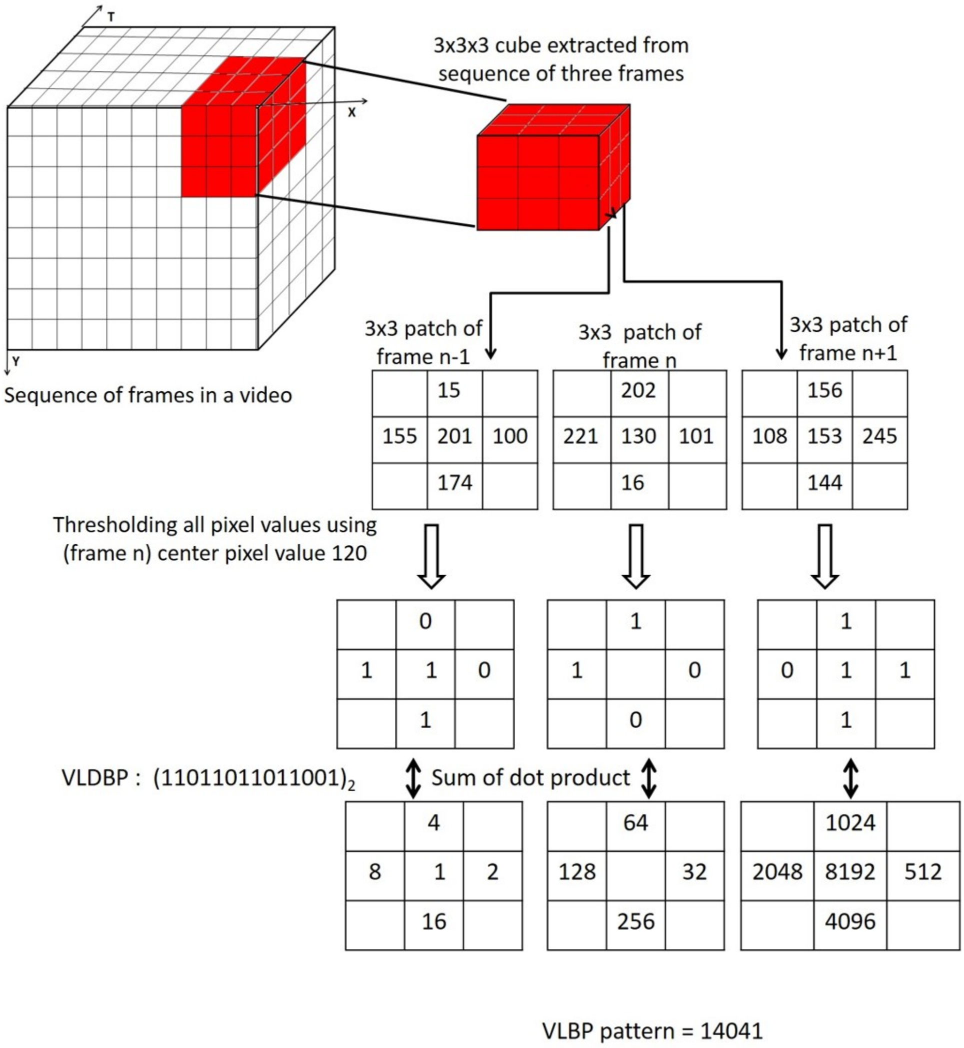


Figure.3. An example calculation of VLBP

D. Pixel Change Ratio Map (PCRM) based motion histogram

Haoran yi et al. have introduced a new motion feature extraction method in [13], known as pixel change ratio map (PCRM). PCRM is useful for applications like video retrieval, classification, and clustering. The procedure employed to extract PCRM from a video is, from each sequence of three frames fn-1, fn, and fn+1 , the summation of absolute, gray value difference of fn, fn-1 and fn+1, fn is computed as shown in equation (2).

$$D_i = | f_n - f_{n-1} | + | f_{n+1} - f_n | \quad (2)$$

PCRM is an array of size equivalent to size of a gray scale frame of the video. Initially, an empty array with zero values is created. Then, a group of three consecutive frames starting from the first frame of the video is considered. For example, a group of three frames (f1, f2, and f3) is selected. Then, the absolute difference between pixel values of (f2, f1), and (f3, f2) is computed, and subsequently summed as showed in equation (2).

Based on the Di pixel value, if a pixel value in Di is greater than a predefined threshold value (in our case it is ten), then the corresponding pixel value in PCRM array is increments by one. The same procedure is applied to all the groups of a sequence of three frames of the video. The resultant PCRM array is normalized using the value of the number of frames present in the video, and a thousand bins histogram is calculated, which results in PCRM based motion histogram.

III. PROPOSED FRAMEWORK

Pictorial representation of our proposed framework is as showed in figure 4. In our framework, the steps followed to extract spatial and temporal features are same, for all the videos in the video database, and the query video. Initial step is extraction of ten key-frames from the input video. Here we employ the algorithm 1, to extract key-frame from a video. The next step is, for the key frame extraction of block-based local binary patterns, which act as the spatial features of the video.

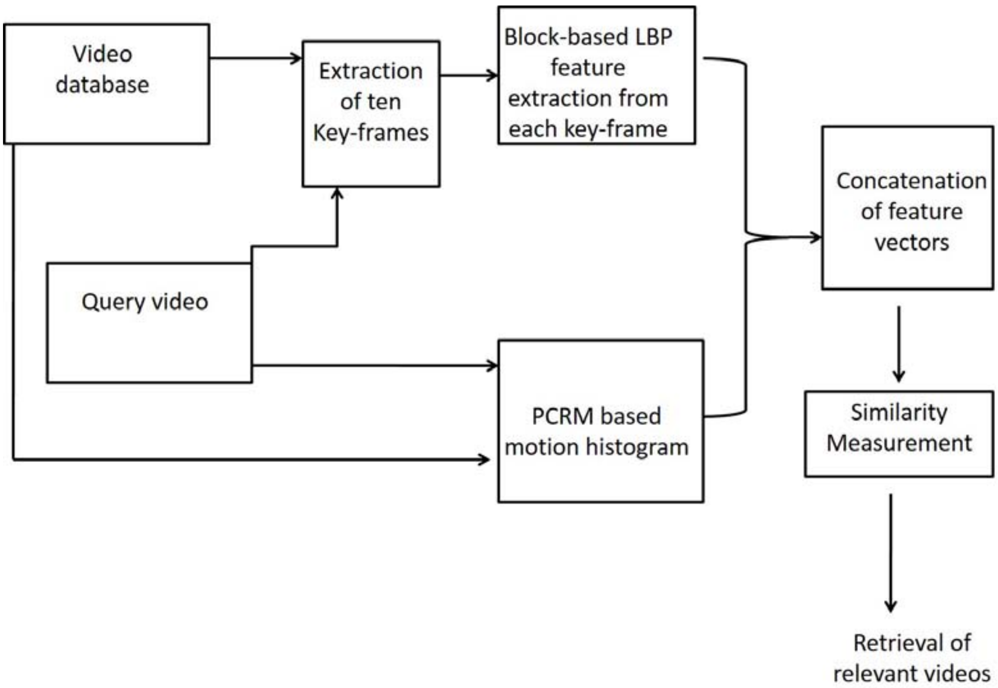


Figure.4. Proposed framework

In our framework, to extract temporal features we employ the PCRM based motion histogram method as discussed in section 2.4. After extracting motion features, the next step is concatenation of spatial and temporal features, which result in a feature vector of the video. The feature vector of the query video is compared with the feature vectors of the videos of the video database and relevant videos are extracted as output.

A. Algorithm 1: Key-frame extraction

Step1: For the input video, extract separate colour channel (R, G, and B) histograms from the first frame of the video.

Step2: Concatenate the histograms, of the frame which are extracted in step 1.

Step3: Apply step 1, and step 2 on all the frames of the video.

Step4: Find the mean histogram, from the histograms of all the frames of the video.

Step5: Extract ten key-frames based on k-nearest neighbour search method applied on histograms of individual frames, and mean histogram computed in step 4.

B. Algorithm 2: Proposed method - BBLBPPCRM

Step1: From the input video, extract ten key-frames using algorithm 1.

Step2: Extract block-based local binary patterns, and compute the feature histogram for the key-frame extracted in step 1.

Step3: Extract PCRM based motion features from the input video.

Step4: Concatenate the features extracted in step 2, and step 3, to formulate the feature vector of the video.

Step5: Apply the steps 1-5, on all the videos of the video database.

Step6: Apply the steps 1-5, on query video.

Step7: Compare the feature vector of the query video with the feature vectors of the videos in the video database using k-nearest neighbour search.

Step8: Retrieve the relevant videos.

IV. EXPERIMENTAL RESULTS

The proposed algorithm is applied on the dataset available at [14], which consists of ten air-plane videos, ten sailing boat videos, ten car videos, and ten war tank videos. The performance of the algorithms is compared with the existing methods, volume local binary patterns, and pixel change ratio map (PCRM) based motion histogram.

The retrieval performance is measured using precision and recall. The notations of precision and recall are as shown in equation (3) and (4) respectively.

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \tag{3}$$

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \tag{4}$$

The comparison of our proposed method results with already existing methods VLBP and PCRM based motion histogram are as shown in table 1 and table 2.

TABLE I. Precision (N=5) (%) (Top five videos retrieved)

Category	**VLBP**	**PCRM based motion histogram**	**Proposed Key-frame BBLBPPCRM**
Airplane	70	84	96
Sailing boat	50	76	76
Car	62	58	74
War tank	48	58	70
Average value	**57.5**	**69**	**79**

TABLE II. Recall (N=10) (%) (Top ten videos retrieved)

Category	**VLBP**	**PCRM based motion histogram**	**Proposed Key-frame BBLBPPCRM**
Airplane	52	72	85
Sailing boat	48	55	61
Car	56	56	64
War tank	40	43	46
Average value	**49**	**56.5**	**64**

Example Query video and its corresponding top five retrieved videos using our proposed method are as shown in below figure 5.

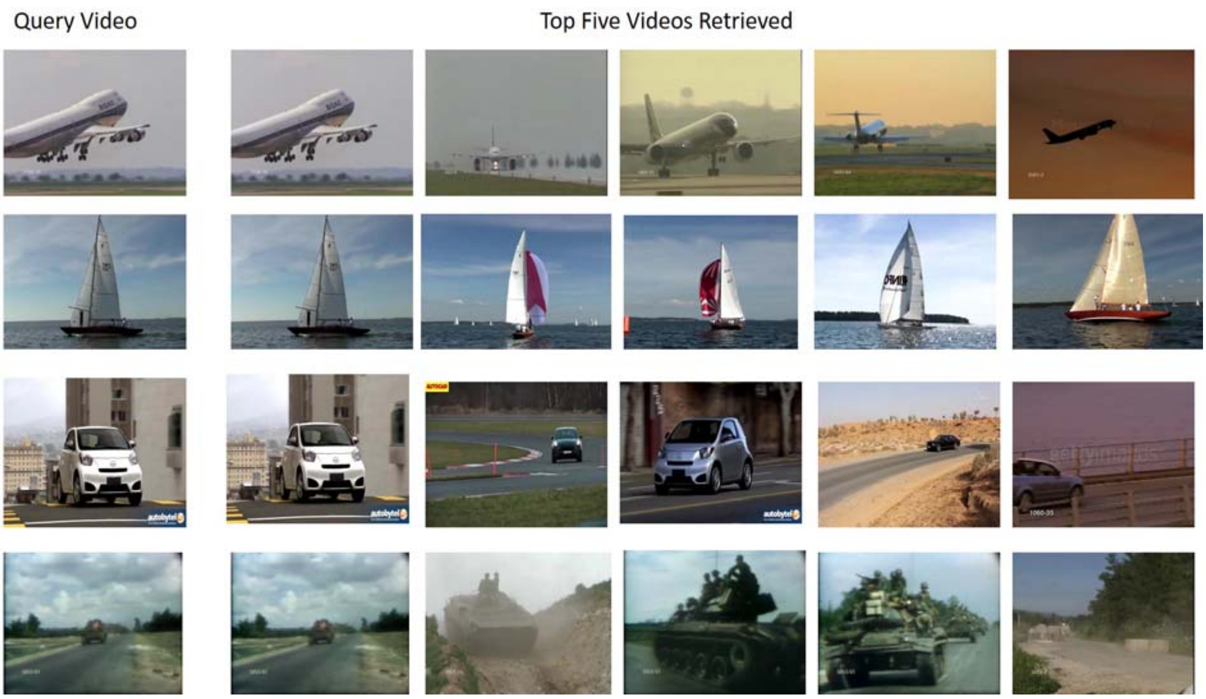
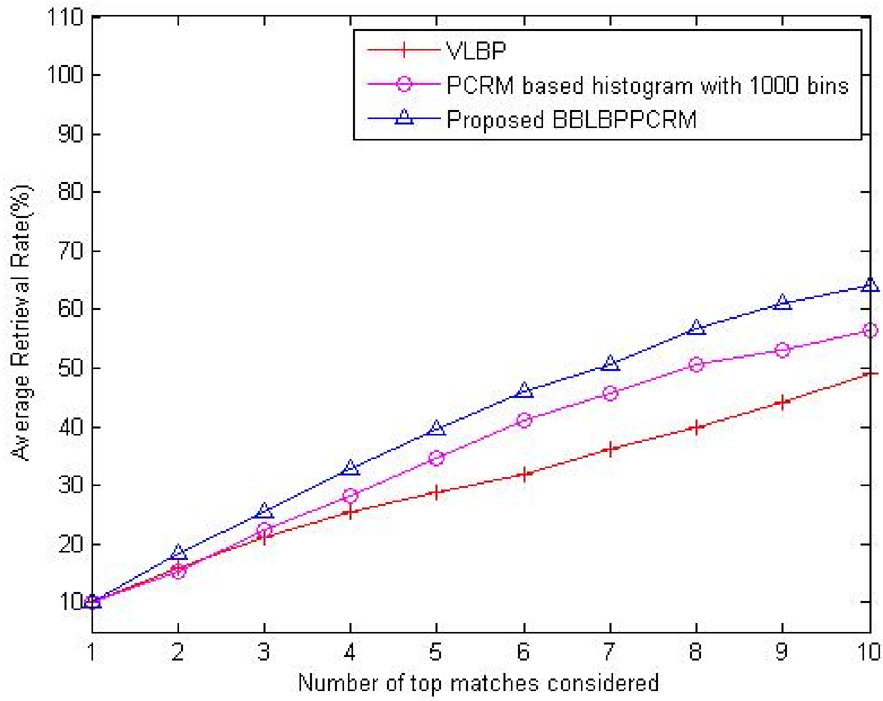
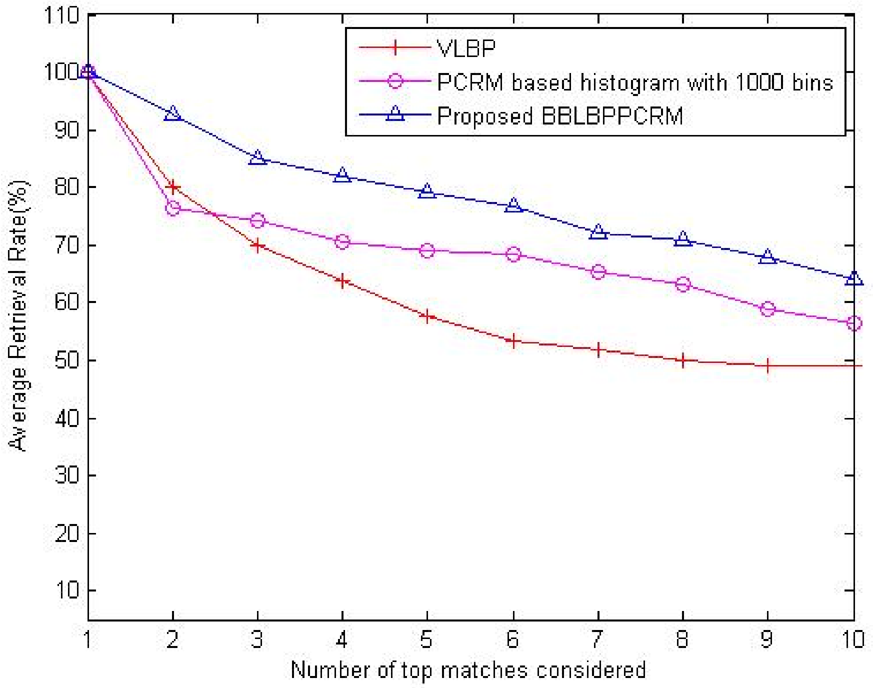


Figure.5. Query video and its corresponding top five videos retrieved

The comparison of our proposed algorithm, with existing VLBP and PCRM based motion histogram methods in terms of average precision and average retrieval rate is as showed in following graphs of figure 6 (a), and 6 (b).



(a)



(b)

Figure.6. Comparison of proposed method, and existing methods VLBP and PCRM based motion histogram, in terms of (a) average precision (b) average retrieval rate

V. CONCLUSION

In this paper, a novel framework meant for content based video indexing and retrieval has been put forward. To evaluate the performance of the algorithm, a video dataset of forty videos have been used. The dataset is publicly available at [14]. Our proposed method BBLBPPCRM has shown reasonably good results reported to existing methods.

REFERENCES

[1] Weiming Hu., Nianhua Xie., Li Li., Xianglin Zeng., and Stephen Maybank (2011), “A Survey on Visual Content-Based Video Indexing and Retrieval”, IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 41, No. 6.
[2] Alan Hanjalic.,Reginald L.Lagendijk.,and Jan Biemond,"A New Method for Key Frame based Video Content Representation".
[3] Yeung M.M., and Liu B(1995),"Efficient Matching and Clustering of Video Shots", Proceedings of ICIP, Vol.1,pp 338-341, Washington, D.C., USA.
[4] Yeo B., and Liu.B (1995),"Rapid Scene Analysis on Compressed Video",IEEE Transactions on Circuits and Systems for Video technology,Vol.5,No.6.
[5] Furth B., Smoliar S.W., and Zhang H (1995),"Video and Image Processing in Multimedia Systems",Kluwer Academic Publishers.
[6] Arman F., Hsu A., and Chiu M.-Y(1993), "Image Processing on Compressed Data for Large Video Databases", Proc.ACM Multimedia,'93, Annaheim, CA.
[7] Pentland et al.(1994),"Video and Image Semantics: Advanced Tools for Telecommunications",IEEE MultiMedia, Summer.
[8] Gerek O.N., and Altunbasak Y (1997),"Key Frame Slelection from MPEG Video Data", Proc.of SPIE Vol.3024, San Jose.
[9] Xiong., Ma R., and Lee J.C.-M (1997), "A Novel technic for Automatic Key Frame Computing", Proc.of SPIE, Vol.3022, San Jose.
[10] Ojala. T., Pietikäinen.M., and Harwood.D (1996), “A comparative study of texture measures with classification based on feature distributions”. Pattern Recognition 29, 51-59.
[11] Valtteri Takala.,Timo Ahonen., and Matti Pietikainen (2005),"Block-Based Methods for Image Retrieval Using Local Binary Patterns",SCIA 2005, LNCS 3540, pp. 882–891.
[12] Guoying Zhao., and Matti Pietikainen, “Dynamic Texture Recognition Using Volume Local Binary Patterns”.
[13] Haoran Yi., Deepu Rajan., and Liang-Tien Chia (2005), “A new motion histogram to index motion content in video segments”, Pattern Recognition Letters 26, 1221–1231.
[14] https://sites.google.com/site/benchmarkvideodata/dataset

AUTHOR PROFILE

M.Ravinder has completed his B.Tech, M.Tech in CSE in 2003 and 2009 respectively. He has more than eight years of teaching experience and currently pursuing PhD in CSE from JNTUK, Kakinada, AP, India.

Dr.T.Venugopal has completed his M.Tech, PhD in CSE in 2003 and 2009 respectively. He has more than seventeen years of teaching experience and currently he is working as Associate Professor in CSE at JNTUHCES, Sultanpur, Medak, Telangana, India.