A Novel Smoke Detection Method Using Support Vector Machine

Hidenori Maruta
Information Media Center
Nagasaki University, Japan
1-14 Bunkyo-machi, Nagasaki-shi
Nagasaki, Japan
Email: hmaruta@nagasaki-u.ac.jp

Akihiro Nakamura
Graduate School of
Science and Technology
Nagasaki University, Japan
1-14 Bunkyo-machi, Nagasaki-shi
Nagasaki, Japan
Email: d709026h@cc.nagasaki-u.ac.jp

Fujio Kurokawa
Faculty of Engineering
Nagasaki University, Japan
1-14 Bunkyo-machi, Nagasaki-shi
Nagasaki, Japan
Email: fkurokaw@nagasaki-u.ac.jp

***Abstract*—An early and certain fire detection is one of the important issue to keep safe infrastructures. Especially, it becomes an urgent problem in large facilities like port facilities, large factories and power plants, due to its large harmful effect to surrounding areas. In these places, smoke is an important and useful sign to detect the fire robustly even in such open areas. In this study, we present a novel and robust smoke detection method based on image information. Firstly we extract moving objects of images as smoke candidate regions in a pre–processing. Because smoke has a characteristic pattern as image information, we treat smoke patterns as textures. Here we use texture analysis to extract feature vectors of images. To classify extracted moving objects are smoke or non–smoke, we use support vector machine(SVM) with texture features as input features. Extraction of moving objects are sometimes easily affected by environmental conditions. So we accumulate the result of the classification with SVM about time to obtain accurate extraction results of smoke regions.**

I. Introduction

It is one of the important issues to detect fire to build and keep safe infrastructures. Port facilities, power plant or large factories, which are large infrastructures in open areas, harm the surrounding areas of them in the case of fire. Especially for these places, an early and certain fire detection method is needed urgently. Computer vision based approaches is useful for these places because they use image information obtained, for example, surveillance cameras, as they are usually set already and there is little cost to set special sensing devices around them. However, the direct detection of fires or flames (e.g. [1], [2]) is sometime very limited because sources of the fire or flames cannot always fall into the field of view due to their positions and sizes especially in open areas. In these cases, smoke is important and useful sign for the fire detection. We can simply and easily observe the smoke of the fire by a camera even if the flames are not visible.

In this study, we present a novel smoke detection method using image information. Previous approaches of the smoke detection, using pixel- and block-based, or color-based image processing methods have been presented[3], [4]. These methods try to detect edge or contour information of smoke. However, they have difficulties in treating characteristic properties of smoke and needing high–cost computations to detect smoke whole of wide-view images or image sequences. Additionally, handling color information is also a problem because it may be affected by sources or fuel types of the fire. In our method, we use texture analysis and non-linear classification with SVM to extract smoke regions in image sequences. Firstly, as a pre–processing stage, we detect moving objects in images as candidates of smoke regions. Next, we apply the texture analysis to the regions of moving objects. We extract texture features which become components of a feature vector. Based on this feature vector as input vector, we apply SVM with non–linear kernel to discriminate whether the region is smoke or not. Moreover, to improve the extraction result, we accumulate classification results of SVM about time on images. This process can be available us to obtain more accurate extraction results of the positional information of smoke. To evaluate our method, we examine it with real–scene images obtained from a surveillance camera under various conditions.

II. Pre-Processing

In our method, we detect moving objects in images as candidates of smoke regions in the pre–processing. As a growth-speed of smoke is considered, we obtain 1 frame-per-second rate image sequences $f(t)$ $(t = 0, 1, 2, \cdots)$ from original ones. The pre–processing consists of following five steps: (i) image subtraction and accumulation, (ii) image binarization, (iii) morphological operation (*opening*) [5] to remove noise-like parts, (iv) extraction of Feret's diameters and their position as *Feret's regions*, moving objects that are obtained as smoke candidate regions, (v) creation of the image mask to obtain the moving objects of gray-scale image $f(t)$. Outline of the pre–processing is shown in Fig. 1. After the pre–processing, we obtain images which contain some moving objects as smoke candidate regions, in which smoke detection is proceeded.

A. Subtraction and Accumulation

We use the image subtraction technique to extract regions of moving objects. A subtracted image frame is written as $g(t) = f(t) - f(t-1)$. As the growth speed is not so fast, smoke in subtracted image not so clear. So we use the image

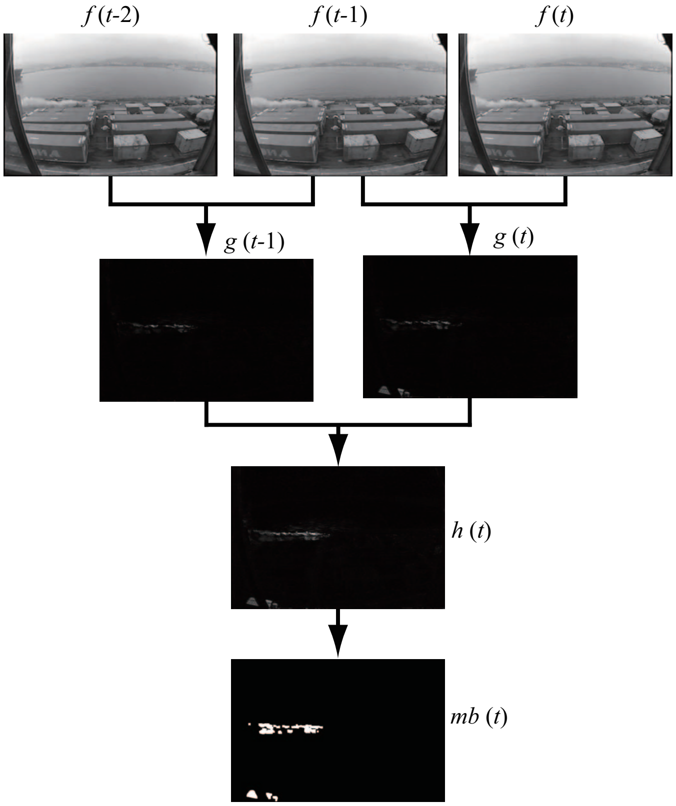


Fig. 1. Outline of the pre–processing described in Sec. II.

frame $h(t)$ which accumulate two subtracted image, i.e. $h(t) = |g(t) + g(t-1)|$, in the following processing.

B. Binarization and Morphological operation

To remove very small noise–like regions in $h(t)$, we operate binarization and opening, an morphological operation which removes the noise-like regions in binary images. We process $h(t)$ to be binary images $b(t)$ with the binarization operation. In this binarization, it is effective and very important to set the adequate threshold value to extract contained smoke regions. Because smoke has the property of semi–transparency, it is affected its background condition. To solve this problem, we use *Otsu*'s automatic threshold selection method[6]. We also use a morphological operation, *opening*, to remove small noisy regions in $b(t)$. Here let $m_b(t)$ be the image frame sequence after opening operation to $b(t)$. $m_b(t)$ is used in the following process to detect the candidates of smoke regions.

C. Extraction of Feret's regions

To determine shapes and positions of moving object regions in $m_b(t)$, we apply a method to extract Feret's diameter of them. We can extract a circumscribed rectangle of one region of objects (moving object region in this case), which is the smallest surrounding rectangle. Lengths of horizontal and vertical edges of this rectangle are called Feret's diameters. When we extract it, we also obtain the position and the approximated shape of the object as the rectangle. We call

Fig. 2. An example image of the pro–processing result. Extracted Feret's regions are superimposed on the subtracted image $g(t)$.

it Feret's region, which is considered as estimated moving objects with the rectangle shape and the positional information. Each $m_b(t)$ may have several Feret's regions, so let $F(t; i)$ be the i th Feret's region of $m_b(t)$. We use the information of $F(t; i)$ as an image mask to estimate moving object regions in the subtracted image frame $g(t)$. To extracted regions with this image mask, we apply the texture analysis to estimated regions and obtain information to examine the existence of smoke in the image sequence in the method described in the following. An example of the pre–processing result described above is shown in Fig. 2. In this example, moving object is only smoke at the center of the image, and obtained Feret's regions are represented as rectangles and superimposed on the subtracted image frame $g(t)$. On the left bottom of Fig. 2, there remain Feret's regions and these are not smoke–regions. To obtain more accurate results, we need further processing described in the next section.

III. Smoke detection method

A. Texture Features

After the pre–processing, we obtain the candidate regions of smoke in $g(t)$. We focus on the texture pattern[7] of the smoke as a feature vector to use to determine which regions are smoke. The texture feature of smoke is defined with the co–occurrence matrix, so it does not depend on the visible size of smoke in images. So the texture feature is suitable for the purpose of our method, which needs the robustness about the size in the images for general conditions. We select 14-types popular texture features[8] that become the components of the feature vector of extracted Feret's regions. We determine which Feret's regions are smoke using this feature vector as a input vector.

B. Support Vector Machine

For the discrimination of the extracted Feret's regions, we use the SVM approach. We use the texture features described in as the input vector of SVM. The texture features described in of smoke sometimes similar values compared to those of non–smoke's. This cause that the input vectors of smoke and non–smoke regions hard to be linearly separable. So we select to use non–linear kernel (RBF kernel) of SVM as Eq. (1).

$$K(x_i, x) = \exp\left(-\gamma ||x_i - x||^2\right) \tag{1}$$

where $x_i \in \mathbb{R}^{14}$ is a texture feature vector of a Feret's' region. To train the SVM, we prepare manually selected Feret's' region of smoke, which we call *ideal smoke*. Using trained SVM, we discriminate whether Feret's' regions are smoke or not. When the SVM's output of the Feret's' region in $g(t)$ is smoke, we set the label $l_g(x, y; t)$ of the point (x, y) in the Feret's' region to be 1 and else to be 0.

$$l_g(x, y; t) = \begin{cases} 1((x, y) \ is \ smoke.) \\ 0(else) \end{cases} \tag{2}$$

Note that if some Feret's' regions which discriminated smoke are overlapped, we set the the label $l_g(x, y; t)$ of the overlapped point (x, y) still to be 1. That is, any point in the $g(t)$ has the label 0 or 1 at time t.

C. Time Accumulation method

In the real situation, we sometime have difficulties using only one index due to effects of several conditions. For example, there exist not only smoke but some moving objects as noises. To obtain the accurate result of the smoke detection, we consider to accumulate the labeling results $l_g(x, y; t)$ with SVM about *time*. The accumulation is defined as follows:

$$\mathcal{A}_g(x, y; t) = \sum_{t=\tau}^{\tau+T_a} l_g(x, y; t) \tag{3}$$

where $l_g(x, y; t)$ is the label of (x, y) described in Eq. (2) and accumulation time T_a is selected manually. Based on $\mathcal{A}_g(x, y; t)$, as well as the information of Feret's' regions, we can extract the smoke areas in $g(t)$. That is, when there exist a point (x, y) whose $\mathcal{A}_g(x, y; t)$ is higher than the manually selected threshold, all Feret's regions which contain that point discriminated as smoke at time t. So the smoke areas in $g(t)$ consist of Feret's regions discriminated as smoke with $\mathcal{A}_g(x, y; t)$.

IV. Experimental Results

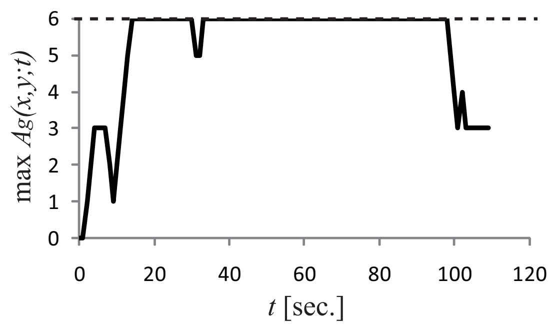
A. Experiment I: Evaluation by same videos used in SVM training

1) Video Used in Experiments: In Experiment I, we evaluate our method using three videos, from which we obtain a data set for SVM training. The size of them is 720×640 and conditions such as the lighting and the wind are different among them. Table I summarizes them used in following experiments.

2) Training of Support Vector Machine: To train SVM, we prepare a training set which consists of 1200 regions of images (600 smoke regions and 600 non–smoke regions) manually selected. All components of the feature (input) vector are normalized, i.e. all components are rescaled to have the zero mean and the unit variance. We use the *soft margin*[9] loss function in the training algorithm of SVM. Hence we need to optimize the kernel parameter γ and the trade–off (penalty) parameter C of the loss function[9]. We search the optimal parameters γ and C using the training set with *10–fold–cross validation*[9], [10]. By changing γ and C we obtain the best cross validation rate 86.09%. In the experiments, we use this trained SVM for the discrimination.

TABLE I
Summary of image sequences used in experiments.

	existence of smoke	other moving object
Video 1	Yes	human, car
Video 2	Yes (transmissive)	human, car
Video 3	No	gantry crane, plastic sheet, human, car

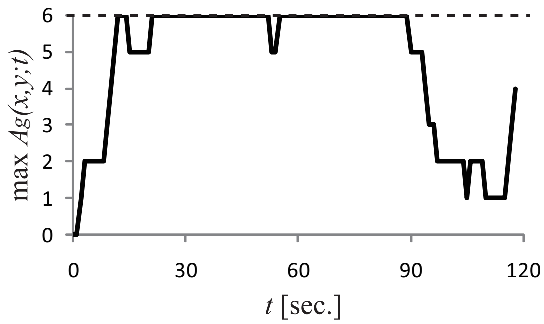


(a)

(b)

Fig. 3. (a) $\max_{(x,y)} \mathcal{A}_g(x, y; t)$ of Video 1. (b)Extracted smoke region at $t = 14$.

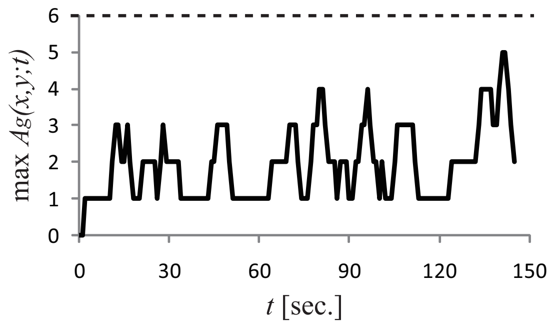
3) Time Accumulate Characteristic: In the experiments, we set the time accumulation parameter $T_a = 6$ in Eq. (3). The experimental results using Video 1 and Video 2 are shown in Fig. 3 and Fig. 4 respectively. Fig. 3(a) and Fig. 4(a) are the time series data of the maximum value of $\mathcal{A}_g(x, y; t)$ at time t in Video 1 and Video 2. From these results, we can observe that the maximum value grows as the smoke appears and grows after the time-lag of the accumulation in both image sequences. In Video 1, the maximum value $\mathcal{A}_g(x, y; t)$ reaches its limit ($\mathcal{A}_g(x, y; t) = T_a = 6$) at $t = 14$ and keeps its limit value during the smoke exists. In Video 2, the maximum value $\mathcal{A}_g(x, y; t)$ reaches its limit ($\mathcal{A}_g(x, y; t) = T_a = 6$) at $t = 12$ and almost keeps its limit value during the smoke exists even



(a)

(b)

Fig. 4. (a) $\max_{(x,y)} \mathcal{A}_g(x,y;t)$ of Video 2. (b)Extracted smoke region at $t = 12$.



(a)

(b)

Fig. 5. (a) $\max_{(x,y)} \mathcal{A}_g(x,y;t)$ of Video 3. (b)The image at $t = 141$ where $\mathcal{A}_g(x,y;t)$ is the maximum.

it is more transmissive compare to Video 1. Fig. 3(b) and Fig. 4(b) are example images of the smoke region extraction at $t = 14$ and $t = 12$ respectively. In both cases, we set the threshold value 6, i.e. Feret's regions which contain the point that has $\mathcal{A}_g(x,y;t) = 6$ are used for the extraction. These result show that our method can extract the accurate smoke region in the image.

The experimental result using Video 3 is shown in Fig. 5 . Fig. 5(a) is the time series data of the maximum value of $\mathcal{A}_g(x,y;t)$ at time t in Video 3. From this result, we can observe that the maximum value cannot keep stable value.Fig. 5(b) is the image at the first maximum point $\mathcal{A}_g(x,y;t) = 5$ ($t = 141$). If, we set the threshold value 6 as same as the case of Video 1 and Video 2, $\mathcal{A}_g(x,y;t)$ of Video 3 cannot reach the threshold. So we can determine that smoke does not exist.

B. Experiment II: Smoke extraction with videos under different conditions and settings

For evaluation of our method under general conditions, we do some experiments using other different videos (Video 4 – Video 10). The videos used here obtained from website[11], or prepared by authors. These videos include smoke and have different sizes, camera settings and environmental conditions. However, in Experiment II, we use same SVM settings and the time accumulation condition as Experiment I. That is, the training set of SVM is obtained from videos in Experiment I, not from videos used in Experiment II.

Final results of smoke extraction are shown in Fig. 6–12 respectively.

From this experiment, we confirm that our method can extract smoke regions in videos under general conditions without re-training of SVM.

V. Conclusions

We present the novel smoke detection method based on the texture analysis and the support vector classifier. We focus on the image information of smoke as the texture pattern, which is not affected the size of smoke in the images. After the pre–processing, which extracts moving objects as candidate regions in images, we also extract 14–types of texture features of these regions. Using this texture features as the input vector, we determine whether the candidate regions are smoke regions or not with non–linear SVM. Moreover, to obtain more accurate extraction results of smoke areas in images, we accumulate the detection result of SVM in each image about time. The time–accumulation works well as smoke has different properties from other moving objects. For the evaluation our method, we examine it with some examples of image sequences. In the experiments, there exist smoke and other moving objects which are considered as obstructions in image sequences. Experimental results show the effectiveness of our method under general conditions.

(a) Original image (Video 4).

(b) Extraction image (Video 4).

Fig. 6. Example frame of the extraction result (Video 4, 320×240).

There remain some problems for further works. It is needed to evaluate our method using more various types of smoke. Additionally, to use in the real situation, we must compile the smoke detection system with other systems. One way is to combine our method with multi camera system to cover the wide target area.

REFERENCES

[1] W. Phillips III, M. Shah and N. V. Lobo, "Flame recognition in video", Pattern Recognition Lett. 23, pp. 319–327, 2002.

[2] B. U. Töreyin, Y. Dedeoglu, U. Güdükbay and A. E. Çetin , "Computer vision based method for real–time fire and flame detection", Pattern Recognition Lett. 27, pp. 49–58, 2006.

[3] B. U. Töreyin, Y. Dedeoglu and A. E. Enis, "Wavelet based real–time smoke detection in video", 13th European Signal Processing Conference (EUSIPCO), 2005.

[4] B. U. Töreyin, Y. Dedeoglu and A. E. Enis, "Contour based smoke detection in video using wavelet", 14th European Signal Processing Conference (EUSIPCO), 2006.

[5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing* 3rd ed., Prentice Hall, 2007.

[6] N. Otsu, "A Threshold Selection Method from Gray–Level Histograms", IEEE Trans. System, Man, and Cybernetics, Vol. SNC–9, No. 1, pp. 62–66, 1979.

[7] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 2nd ed., Academic Press, 2003.

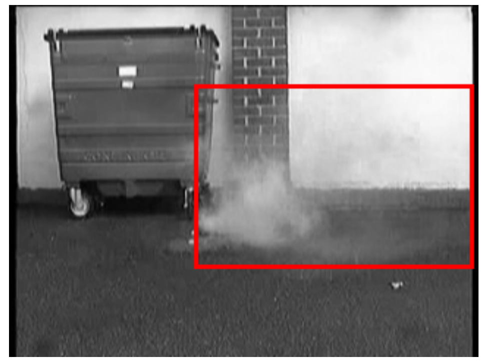
[8] S. Umbaugh , *Computer Imaging: Digital Image Analysis and Processing*, CRC Press, 2005.

[9] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[10] C. Chang, C. Lin, LIBSVM , a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm .

[11] http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SampleClips.html

(a) Original image (Video 5).



(b) Extraction image (Video 5).

Fig. 7. Example frame of the extraction result (Video 5, 320×240).

(a) Original image (Video 6).

(b) Extraction image (Video 6).

Fig. 8. Example frame of the extraction result (Video 6, 320×240).

(a) Original image (Video 7).

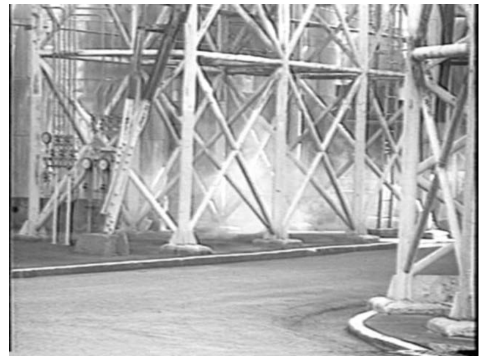
(b) Extraction image (Video 7).

Fig. 9. Example frame of the extraction result (Video 7, 720×480).

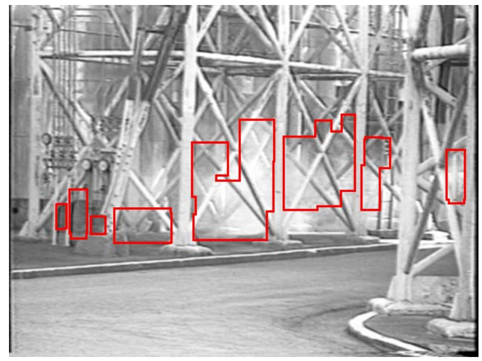
(a) Original image (Video 8).

(b) Extraction image (Video 8).

Fig. 10. Example frame of the extraction result (Video 8, 640×480).

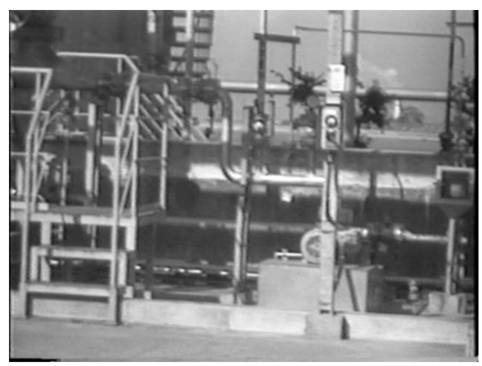


(a) Original image (Video 9).

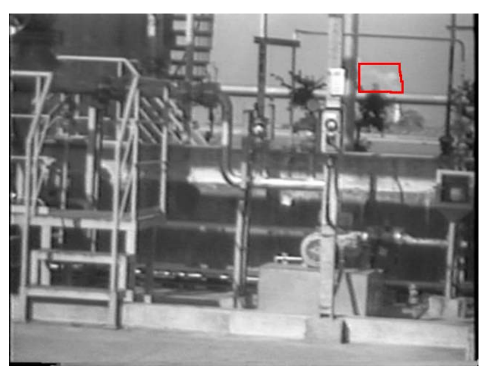


(b) Extraction image (Video 9).

Fig. 11. Example frame of the extraction result (Video 9, 640×480).



(a) Original image (Video 10).



(b) Extraction image (Video 10).

Fig. 12. Example frame of the extraction result (Video 10, 640×480).