Analysis of Volume Local Binary Patterns for Video based Smoke Detection

Gaohua Lin, Qixing Zhang, Gao Xu, Lan Qian, Yongming Zhang
State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei, Anhui 230026, PR China

Abstract

A video based smoke detection method using dynamic texture feature extraction with volume local binary patterns is studied. Dynamic texture features were extracted from blocks and irregular motion regions, respectively. Different operators were used to extract dynamic texture for smoke detection to study their characteristics.

The results show that dynamic texture is a reliable clue for video based smoke detection. Irregular motion regions based method reduces adverse impacts of block size and motion area ratio threshold. It is generally conducive to reducing the false alarm rate by increasing the dimension of the feature vector. Additionally, it is found that the feature computing time is not directly related to the vector dimension, which is important for the realization of real-time detection.

Keywords: Smoke detection, Video sequences, Volume local binary pattern, Dynamic texture, Support vector machine

Introduction

Compared to conventional point smoke detector, video based fire detection system shows advantages in being usable in large open spaces, detecting fire immediately, providing more information such as the fire development and location. Smoke often emerges before flames, and is a more efficient clue for early fire detection. Video based smoke detection methods distinguish smoke from non-smoke objects based on some distinctive features such as motion, edge, color and texture [1].

Local binary pattern (LBP) is one of the most prominent methods in the field of texture analysis [2]. LBP based smoke detection methods have been studied, but the previous studies have focused on image based local feature extraction. Volume local binary pattern (VLBP) [3] can combine spatial and temporal features of smoke. Fig. 1 shows the entire computing procedure for $VLBP_{1,4,1}$.

The LBPTOP is a simplified VLBP calculating the binary number separately for three orthogonal planes around the center pixels. The CVLBP adds a global texture feature that uses center pixel information combined with a global mean difference as a threshold to modify [4].

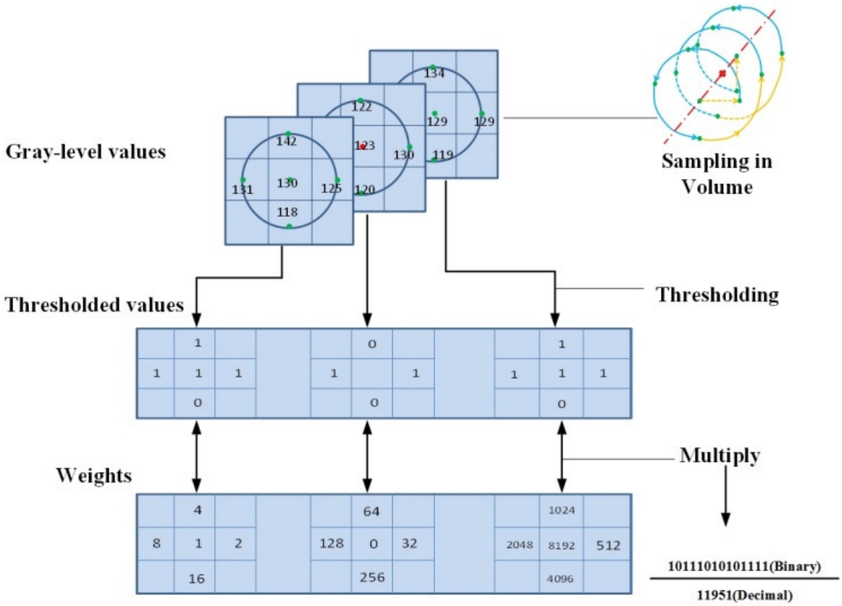

Fig. 1. Computing procedure for $VLBP_{1,4,1}$ [3].

Block based dynamic texture extraction

HD video network cameras were used in our study to shoot a group of smoke videos and non-smoke videos with a size of 1920 x 1080. The videos can be downloaded on our website (http://smoke.ustc.edu.cn/). The videos used for training consisted of 10 smoke videos and 5 non-smoke videos named Video 1 - 15. The testing videos were named Video A - O.

We divided the video image into 100×100 non-overlapping blocks. A total of 4,805 smoke blocks and 11,842 non-smoke blocks were labeled in video samples 1 - 15. Fig. 2 is a schematic of some samples. Then the $VLBP_{1,8,1}$ operator was used to extract the dynamic texture features of each block. A total of 2,366 smoke blocks and 5,989 non-smoke blocks were randomly selected from the labeled blocks and used as the training data for the SVM.

Firstly, a block was considered as the object to conduct the test. The results are shown in Fig. 3 for the detection rate (DR) and false alarm rate (FAR). Some smoke blocks in diffuse white smoke Videos 6 and 7 were identified mistakenly, and some non-smoke blocks in Video 6 that contained extremely thin smoke caused a high FAR.

Video 8 contained a thin black smoke plume, and the blocks could not frame out the smoke perfectly, thus the detection performance was the worst in this case.

Fig. 2. Positive samples (left) and negative samples (right).

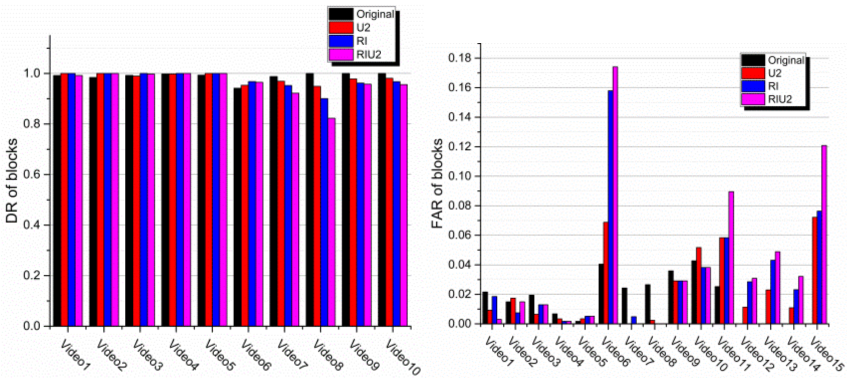

Fig. 3. DR and FAR of smoke block detection.

After the blocks were classified, the next procedure was to determine the category of each whole frame. We used the neighboring block rule [5] to issue a smoke alarm based on a frame.

For testing video A - O, we first extracted a motion region using background subtraction method and selected the candidate blocks from the extracted region subsequently. As the detection result is sensitive to motion area ratio threshold, we set the thresholds as 0, 0.1, 0.2, 0.3, 0.4 and 0.5 to select the candidate block. A motion area ratio of zero means that all the blocks in the frame are treated as the candidate blocks.

The total DR and FAR of frames for testing video A - O aganist different thresholds are shown in Fig. 4. Obviously, a low threshold causes more blocks recognized as candidate smoke blocks which resulted in high DR and FAR of frames. The high FAR will decrease with increase in the motion area ratio accompanied by a lower DR simultaneously.

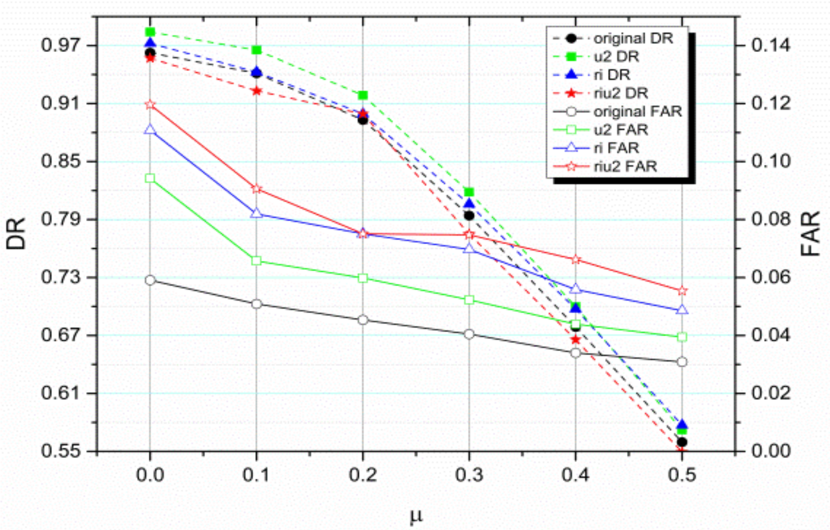

Fig. 4. DR and FAR of frames based on blocks.

Compared to the smoke blocks with relatively consistent dynamic texture features, the non-smoke blocks had a variety of feature histograms, with the result that the larger the feature dimension is, the more non-smoke blocks could be excluded. Thus, it can be found that the original mode had the lowest FAR compared with the other three modes in Fig. 4 distinctly.

Dynamic texture extraction based on irregular regions

Both the motion area ratio threshold and frame alarm rule have a great influence on the frame DR and FAR. What even worse is that as HD cameras have wider view coverage and greater monitoring distance, the smoke region in the HD video frame have a large variable range. It is very difficult to determine a reasonable block size. We proposed a dynamic texture feature extraction method based on irregular region to avoid these problems. The moving regions with irregular shape in a frame are obtained using the same motion region extraction method as before. Then, the irregular motion regions in a frame are treated as an entire target, and dynamic texture was only obtained from these regions. The procedure is shown in Fig. 5.

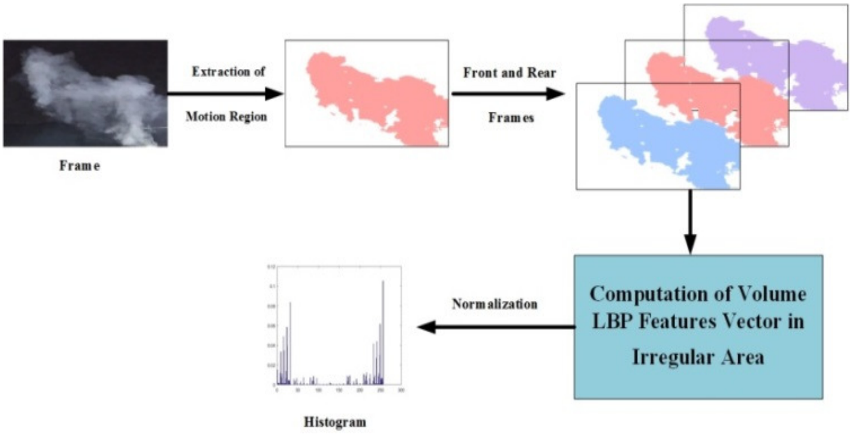

Fig. 5. Extraction of the VLBP texture features from irregular regions.

We also extracted the $VLBP_{1,8,1}$ dynamic texture from irregular regions of the training frames in Video 1–15 and testing frames in Video A - O to conduct a comparison with the blocks based method. SVM was used to train and test the feature vectors of those frames and the result is shown in Fig. 6. It can be observed that the regions based method could greatly reduce the FAR while keeping the DR at a high level at the same time. Additionally, it is conducive to reduce calculation as the regions based method only extracts texture features from motion regions strictly.

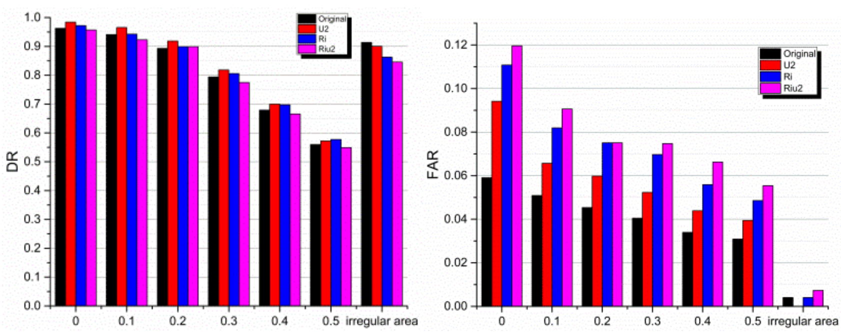

Fig.6. DR and FAR of the irregular regions based method compared with the blocks based method.

Characteristics of volume local binary pattern operators

To study the characteristics of volume local binary pattern, we used the LBPTOP, VLBP, CLBPTOP and CVLBP operators to extract dynamic texture for smoke detection respectively. Fig. 7 shows some parameters of some operators.

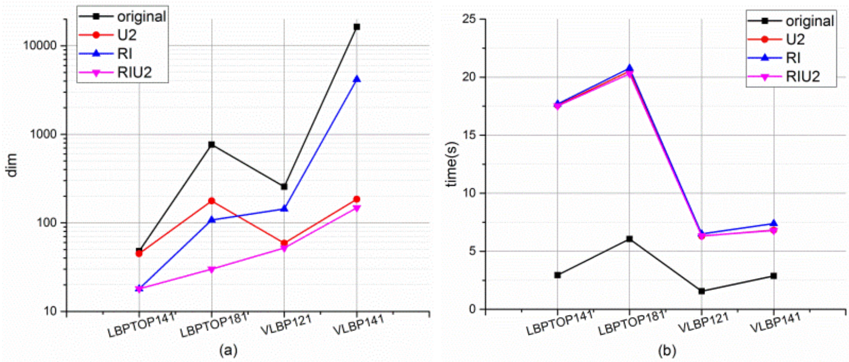

Fig. 7. (a) dimension of the feature vector for each operators, (b) the time for calculating the feature vector for an identical video.

It can be find that the computation time is not directly related to the feature dimension, for example, the original dimension of $LBPTOP_{1,8,1}$ is 768 and the calculation time is 6.06 seconds; however, when the original dimension of $VLBP_{1,4,1}$ reaches up to 16384, the calculation is only 2.88 seconds. The U2, Ri and Riu2 modes greatly reduce the feature dimension compared with the original mode. However, both of them are calculated based on the original one. Thus, the calculation time is obviously longer. In fact, the feature dimension of the operator only directly affects the training time of the SVM, which is greatly shorter than the time required for feature vector computation.

The LBPTOP is a simplification of the VLBP with the purpose of reducing the feature dimension and the computational complexity. Nevertheless, as previously described, the feature dimension does not directly affect the time spent on feature extraction. Fig. 8 shows the detection results of the LBPTOP and VLBP operators. On increasing the number of sample points P, FAR for both the LBPTOP and VLBP decreased sharply. The DR for the LBPTOP with P = 8 was higher than that with P = 4. However, increasing the number of sample points P did not improve the DR for the VLBP operator.

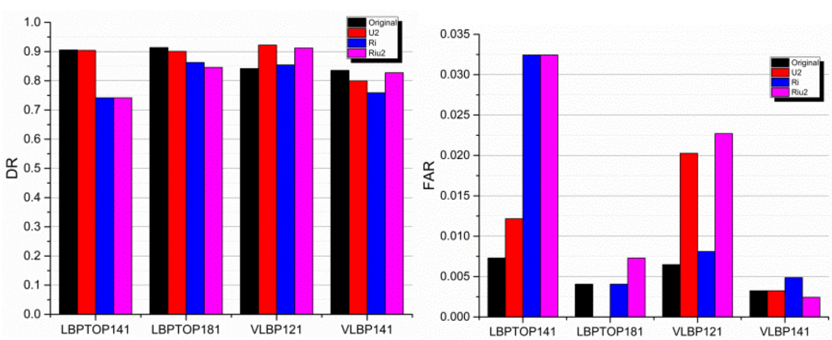

Fig. 8. DR and FAR of the LBPTOP and VLBP

In [5], the volume block for VLBP dynamic texture extraction contained 64 frames. Given that the time and storage space required for the calculation increase with the number of frames, it is difficult to achieve a real-time detection. In this paper, we used only 3 frames to extract dynamic texture with irregular regions. A comparison between the detection results based on 3 frames and 16 frames is shown in Fig. 9. It is obvious that more frames result in a better detection capacity with higher DR and lower FAR. Meanwhile, more frames means more calculation. Thus, the detection requirements and hardware performance should be considered synthetically to determine the optimal number of frames for dynamic texture extraction.

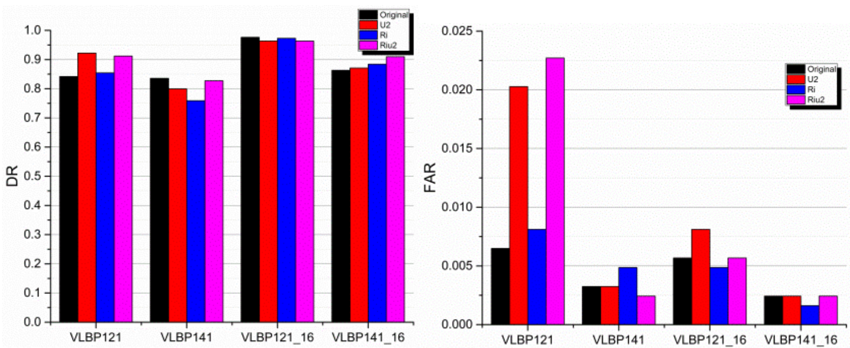

Fig. 9. DR and FAR based on 3 frames and 16 frames.

Fig. 10 show the results of CVLBP compared with VLBP. It indicated that adding magnitude features could exclude more false alarms. However, for the smoke frames, CVLBP did not always demonstrate better recognition ability for all modes. Only the DR for the four modes of $VLBP_{1,2,1}$ were promoted visibly.

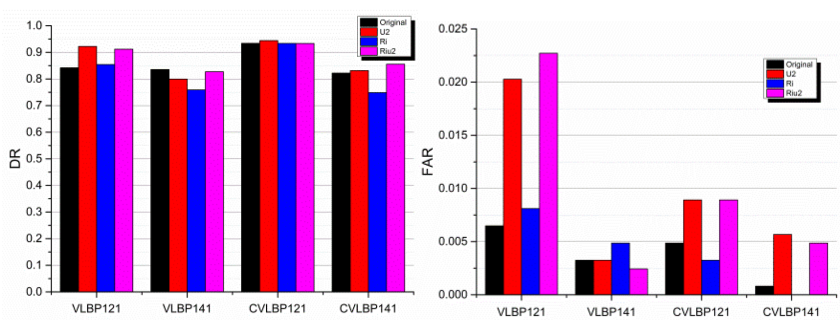

Fig. 10. DR and FAR of the VLBP and CVLBP.

The movement characteristics of smoke are an important feature to distinguish it from non-smoke objects. Smoke detection method based on dynamic texture extracts movement characteristics by extending the image texture to the spatiotemporal domain. As the motion information of smoke is reflected in the difference between the frames, the frame interval L influences dynamic texture extraction. We used L = 2 for all operators mentioned above to conduct the training and testing. The DR and FAR results are shown in Fig. 11. We can observe that the general trends of FAR for the two conditions are similar, whereas the amplitudes for L = 1 are smaller in most cases. For the DR of the smoke frames, neither L = 1 nor L = 2 has obvious advantages over the other. The optimal frame interval value is depends on the operator.

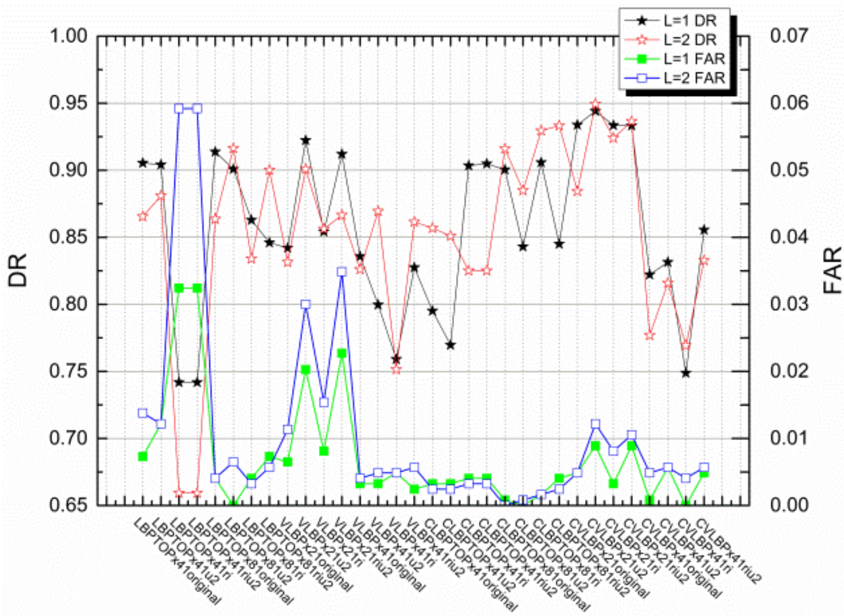

Fig. 12. DR and FAR of frame interval 1 and frame interval 2.

Conclusion

Using dynamic texture features to detect smoke was an effective method, especially for clear white smoke plume with an extremely high DR. Diffusion and thinning of smoke reduced the texture discrimination of smoke.

HD videos provide more details for smoke detection, but the size of smoke in HD videos also become more changeable. Thus It is very difficult to determine a reasonable block size. The threshold of the candidate smoke block also greatly affected the smoke DR and FAR. A method based on irregular regions to extract the dynamic texture feature from motion areas directly was used, which ensured a relatively high DR and greatly reduced the FAR.

Generally, as the non-smoke samples had various texture distributions, an operator with high feature dimension (achieved by increasing the number of sample points or increasing the complexity of the method) was more conducive to excluding non-smoke samples, thereby reducing the FAR.

Additionally, the feature extraction computing time of an operator was not directly related to the size of the vector dimension. In fact, the Original and U2 modes were more suitable for real-time smoke detection in most cases.

References

[1] Cetin, Enis A, Gunay, et al. *Video fire detection – Review[J]*. Digital Signal Processing, 2013, 23(6):1827-1843.

[2] Ojala T, Pietikäinen M, Mäenpää T*. Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns[J]*. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7):971-987.

[3] Zhao G, Pietikinen M. *Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions[J]*. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29(6):915-28.

[4] Tiwari D, Tyagi V. *Dynamic texture recognition based on completed volume local binary pattern[J]*. Multidimensional Systems & Signal Processing, 2015, 27(2):1-13.

[5] Tian H, Li W, Ogunbona P, et al. *Smoke detection in videos using Non-Redundant Local Binary Pattern-based features.[C]*, IEEE, International Workshop on Multimedia Signal Processing. 2011:1-4.