Target-Tracking Based Early Fire Smoke Detection in Video

Zheng Wei, Xingang Wang, Wenchuan An, Jianfeng Che
Institute of Automation of the Chinese Academy of Sciences
95 Zhongguancun East Road, 100190, BEIJING, CHINA
zheng.wei@ia.ac.cn, xingang.wang@ia.ac.cn, an.wenchuan@gmail.com, jfche.casia@hotmail.com

Abstract

The paper proposed a target-tracking based fire smoke detection method for early fire-alarming system at large or open space. The method utilizes an improved Gaussian mixture model positioning algorithm, an efficient target tracking algorithm as well as three effective static and dynamic smoke visual features: brightness consistency, motion accumulation and spread. Finally, an algorithm combined temporal and spatial information to assess the fire alarm algorithm is implemented by considering the performance requirements of the fire-alarming system. Experimental results show that the method has low time complexity and is able to rule out the major interference sources.

1. Introduction

Fire detection and alarm technology in video, which makes use of cameras as probe, the scene of the image information and digital image processing technology, identify fire flame and smoke and then determine whether the fire occurred. "Smoke is the fire's beginning." It is necessary to detect the smoke of the early stages of fire for more sensitive and credible of smoke detection. Since the technology above can be applied to early warning and outdoor fire detection such as large space, it gets more and more attention in recent years and obtains good results in particular experimental environment.

However, in practical applications, due to fire derivatives variability and complexity of the scenes, most of fire smoke detection methods [1, 2, 3, 4, 7, 9] have a relatively high false alarm rate and lack enough robustness and adaptability.

To reduce false alarm rate, we have carried out detailed analysis of the fire smoke image information, summarized most of the possible major interference sources and done a classification. According to the characteristics of various types of interference sources in combination with fire smoke image information, we find that brightness consistency, motion accumulation and spread are effective static and dynamic features in eliminating interference. According to the experimental results above, we propose a target-tracking based early fire smoke detection method. This method considers both time and space complexity, which meets the accuracy and speed needs of practical engineering requirement. In the following, we will present more detail of our detection method.

2. Fire smoke detection

Above all, our method using ordinary color cameras assumes that the stationary camera monitoring the scene is based on YUV color model. Smoke detection algorithm consists of four steps: (1) The moving pixels or regions in the current frame of a video are segmented; (2) The connected regions are extracted and marked corresponding number; (3) By tracking each connected region, static and dynamic features are extracted to match fire smoke features criterions; (4) Fire alarm rule, an algorithm, combined temporal and spatial information is used to decide the fire alarm.

2.1. Moving regions segmentation

Segmentation of moving regions in video is a fundamental and key step in fire smoke detection. A typical method is background subtraction. Background subtraction involves calculating a reference image, subtracting each new frame from this image and threshold the result.

Although many papers have been referred to for introducing various background subtraction algorithm, the problem, moving regions segmentation in complex environment, was not solved completely. This is mainly due to both the indoor and outdoor environment. At any time, a lot of changes exist, while these changes may affect only part of the background or the whole background. It needs the compared background image updated to adapt to these changes. Grimson et al and P. KaewTraKulPong et al employed an adaptive nonparametric Gaussian mixture model to solve these problems [5, 6]. Their model can cope well with the illumination changes and lessen the effect of small repetitive motions; for example, moving vegetation likes trees

978-0-7695-3883-9/09 $26.00 © 2009 IEEE
DOI 10.1109/ICIG.2009.173

and bushes as well as small camera displacement. In our method we use the background estimation method developed by P.KaewTraKulPong et al[6].

Each pixel in the scene is modeled by a mixture of Gaussian distributions using YUV color model. The probability that a certain pixel has a value of at time N can be written as

$$p(x_N) = \sum_{j=1}^{K} w_j \eta(x_N; \theta_j) \tag{1}$$

$$\begin{aligned}\eta(x;\theta_k) &= \eta\Big(x; u_k, \sum_k\Big) \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \left|\sum_k\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \sum_k^{-1}(x-\mu_k)}\end{aligned} \tag{2}$$

where w_j is the weight of the k^{th} Gaussian, $\eta(x_N;\theta_j)$ is the distribution of the k^{th} Gaussian, μ_k is the mean and $\sum_k = \sigma_k^2 I$ is the covariance of the k^{th} component. The K Gaussians are ordered by the fitness value of $\frac{w_k}{\sigma_k}$ and the first B distributions are used as a model of the background of the scene where B is estimated as:

$$B = arg \min_b \left(\sum_{j=1}^{b} w_j > T \right) \tag{3}$$

the threshold T is the minimum fraction of the background model. In other words, it is the minimum prior probability that the background is in the scene. Background subtraction is performed by marking a foreground pixel any pixel that is more than 2.5 standard deviations away from any of the B distributions. The first Gaussian component that matches the test value will be updated by the update equations using L-recent window version.

2.2. Connected regions extraction

The results of moving regions segmentation may include many connected regions. It is necessary to analyze these connected regions further to determine if the motion is due to smoke or an ordinary moving object. For detecting fire smoke, each connected region needs to be extracted and tracked in order to match static and dynamic features of fire smoke in every video frame. For this purpose, we developed an extraction and marking algorithm based on tree structure. The input is moving regions segmentation result, a binary image, and the output is marked by a number for each region. Basically, the extraction and marking algorithm consists of three steps realized by tree structure:

1. Initialization. A unique label number is given to each connected region in one row, and record the region's relation with its upper regions. Each entire connected region is described as a tree.

2. Linking and merging. Merge all the connected regions and give new labels, by analyzing the region relations recorded in step 1. In this process, a connected region may be linked to two or more upper regions, which means that through this area more than one upper region can be connected as one region. According to the tree theory, we connect one tree's root node to another's.

3. Rescanning and marking. Rescan the binary image, and replace all the labels by given final labels. The actual extraction and marking is obtained by using the tree structure of the created links. All the regions that are connected to a common root are classified as a single region, which labeled a distinct number.

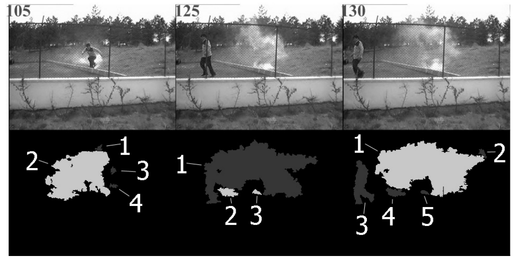


Figure 1: Segmentation and extraction results.

Figure 1 shows the results of segmentation and extraction from a video sequence in which a person is moving away from the outdoor scene. The top three images show the original video frames, the number in top left corner presents the frame number in the video, similarly hereinafter, and the bottom images show the results, using different gray scales for distinction.

2.3. Fire smoke visual features

The smoke visual features we used for smoke detection include the static and dynamic features, brightness consistency, motion accumulation [10] and spread features [8].

The brightness consistency feature in a smoke region is a good indicator. We found the smoke usually displays self-similarity in brightness. By the experimental statistics, the brightness variance of smoke area maintains at around 900, standing for the changes in pixel brightness values ranging around 30.

In point of dynamic features, we investigate motion accumulation and spread features.

Smoke often emerges continually from the place of catching up. An accumulation model is presented to extract these dynamic visual features of early fire over a time window. The model synthesizes information from several frames, so it can mostly suppress noises.

The fire smoke in the early stage is generated around the fuel and does not produce significant overall movement in a short period. From the timeline, the moving target in smoke region has a much larger possibility to have value one in the extracted binary image. In order to measure this feature, we set up a motion accumulation gray image by accumulating the image in a time window has a length of W. The following formula is used to get this gray image:

$$A(x,y,t) = \sum_{i=0}^{W-1} M(x,y,i) \tag{4}$$

where $A(x,y,z)$ represents the motion accumulation image and $M(x,y,z)$ represents the extracted binary image at time t. In our method, the motion accumulation gray image pixels are between $[0, W \times fps]$ and threshold $T_A = 0.75W \times fps$, the motion accumulation feature can be extracted by estimating if the pixel value in $A(x,y,z)$ is greater than T_A as follows:

$$L(x,y,t) = \begin{cases} 1 & A(x,y,t > T_A) \\ 0 & else \end{cases} \tag{5}$$

Figure 2 shows the result of motion accumulation feature extracted from two video. Subfigure 2a shows a person moving slowly and subfigure 2b shows the scene of smoke spreading. In each subfigure, the first two images are the beginning and the ending frames of the video sequence and the third image is the motion accumulation result, gray images.

The spread feature is also a crucial dynamic feature for enhancing the reliability and reducing the false alarm rate of smoke detection. Due to the existence of diffusion in smoke generation process, the smoke region in the video will gradually increase. The increment rate of the smoke region is defined as

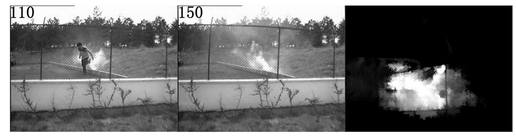
$$V_{area} = \frac{N_t - N_{t-\Delta t}}{\Delta t} \tag{6}$$

where V_{area} is the smoke area increment rate at the interval between time t and $t - \Delta t$, in addition N_t and $N_{t-\Delta t}$ represent the pixel quantity of the smoke area and n represents the number of frames in Δt. Considering that during the growth process, the smoke area cannot change much quickly or remain exactly the same, so there are low-bound and high-bound thresholds of growth rate, respectively like this $V_{min} < V_{area} < V_{max}$.

The thresholds of V_{min} and V_{max} are determined by the statistical data of experiments. We choose a typical video sequence including smoke spreading and disappearing. The smoke spreading and disappearing process in this video approaches the boundary condition in our experiments, so we show the statistical data results of the smoke region in this video in Figure 3 for clarity.

As shown in Figure 3, in smoke spreading process the area of smoke grows as time increases(e.g. 0~10s). We choose the upper and lower bound of the area increase rate

(a)



(b)

Figure 2: Motion accumulation results.

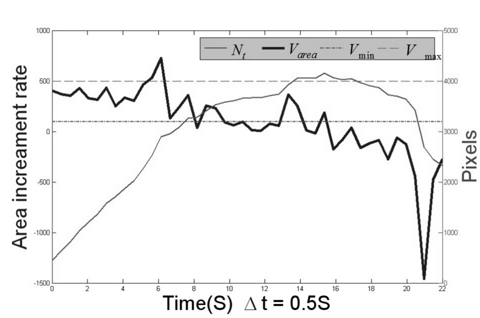


Figure 3: Spread feature statistical results.

as the V_{min} and V_{max} separately after we removes the impulses on the area increasing rate curve using filtering algorithm such as average filtering.

2.4. Target tracking based smoke detection

According to the static and dynamic features above, this paper proposes a target-tracking based smoke detection method, a robust and broad applicability smoke detection method provides an early alarm. To be more specific, track the connected region by a simple matching algorithm based on overlapped region area, extract the features including brightness consistency, motion accumulation and spread, and decide if this region containing fire smoke. Figure 4 shows the flow diagram of the smoke detection method above.

The detail algorithm of connected regions matching and features extracting is as follows:

1. Detect the connected regions of the current frame every $\Delta t (\Delta t = 0.5s)$ interval.

2. Compute the overlapping area of the current connected regions $fd(x,y,t)$ and the previous connected regions $fd(x,y,t-\Delta t)$.

3. Sort all the connected regions in descending order based on overlapping area.

4. Traverse from the largest overlapping area regions

(a) If neither of the overlapping regions has been marked, mark the regions as matched.

(b) Turn to the next regions, and repeat (a) until all the regions have been matched or no overlapping area regions left.

5. For each pair of matched regions

$$fd(x, y, t - \Delta t) \to fd(x, y, t)$$

(a) If the previous region is not marked as suspicious smoke region, while the current region satisfies the brightness consistency and spread features, set $A(x, y, t) = A(x, y, t - \Delta t) + 1$. If $A(x, y, t) > T_A$, mark the current region as suspicious smoke region.

(b) If the previous region is already marked as suspicious smoke region, lower the threshold of V_{min} and increase the threshold of brightness variance. At the same time, if the current region satisfies the brightness consistency and spread features, set $A(x, y, t) = A(x, y, t - \Delta t) + 1$. If $A(x, y, t) > T_A$, mark the current region as suspicious smoke region.

6. Mark all of the non-match regions as non-suspicious smoke regions

7. Repeat step 1.

2.5. Fire alarm rule

The ultimate goal of the smoke detection is to be used in reality. Considering the accuracy, real-time process and low false alarm rate requirements of practical application, we proposes an alarm assessing algorithm based on time and space statistical model. Firstly, divide the suspicious regions into sub-block regions of size 8×8. And compute the area of each sub-block regions. In digital image processing, the area can be represented by the pixel quantities. To the binary image, pixels valued 1 all together stand for the sub-block region area as follow,

$$A_i^t = \sum_{(x,y,t) \in R_i} D(x, y, t) \tag{7}$$

where $D(x, y, t)$ represents the binary image of suspicious region, $R_1, R_2, \cdots, R_i$ the sub-block regions and A_i their area. If the sub-block region area is greater than the threshold T_A, set the decision result $S = 1$, as follow,

$$S = \cup_i (A_i^t > T_A) \tag{8}$$

T_A is directly proportional to the size of the sub-block region, in our method $T_A = 8 \times 8 \times 25\%$. To reduce false alarm further, those fire decision with short duration $(S = 1)$ is regard as noise. So we set a time window length of W, $W = 10s$ in our method, and a threshold

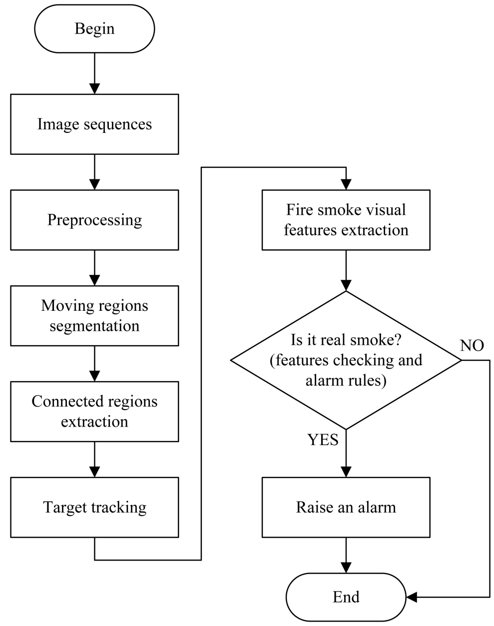


Figure 4: Smoke detection method.

$T_S = W \times 50\%$, if

$$S_t = \begin{cases} 1 & if \sum_{t^t=0}^{w} S_{t^t} > T_S \\ 0 & else \end{cases} \tag{9}$$

confirm that the sub-block contains fire smoke pixels and a smoke alarm is given.

3. Experimental results

The proposed fire smoke detection method is implemented on a PC with an Intel Core 2 1.80GHz processor and tested for a large variety of real-time videos containing only smoke and videos with different interference source. The processing time per frame is about 45 MSEL for frames of size 320 by 240 pixels. We choose six typical videos including various types of interference source in outdoor such as snow, shaking branch, walking people and car lights etc. Detection results for the test videos are presented in Table 1. Smoke is successfully detected in all of the videos containing smoke and no false alarms are issued. For contrast, we do the same experiment on method 2 proposed in [4] and method 3 proposed in [8]. The results are listed in Table 1. Two false alarms occurred in method 2 for the invalidation of wavelet method which is the key point in method 1; method 3 initialed 3 false alarm in video 4-6, for mostly focusing the dynamic features.

Figure 5 shows two videos with smoke regions marked. The top video is a snow scene with slight shaking leaves. The bottom video shows a person moving out of the scene.

No.	FSA	FSD	Method 1	Method 2	Method 3	Description
1	167	192	Y	Y	Y	Snow scene with moving person
2	128	150	Y	Y	Y	Branch shaking and automobile moving in the distance view
3	190	246	Y	Y	Y	Smoke generated in the distance view
4	0	0	N	Y	Y	Car light
5	0	0	N	Y	Y	Person moving wearing grayish clothes
6	0	0	N	N	Y	Flame

Table 1: Detection results of the test videos. FSA is short for frame of smoke appears, FSD is short for frame of smoke detected in our method and Method 1 is our method.

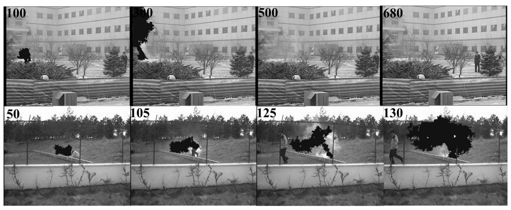


Figure 5: Experimental results.

In these two video sequences, the smoke is detected successfully with high sensitivity ruling out all interference sources by our method.

4. Conclusion

A novel target tracking based smoke detection method is proposed to provide an early alarm for the fire accidents. Static and dynamic features of smoke along with an alarm-assessing algorithm are involved into the decision to improve the reliability of smoke detection in video.

Experiments show that this method can provide a reliable and cost-effective solution for smoke detection. In addition, the tracking based method can be used for multi-source fire detection by a little bit modification.

References

[1] A. E. Cetin and R. Ansari. Signal recovery from wavelet transform maxima. 42(1):194–196, Jan. 1994. 1

[2] T.-H. Chen, C.-L. Kao, and S.-M. Chang. An intelligent real-time fire-detection method based on video processing. In *Proc. IEEE 37th Annual 2003 International Carnahan Conference on Security Technology*, pages 104–111, Oct. 14–16, 2003. 1

[3] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou. An early fire-detection method based on image processing. In *Proc. International Conference on Image Processing ICIP '04*, volume 3, pages 1707–1710, Oct. 24–27, 2004. 1

[4] N. Dedeoglu, B. U. Toreyin, U. Gudukbay, and A. E. Cetin. Real-time fire and flame detection in video. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, volume 2, pages 669–672, Mar. 2005. 1, 4

[5] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee. Using adaptive tracking to classify and monitor activities in a site. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 22–29, June 23–25, 1998. 1

[6] P. KaewTraKulPong and R. Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. *Video-based Surveillance Systems: Computer Vision and Distributed Processing*, pages 135–144, 2002. 1, 2

[7] S. Noda and K. Ueda. Fire detection in tunnels using an image processing method. In *Proc. Vehicle Navigation and Information Systems Conference*, pages 57–62, Aug. 31–Sept. 2, 1994. 1

[8] X. Wang, Z. Wei, D. Liu, and X. Zheng. Smoke dynamic features based real-time fire detection. *Computer Technology and Development*, 11:9–17, 2008. 2, 4

[9] H. WB and P. JW. Real-time fire detection from video: A preliminary report. In *14th IPPR, Computer Vision, Graphics and Image Processing*, 2001. 1

[10] F. Yuan. Motion accumulation and translucence based model for video smoke detection. *Shuju Caiji Yu Chuli/Journal of Data Acquisition and Processing*, 22(4):396 – 400, 2007. 2