Fire Smoke Detection Based on Contextual Object Detection

Xuan Zhao
Nanjing University of Posts and Telecommunications
Nanjing, China
e-mail: 2458900120@qq.com

Hang Ji
Nanjing University of Posts and Telecommunications
Nanjing, China
e-mail: 312484203@qq.com

Dengyin Zhang
Nanjing University of Posts and Telecommunications
Nanjing, China
e-mail: zhangdy@njupt.edu.cn

Huanhuan Bao
Nanjing University of Posts and Telecommunications
Nanjing, China
e-mail: 335785739@qq.com

***Abstract*—Smoke detection based on automatic visual system has been applied to fire alarm in open spaces where traditional smoke detection system is not suitable for it. However, detecting the course of smoke posed great challenges for both systems. To address this problem, we propose a new method that combines context-aware framework with automatic visual smoke detection. The strategy is evaluated on dataset and the results demonstrate the effectiveness of the proposed method.**

Keywords-fires;alarm systems; smoke detectors; context

I. INTRODUCTION

Automatic visual object detection has been fully used of in order to alarm fire disaster as soon as possible for open space. Considering the traditional fire detection system's monitoring region's limitation and data transporting latency, along with the development of visual processing technology, vision technology based on fire detection has attracted researchers' attention. Compared with traditional fire detection method, vision technology based on fire detection has three main advantages: availability, controllability and instantaneity. First, the installation of various surveillance cameras makes sure monitored regions can be available to visual detector system, which enable security attendants to master real-time situation. Next, controllability is reflected in videos being stored via transmission once fire disasters happened. Finally, limited computational cost and efficient algorithm ensure the instantaneity of fire alarm.

As visual detectors, the characteristic of targets to be detected should be extracted precisely. The key to detecting fire alarm automatically based on vision is to express fire situations' features definitely. There have been considerable procedures in extracting the features of fire disaster images by far: the existing work focuses mainly on early detection of smoke and fire flame. In [1-5], smoke in early fire can be extracted from image mainly for its color, contours and motion orientation, while fire flame's distinctive characteristics includes color and motion frequency. Owing to the visibility and widely spreading of smoke in fire disasters, researchers are inclined to use smoke as the detection target to get fire alarm.

However, smoke can be detected in some certain situations rather than fire disaster, such as setting off fireworks, which could result in the false detection. In this literature, our solution to this problem is to introduce contextual objects detection into fire alarm verification. Smoke is detected followed by contextual objects detection, which aims to form consistent goal pairs between smoke and remarkable objectives to estimate the smoke scene. In smoke detection is employed using the characteristics of color histogram and fuzzification. We detect contextual objects using Hoff transform [6] in certain regions where smoke at the center in order to verify whether fire disaster occurred or not.

Implemented in this approach, fusion of these two descriptors is the key to decrease fire alarm false positives. Amounts of databases, such as video images captured online, data sets provided for deep learning training and fire smoke videos recorded in reality in open spaces can be determined by the proposed category for best results in the case of fire alarm precision rate of 87.6%.

In the following sections the method structure containing motion objects detection, smoke detection and contextual objective detection in detail is described in Section II. Section III gives experimental results while conclusions are drawn in Section IV.

II. DESCRIPTION OF ALGORITHMS

In order to eliminate the non-disaster fire situation, there are three elements to fire smoke alarm: motion objects detection, smoke region verification and contextual objects detection.

A. Motion Objects Detection

Owing to smoke's dynamic characteristic, this paper uses an adaptive background subtraction algorithm to identify moving objects. As is shown in Fig. 1, the potential problem of conventional background algorithm is that if the pixel in the motion place, it will be updated with the next sampled pixel. To make sure the background modeling closer to the current image's background, we propose an adaptive algorithm. Different from [7], the gray-scale value of pixel at

978-1-5386-4991-6/18/$31.00 ©2018 IEEE

(x, y) in motion region is updated with the value between the former background gray-scale value and the current image gray-scale value:

$$B_{i+n}\{x,y\} = \begin{cases} \alpha B_i\{x,y\} + (1-\alpha) I_i\{x,y\} & if\ (x,y)\ in\ motion\ region \\ B_i\{x,y\} & else \end{cases} \tag{1}$$

where $B_{i+n}\{x,y\}$ and $B_i\{x,y\}$ are the background gray-scale value of pixel at (x,y) in frame $i+n$ and i separately, $I_i\{x,y\}$ represents the current image's gray-scale value at (x,y) and α is a number between 0 and 1. Because the sluggish motion rate of smoke region, n, defined as a sampling rate, is recommended equal to at least ten to distinguish sampled image from the former.

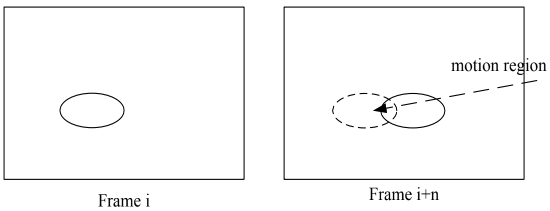

Figure 1. A background subtraction model

For each frame, the motion region is extracted as the subtraction between the gray-scale value of background and sampled frame:

$$M_i(x,y) = \begin{cases} 1 & \text{if } \left|I_i\{x,y\}\text{-}B_i\{x,y\}\right| \geq th \\ 0 & \text{else} \end{cases} \tag{2}$$

where th is the number of the subtraction's threshold value, which can be adjusted to reduce the influence of illumination variation.

B. Smoke Region Verification

Since smoke has various features compared with other moving physical objects, allying smoke features with each other is a robust way to model the behavior of smoke. In this method, two apparent features including color and fuzzification are applied to filter out non-smoke motion regions. Only regions that meet both color analysis and fuzzification peculiarity can be selected as candidate smoke regions.

Smoke-color is one of the most important features and several techniques have been proposed for color feature modeling in [8,9], some of which work well in some particular situations. The color of burning things ranges from white to gray, whose characteristic is that the gray level statistics of B,G,R channels tend to gather together. In Fig. 2, the color histograms of smoke regions in three different channels are similar to each other while non-smoke region is not. We choose to use the color histogram method in [10], which is suited for detected moving region to classify into color histogram descriptor, and three elements are added to it: $|R-G|$, $|G-B|$ and $|R-B|$.

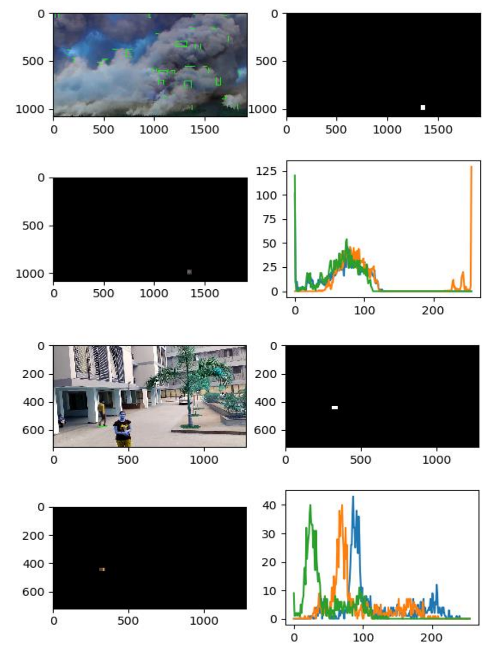

Figure 2. Color histograms of a sampled smoke region and non-smoke region at B,G,R channel

Based on smoke's fuzzification characteristic, we consider the high frequency energy of smoke region as a feature because smoke regions contain more high frequency energy than usual. For the reason that high frequency energy mainly concentrates on texture and edge information in images and smoke regions conclude more texture and edge. Differ from the method mentioned in [11], this process uses the subtraction of high frequency energy between original pixels and wavelet transformed ones, $\triangle E$, to establish a feature vector.

In this research, the wavelet transformation is implemented using the approach applied in [11], where one coefficient of low frequency and three coefficients of high frequency are constructed and represented with LL, LH, HL, HH respectively. The high frequency of pixel at (x,y) is calculated as:

$$\mathrm{HW}_{(x,y)} = \left|HL_{(x,y)}\right|^2 + \left|LH_{(x,y)}\right|^2 + \left|HH_{(x,y)}\right|^2 \tag{3}$$

This paper uses Fourier transform to compute the high frequency energy of original image region, and the threshold of high frequency is set based on training data sampled from fire-smoke video sequences, the pixel of which at (x,y) is represented with $\mathrm{H}_{(x,y)}$. Hence $\triangle E$ at the pixel (x,y) is as:

$$\triangle E_{(x,y)} = \left|\mathrm{H}_{(x,y)} - \mathrm{HW}_{(x,y)}\right| \tag{4}$$

So that, in addition to the above three elements: $|R-G|$, $|G-B|$ and $|R-B|$, the approach uses a combination from $\triangle E$ to establish two 4-feature vectors for a pixel as following:

1) $X^{\mathrm{I}} \in$ 4-dimension space of $[R, G, B, \triangle E]$

2) $X^{\mathrm{II}} \in$ 4-dimension space of $[|\mathrm{R}-\mathrm{G}|, |\mathrm{G}-\mathrm{B}|, |\mathrm{R}-\mathrm{B}|, \triangle E]$

Using these two feature vectors of a pixel in motion regions, KNN classifier in [12] investigates and extracts candidate smoke regions. As it is shown in Fig. 2, detected smoke regions gather together and locate at the edge of smoke mostly because of the diffusivity of smoke. For this reason, one filter condition is added to accurately detect smoke, that if the totality of candidate smoke regions is less than three in one limited area, the candidate smoke regions will not belong to smoke regions.

C. Contextual Objects Detection Modeling

Smoke can be caused by non-disaster fire situations, such as setting off fireworks, etc. For the purpose of solving the problem, contextual objects detection is proposed to combine the detected smoke regions with symbolic objects around it to make an identification that what the sources of smoke is and determine whether it is caused by fire disaster or not. The relation between smoke regions and some enumerated contextual objects is described as follows in Table I:

TABLE I. RELATION BETWEEN SMOKE AND CONTEXTUAL OBJECTS

Objects	**Objects characteristics**		
	Location	***Possible situation***	***Category***
fireworks	below	Setting off fireworks	A
people	around	Outdoor barbeque	A
chimney	below	Factory	B
car	surround	Exhaust emission	B
censer	below	Temple	A/B

In this implementation, because scenes mentioned in Table I can be classified to people-existing (category A) and non-people-existing (category B) scene, detecting people firstly can increase the efficiency of contextual objects detection. We use the background subtraction method mentioned in [7] and extract human boxes with the ratio of height and width of detected box firstly. So once the moving region boxes get from part A in Section II identified as smoke regions, the rest of them should be the input of detecting people model process. Next, fire smoke images with people existing and non-people existing are classified into particular scenes with Hoff transform method according to corresponding objects respectively in Table I, other wisely alarm will be raised to draw supervisory personnel's attention to make decision that if it is a disaster. Finally, the sampled images will be collected into databases in addition with new label, if the debatable images are judged into non-disaster fire smoke scene, so that scenario category is enlarged to ensure integrity of the category and increase the robustness of supervisory system. The procedure of contextual objects decision is shown as:

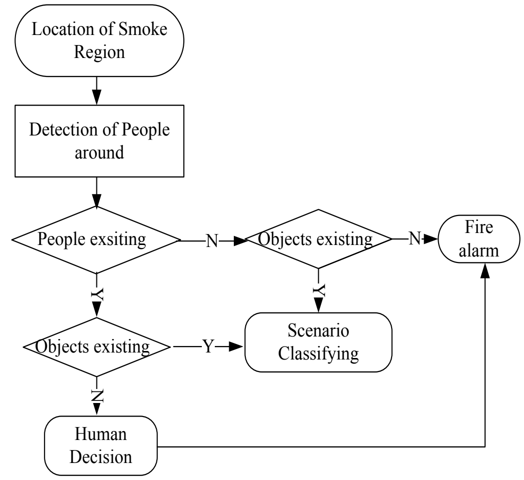

Figure 3. Contextual objects detection process

After the selection of scene to determine whether human exists, the corresponding contextual objects mentioned in Table I are detected respectively using the method Hough transform [6]. Besides, in order to overcome the shortcoming of insufficient information and extreme variation in detected objects, we consider the locational relationship of the detected smoke regions as the basic of verifying contextual objects referring to the possible situation column in Table I. If the locational relationship between detected contextual objects and smoke regions fails to conform to the given principle, the detected contextual objects will be dropped. In this way, a more reliable contextual objects detector strategy can be provided unambiguously. Fire alarm will not be triggered if the non-disaster objects detected, which increases the accuracy rate of fire alarm.

III. EXPERIMENTAL RESULTS MAINTAINING THE INTEGRITY OF THE SPECIFICATIONS

Experiments were carried out based on 800 minutes videos captured online, recorded in reality along with machine learning training datasets and platform with Python language. 70% of datasets are used to train with k-Nearest Neighbor (KNN) classifier mentioned in [13-14] and the rest of them are tested based on the system. The system can detect motion regions, verify smoke regions, filter out non-disaster scenes automatically and trigger fire alarm other wisely at 14 frames every second over a 320*240 pixels image sampled from videos every ten frames.

The following figures show results of detecting motion regions and human respectively:

Figure 4. Motion objects detection in smoke and non-smoke scenes.

Figure 5. Human detection in smoke scenes.

Fig. 6 shows enumerative results of detected smoke regions, which are marked in boxes. Note that non-smoke motion objects are correctly filtered on account of two four-dimensional vectors. In Fig. 7, the results of contextual objects detection in images containing smoke are exhibited along with labeled mark.

Figure 6. Examples of smoke region detection results.

Figure 7. Examples of contextual objects detection and marking.

However, the disadvantage of this method is that amounts of annotation tasks have to be done previously at the data training phase. In this way, objects can be detected at the rate of 87.6% on average and marked precisely. The system algorithm applied to the left 30% of databases and the accuracy rate of fire alarm in different scenes shown in Table II:

TABLE II. FIRE ALARM ACCURACY RATE IN DIFFERENT SCENES

Scene	*Tot*	*Fire-alarm accuracy rate*
Setting off fireworks	303	86.3%
Outdoor barbeque	217	90.6
Chimney smoke	100	87.1%
Temple	50	85.6%
exhaust emissions	50	85.7%

IV. CONCLUSION

The primary element to make fire-alarm accuracy rate robust is the contextual objects detection based on relationship with detected smoke regions. The system effectively connects particular scenes with space-domain statistical methods to classify target objects. The models of contextual targets are not only based on the shape of them but also depend on the relative locations with smoke regions and so that scene can be described effectively.

It is almost impossible for a single detected smoke region to determine the alarm of fire disaster behavior. In this paper, contextual objects detection is proposed to filter out non-disaster situation. However, undefined scene is a challenge because of the variety of situations containing smoke. This problem at present is solved with human interference and introduce new categories into system manually, which can be improved with iteration.

REFERENCES

[1] Hongda Tian, Wanqing Li, Philip O. Ogunbona, Lei Wang, Detection and Separation of Smoke From Single Image Frames. IEEE Transactions on Image Processing. Vol. 27, March 2018, pp. 1164-1177, DOI: 10.1109/TIP.2017.2771499.

[2] Rui Chi, Zhe-Ming Lu, Qing-Ge Ji, Real-time multi-feature based fire flame detection in video. IET Image Processing. Vol. 11, April 2017, pp. 31-37, DOI: 10.1049/iet-ipr.2016.0193.

[3] Pasquale Foggia, Alessia Saggese, Mario Vento, Real-Time Fire Detection for Video-Surveillance Applications Using a Combination of Experts Based on Color, Shape, and Motion. IEEE Transactions on Circuits and Systems for Video Technology. Vol. 25, Jan 2015, pp. 1545-1556, DOI: 10.1109/TCSVT.2015.2392531.

[4] Caixia Cheng, Fuchun SunOne , Xinquan Zhou, One fire detection method using neural networks. Tsinghua Science and Technology. Vol. 16, Feb 2011, pp. 31-35, DOI: 10.1016/S1007-0214(11)70005-0 .

[5] Tian Qiu, Yong Yan, Gang Lu, An Autoadaptive Edge-Detection Algorithm for Flame and Fire Image Processing. IEEE Transactions on Instrumentation and Measurement. Vol. 61, May 2012, pp. 1486-1493, DOI: 10.1109/TIM.2011.2175833.

[6] C. Li, D. Parikh, Tsuhan Chen, Automatic discovery of groups of objects for scene understanding[C]. IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, July 2012, pp.2735–2742.

[7] D.S. Alex, A. Wahi, BSFD: background subtraction frame difference algorithm for moving object detection and extraction. Journal of Theoretical & Applied Information Technology, vol. 60, Feb 2014, pp. 623-628.

[8] Tom Toulouse, Lucile Rossi, Moulay Akhloufi; Turgay Celik; Xavier Maldague, Benchmarking of wildland fire colour segmentation algorithms. IET Image Processing, Vol. 9, Nov 2015, pp. 1064-1072, DOI: 10.1049/iet-ipr.2014.0935.

[9] Zhijian Yin, Boyang WanA, Feiniu Yuan; Xue Xia; Jinting Shi, Deep Normalization and Convolutional Neural Network for Image Smoke Detection. IEEE Access, Vol. 5, August, 2017, pp. 18429-18438.

[10] Zhang wei: A study on the detection of fire smoke detection based on the weighted feature group. submitted to Journal of Wuhan University of Science and Technology (2013).

[11] Huihui Bai, Ce Zhu, Yao Zhao: Optimized Multiple Description Lattice Vector Quantization for Wavelet Image Coding. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, July 2007, pp. 912-917, DOI: 10.1109/TCSVT.2007.898646.

[12] Jian Hou, Huijun Gao, Qi Xia, Naiming Qi, Feature Combination and the kNN Framework in Object Classification. IEEE Transactions on Neural Networks and Learning Systems, Vol. 27, August 2016, pp. 1368-1378.