第47卷第8期 中国科学技术大学学报 Vol.47,No.8
2017年8月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Aug. 2017

文章编号:0253-2778(2017)8-0653-12

Features selection for video smoke detection using random forest

WEN Zebo[1,2], KANG Yu[1,2], CAO Yang[2], WEI Meng[1], SONG Weiguo[1]

(1. *Department of Automation, University of Science and Technology of China, Hefei* 230027, *China*;
2. *State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei* 230027, *China*)

Abstract: Using the random forest algorithm, a video smoke detection method with features selection was proposed. The method first extracted four original smoke image features including color features in RGB space, wavelet high frequency sub-images, multi-scale local max saturation, and multi-scale dark channel to input the random forest(RF). Then it utilized haze image formation model to make the synthetic smoke images from non-smoke images and partitions these images into blocks as the samples for RF. Thirdly, it trained RF to get the selected features from the original features and used support vector machine(SVM) to get a classifier which recognizes the smoke blocks and the non-smoke blocks. And then the smoke region candidate can be extracted from video images by the classifier. Finally, the method analyzed the detected smoke region with the features of the growth rate and the perimeter to area ratio to make the final decision on video smoke detection. The experimental results show that the proposed method can detect the smoke timely and give a fire alarm with a lower false-alarm rate.

Key words: smoke detection; random forest; SVM; features selection; wavelet transform; smoke growth; dark channel

CLC number: O157.4 **Document code**: A doi: 10.3969/j.issn.0253-2778.2017.08.004

Citation: WEN Zebo, KANG Yu, CAO Yang, et al. Features selection for video smoke detection using random forest[J]. Journal of University of Science and Technology of China, 2017,47(8):653-664.
文泽波,康宇,曹洋,等. 基于随机森林特征选择的视频烟雾检测[J]. 中国科学技术大学学报,2017,47(8):653-664.

基于随机森林特征选择的视频烟雾检测

文泽波[1,2],康　宇[1,2],曹　洋[2],魏　梦[1],宋卫国[1]

(1.中国科学技术大学自动化系,安徽合肥 230027;2.中国科学技术大学火灾科学国家重点实验室,安徽合肥 230027)

摘要:利用随机森林算法,提出了一种基于随机森林特征选择的视频烟雾检测方法.首先,提取四种表征烟雾的特征:RGB颜色特征,小波变换高频子图,多尺度局部最大饱和度,多尺度暗通道;其次,根据烟雾图像信息模型利用无烟图片合成烟雾图片并分块得到随机森林训练样本;第三,训练随机森林进行特征选择并通过训练支持向量机得到识别烟雾块和非烟雾块的分类器,并由此得到视频图像帧的疑似烟雾区域;最后通过

Received: 2016-10-09; **Revised**: 2016-11-07

Foundation item: Supported in part by National Natural Science Foundation of China (61422307, 61673361).

Biography: WEN Zebo, male, born in 1992, Master candidate. Research field: Control theory. E-mail: dilong@mail.ustc.edu.cn

Corresponding author: KANG Yu, PhD/professor. E-mail: kangduyu@ustc.edu.cn

视频烟雾区域的凸形度和增长率分析,得到烟雾检测的结果。实验结果表明,该方法能够及时的预警烟雾同时降低火灾预警的误报率.

关键词:烟雾检测;随机森林;支持向量机;特征选择;小波变换;烟雾增长率;暗通道

0 Introduction

Fire accidents usually cause economic and ecological damage as well as endangering people's lives. So it's very necessary to send a fire alarm before the fire breaks out. Since smoke is a forecasting sign of fire, vision smoke detection method are an effective way to issue fire alarms.

Vision smoke detection methods generally consist of two steps: Extracting the smoke features and determining the smoke regions. The first step, the extracted features are always the static features and dynamic features in video smoke images. But the in the second step, most of the conventional smoke region decisions depend on the empirical preset threshold. For instance, by extracting the RGB color features and analyzing the motion characteristics of smoke regions, Chen et al.[1] proposed a I-step method to detect smoke, but the method is excessively dependent on the smoke color, the set thresholds are dependent on experimental statistics. Töreyin et al.[2] proposed a real-time smoke detection method based on smoke image wavelet transform, and the method also depends on empirical thresholds. Alejandro, Yuan et al.[3-4] used the block motion orientation model to estimate the motion orientation of the smoke in video and proposed a smoke detection method, but the method has a high complexity by computing the accumulative motion orientation, and the thresholds of smoke motion direction estimation are also dependent on experimental statistics. These smoke detection methods are all empirical thresholds based, and this leads their lack of universal applicability. Because of this drawback, more and more researchers using the machine learning approaches to detect smoke, and treating the smoke detection problem as a smoke recognition problem.

Many machine learning approaches based smoke detection methods have been proposed. By extracting the static and dynamic features of smoke in video, Xu et al.[5] proposed an automatic fire smoke detection method by training BP neural network. Yang et al.[6] proposed a visual based smoke detection method using SVM. They trained the SVM with extracted features of the changing unevenness of smoke density distribution and the changing irregularities of the smoke contour to get a smoke classifier to detect smoke. Horng et al.[7] used the features of the HSI components of fire to determine fire pixels and non-fire pixels with BP neural network to get fire area candidates. Tung et al.[8] proposed a four-stage smoke detection algorithm by training the SVM to get a classifier to detect smoke. The smoke detection methods mentioned above are all based on the machine learning methods of BP neural network and SVM. By training the BP neural network and SVM with extracted features one can aim to detect smoke and reduce the dependence of experimental threshold. However, these methods cannot analyze the extracted smoke features automatically and need lots of videos to get smoke and non-smoke samples. Besides, it is difficult to get the pure smoke samples.

In order to automatically analyze the extracted smoke features, and reduce the complexity and difficulty to get training samples, we propose a video smoke detection method with random forest features selection. The method consists of five steps: ① Extracting four original features to train RF (random forest): RGB components, the wavelet sub-images, dark-channel and saturation of smoke image. ②Synthesizing the smoke images with non-smoke images, and partitioning all the images into blocks as samples of RF. ③ Training the RF to select features from the original features. ④ Training SVM with the selected features to get a classifier. ⑤ Extracting the smoke region in video with the classifier, and analyzing the smoke growth rate and the perimeter to area ratio to complete smoke detecting and according to the set sampling termination parameter, we can get a fire alarm.

1　Smoke features extracted for RF

1.1　Color features in RGB space

Generally, the color of fire smoke depend on the fuel categories and combustion adequacy. However, smoke of many combustible materials usually exhibits a grayish color, and in RGB color space, the RGB components of smoke images are approximately close to each other[1]. By analyzing the smoke color and the transform between RGB color model and HSI color model, we can roughly determine the smoke pixels through the rules that follow:

$$\left.\begin{array}{l} r_1: R \pm \alpha = G \pm \alpha = B \pm \alpha \\ \quad r_2: L_1 \leqslant I \leqslant L_2 \\ \quad r_3: D_1 \leqslant I \leqslant D_2 \\ \text{if: } (r_1) \text{ and } \{(r_2) \text{ or } (r_3)\} \\ \text{smoke pixel;} \\ \text{else:} \\ \text{non-smoke pixel} \end{array}\right\} \tag{1}$$

where R, G, B donates the RGB components in RGB color space, respectively. I is the intensity of smoke images. $\alpha, L_1, L_2, D_1, D_2$ are the given thresholds.

Chen et al.[1] extracted the smoke pixels by using these rules to detect smoke in video, so the RGB components of smoke images can be extracted as a smoke characteristic. And we extract the smoke color features in $v_1 = [R, G, B]$.

1.2　The wavelet sub-images

In general, smoke is semi-transparent. It can obstruct the textures and edges in the background of an image. Since the edges and textures contribute to the high frequency information of the image, energies of the wavelet high frequency sub-images drop due to smoke in an image sequence. And this can be a feature for detecting smoke.

In Fig. 1, we get the wavelet sub-images of a smoke image and a non-smoke image. As shown in the sub-images, compared to the non-smoke image, the edges information of the smoke image is decreasing and the textures and edges of the smoke image are more smooth and fuzzy because of the smoke. So, we can utilize the wavelet sub-images to detect the smoke, and the wavelet coefficients can be a characteristic of smoke. And we extract the wavelet coefficients as another smoke feature in $v_2 = [cA, cH, cV, cD]$.

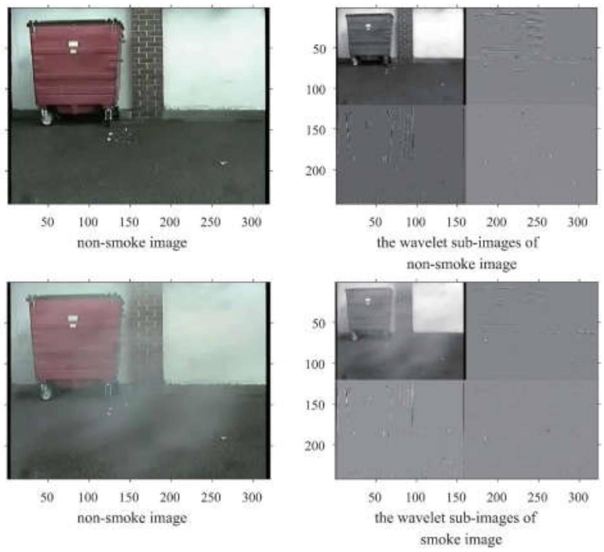

Fig.1　The wavelet sub-images of smoke image and non smoke image

1.3　Dark channel features

The dark channel features are a kind of statistics of outdoor images without smoke and haze. It is based on a key observation: most local patches (except the white wall regions and the sky regions in image) in outdoor non-smoke and non-haze images contain some pixels whose intensity is very low in at least one color channel (R, G, B), and the intensity is sometimes close to 0. According to this kind of statistics, we can extract the dark channel of the smoke image to be a smoke characteristic.

In order to describe the dark channel clearly, we first get the smoke and haze image in formation model as follows:

$$\left.\begin{array}{l} I(x) = J(x)t(x) + A(1 - t(x)) \\ t(x) = e^{-kd(x)},\ t \in (0,1) \end{array}\right\} \tag{2}$$

where I is the observed intensity, J is the scene radiance, x is the pixel of image, A is the global atmospheric light, and t is the medium transmission. $t(x)$ donates the portion of the light that is not scattered and reaches the camera, and is dependent on the distance of the objects to the camera $d(x)$ and the scattering coefficient k[9-10].

In Ref.[9], the dark channel of the non-smoke image is defined as follows:

$$\left.\begin{array}{l} J^{\text{dark}}(x)=\min\limits_{y\in\Omega(x)}\left(\min\limits_{c\in\{R,G,B\}} J^{c}(y)\right) \\ J^{\text{dark}}\rightarrow 0 \end{array}\right\} \quad (3)$$

where c donates one of the components of R,G,B, J^{dark} is the dark channel, $\Omega(x)$ donates a local region in the image.

As shown in Eqs.(2) and (3), we can compute the dark channel of the smoke image as follows:

$$I^{\text{dark}}(x)=J^{\text{dark}}(x)t(x)+A(1-t(x)) \quad (4)$$

In Eq.(4), J^{dark} is close to 0, when there is smoke in the image, t is decreasing due to the scattering of the smoke particles, and t will be 0 as the smoke is very dense. On the other hand, A is almost environment lighting and it is approximately close to [1, 1, 1] (assuming the image is normalized). Therefore, the dark channel of smoke image I^{dark} is brighter than the dark channel of non-smoke image J^{dark}.

According to Eqs.(2), (3) and (4), the dark channel of the smoke image is generally brighter than the dark channel of the non-smoke image. We can define the multi-scale dark channel as follows:

$$D_r(x;I)=\min_{y\in\Omega_r(x)}\min_{c\in\{R,G,B\}} I^{c}(y)) \quad (5)$$

where $\Omega_r(x)$ denotes the $r\times r$ region of the image, c denotes one of the components of R,G,B.

In Fig. 2, we calculate the multi-scale dark channel D_1, D_4, D_{10} of the non-smoke image and the smoke image respectively. We find that the multi-scale dark channel of the non-smoke image is darkes than the smoke image and this becomes more obvious as the scale increases. So, the multi-scale dark channel can be used as a feature of the smoke image, and we extracted the multi-scale dark channel feature in $v_3=[D_1, D_4]$.

1.4 Saturation features

The saturation of smoke images can be defined as follows:

$$S(x;I)=1-\frac{\min\limits_{c\in\{R,G,B\}} I^{c}(x)}{\max\limits_{c\in\{R,G,B\}} I^{c}(x)} \quad (6)$$

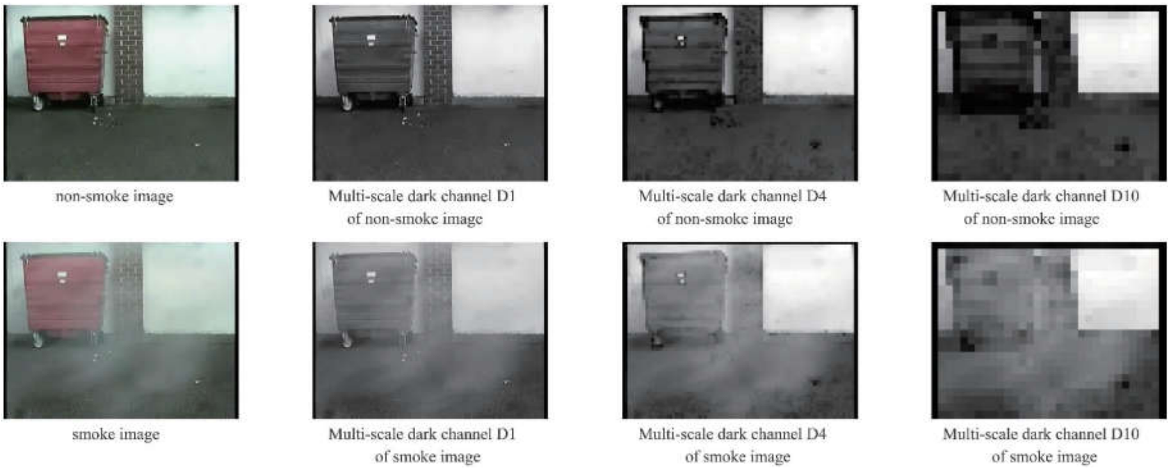

Fig.2 The multi-scale dark channel of smoke image and non smoke image

As shown in Eq.(2), we can get Eq.(7) by both sides of differential of equation:

$$\sum_x \|\nabla I(x)\| = t\sum_x \|\nabla J(x)\| \leqslant \sum_x \|\nabla J(x)\| \quad (7)$$

Therefore, compared to the saturation of the non-smoke image, the saturation of the smoke image is decreasing. We define the multi-scale local max saturation as follows:

$$S_r(x;I)=\max_{y\in\Omega_r(x)}\left(1-\frac{\min\limits_{c\in\{R,G,B\}} I^{c}(y)}{\max\limits_{c\in\{R,G,B\}} I^{c}(y)}\right) \quad (8)$$

where $\Omega_r(x)$ donates the $r\times r$ region of the image.

In Fig.3, we calculate the multi-scale local max saturation S_1, S_4, S_{10} of the non-smoke image and the smoke image respectively. We find that the multi-scale local max saturation of the non-smoke image is more white than that of the smoke image and this becomes more obvious as the scale increases. So, the multi-scale local max saturation can be used as a feature of the smoke image, and we extracted the multi-scale local max saturation feature in $v_4=[S_1, S_4]$.

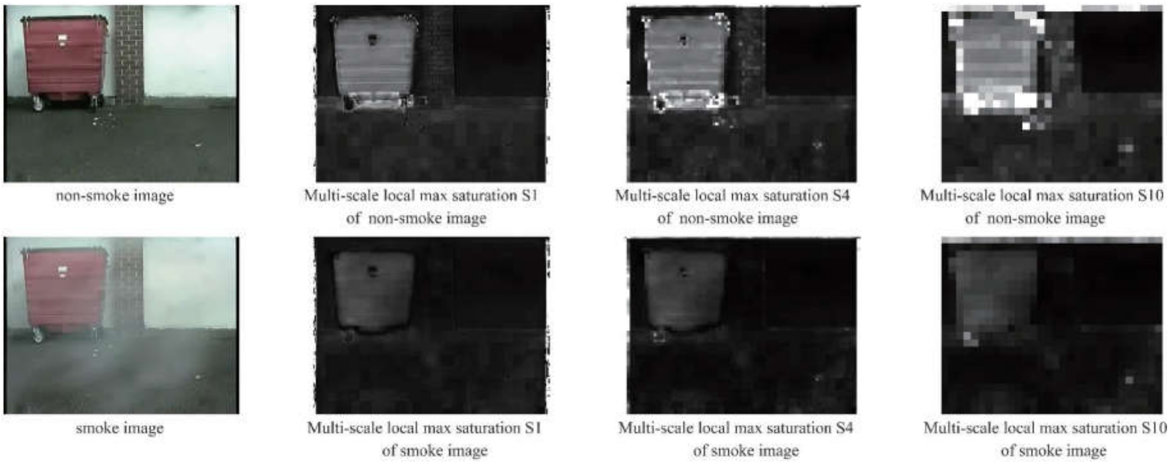

Fig.3 The multi-scale local max saturation of smoke image and non smoke image

Fig.4 The non-smoke block samples in partitioned non-smoke video images

2 Sample preparation, feature selection and SVM training

Although we have extracted the smoke features, we do not yet know which feature is important or not really important. To determine the smoke features we need, we analyze the extracted smoke features with RF, find the necessary features and discard the redundant features. As with the found necessary features, we can train the SVM to get a classifier.

2.1 Preparing training samples

The purpose of training the samples is to get the smoke region in video smoke images. In this paper, we extract the smoke regions by the block-matching method. Therefore, the samples we prepare to get are the blocks in video images, i.e. the samples are smoke block samples and non-smoke block samples.

For non-smoke block samples, we can get the samples by partitioning the non-smoke images into blocks. As shown in Fig. 4, the non-smoke block samples are the partitioning blocks from video non-smoke images.

For smoke block samples, it is difficult and impractical to get the samples from the video images directly. In this paper, we acquire the smoke block samples by synthesizing and partitioning smoke images from non-smoke images. And the synthetic smoke images can be obtained by using Eq.(2) based on two assumptions: ①image content is independent of scene depth or medium transmission, i.e. the same image content can appear at different depths in different images; ②depth is locally constant, i.e., image pixels in one small patch tend to have similar depth values. With these two assumptions, assuming $A=[1,1,1]$ (the image is normalized) and selecting the random values of t ($t\in(0,1)$), we can get a series synthetic smoke images as shown in Eq.(2).

In Fig.5(a) and (b), the synthetic smoke images at different values of t are given.

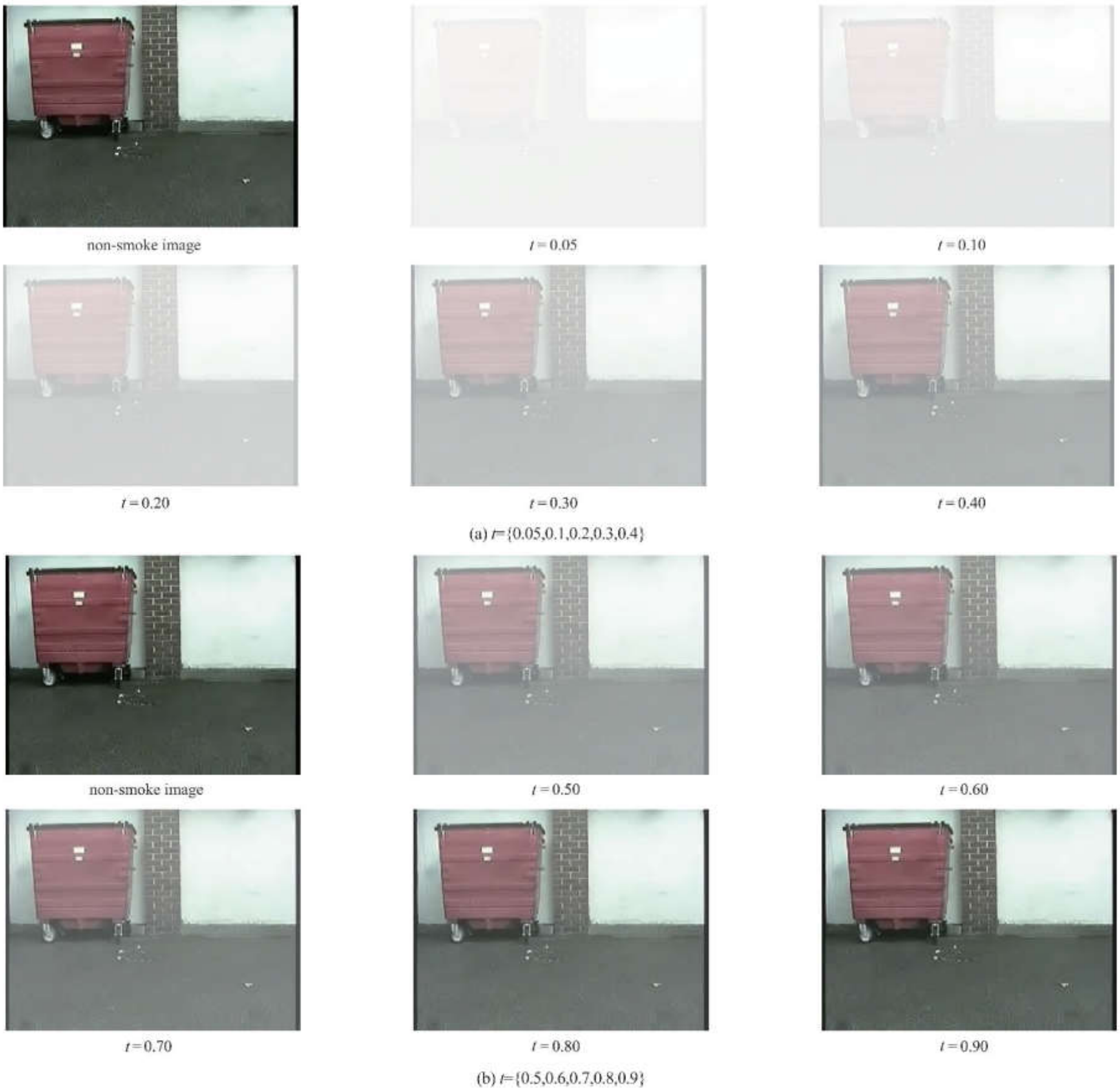

Fig.5 The synthetic smoke images in different t

As shown in Fig.6, the synthetic smoke images are more uniform than the real smoke images. This is because we assume the local regions have the same depth and the same medium coefficient t. This evenness of smoke cannot reflect the characteristic of real smoke images. However, if the values of t are randomly selected and ensure a certain quantity of t values, the synthetic smoke images could reflect the features of real smoke images in small block regions[10]. So we can get the smoke block samples by partitioning the synthetic smoke images into blocks.

In Fig. 7, the smoke block samples are the partitioning blocks from synthetic smoke images.

2.2 Features selection using random forest

With the block samples got from partitioning images and the features in Eq. (9), the necessary features can be selected by training the RF.

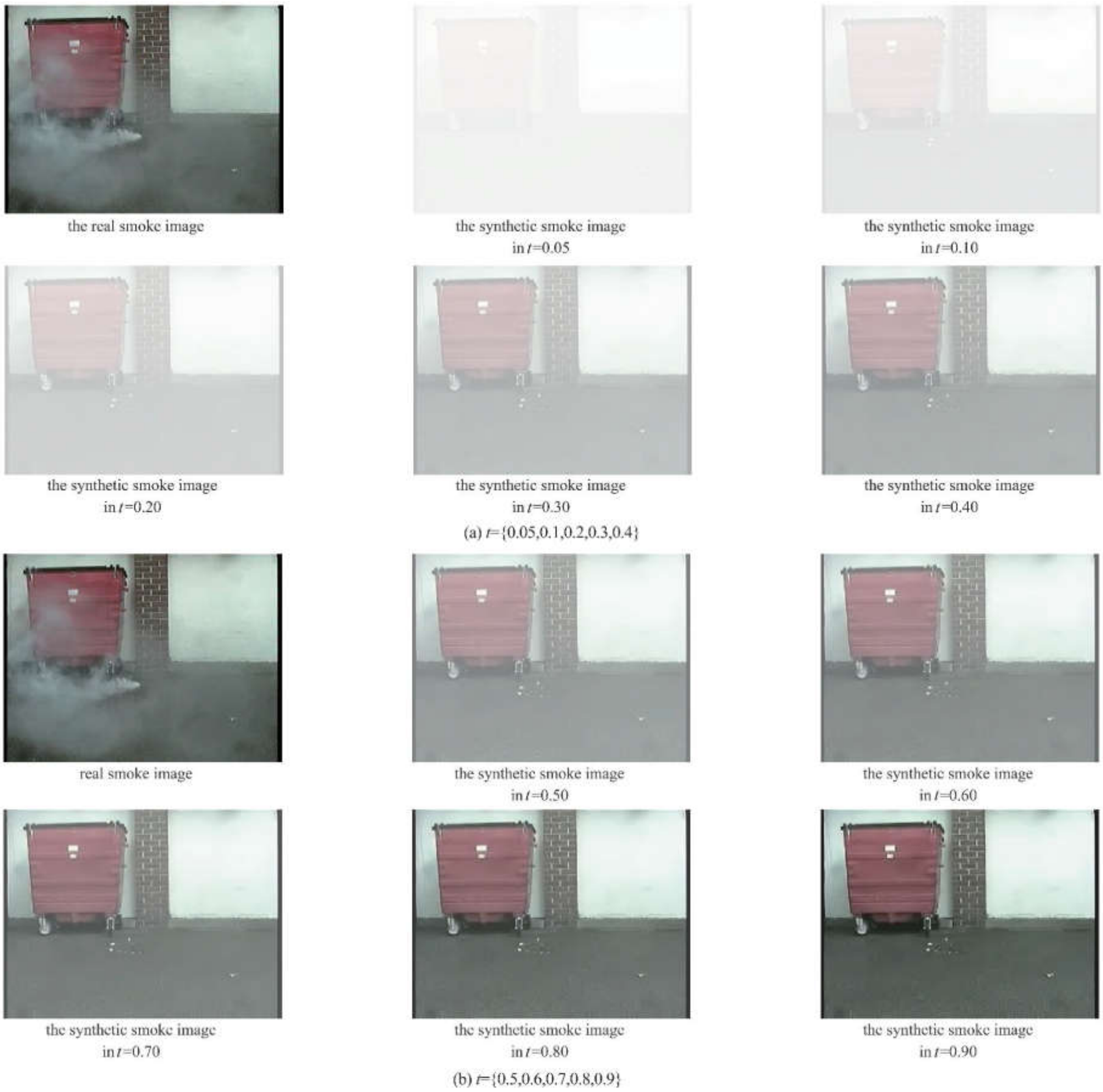

Fig.6　The synthetic smoke images and the real smoke image

$$\left.\begin{aligned} v_1 &= [R, G, B] \\ v_2 &= [cA, cH, cV, cD] \\ v_3 &= [D_1, D_4] \\ v_4 &= [S_1, S_4] \end{aligned}\right\} \tag{9}$$

In Eq.9 the wavelet features v_2 are extracted from the wavelet transform of luminance(Y) images, and the filters used are LPF=[0.25,0.5,0.25] and HPF=[−0.25,0.5,−0.25].

To train the RF[11], we capture non-smoke images from 10 videos, and partition the images into 120 000 blocks for the non-smoke block samples, each block size being 8×8 and marked as negative samples. Simultaneously, we synthesize 100 frames of the smoke images and also partition the images into blocks, and finally get 40 000 pieces of 8×8 blocks as the smoke block samples as well as positive samples. The features input RF are $v = [v_1, v_3, v_4, v_2]$ (Adjusting the feature sequence for a better analysis). Inputting the marked samples to RF for training and the feature selection process is as follows[12]:

Step 1　Inputing the training samples: N and the extracted features: M;

Step 2　Using GINI by coefficient to make a decision tree by random sampling N and M;

Step 3　Repeating step 1 to make K trees to constitute the forest;

Step 4　Calculating the classification error of out-of-bag data for each tree:

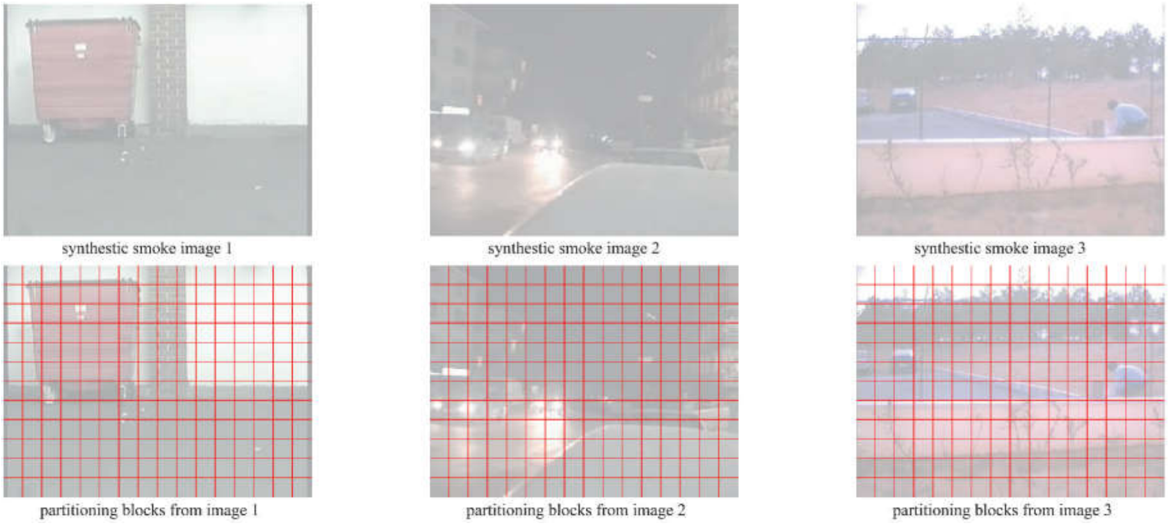

Fig.7 The smoke block samples in partitioned synthetic smoke images images

error OBB_1, error OBB_2, ⋯, error OBB_K

Step 5 Randomly changing the values of x_i (x_i donates the i^{th} attribute of M), re-calculating the classification error of out-of-bag data:

error OBB_1^i, error OBB_2^i, ⋯, error OBB_K^i

Step 6 Calculating the importance of feature attribute x_i:

$$\mathrm{Imp}^{x_i} = \sum \frac{\mathrm{errorOBB}_j^i \text{-} \mathrm{errorOBB}_j}{K}$$

step 7 Repeating step 6 to get the importance of all attributes and sorting in Imp, then outputting the most important features of the first m. The selected feature is in Fig.8.

As shown in Fig.8, by training the RF with the samples and the extracted features, we get the sorted features with feature importance. And in our experiment, we select the first m most important features to be the needed feature. Here, m is the values of $\lceil \sqrt{M} \rceil$ and M is the number of feature attributes.

2.3 Training SVM classifier

Training SVM with the partitioned block samples and the selected features from RF, we can get a classifier to recognize the smoke blocks and non-smoke blocks.

To train SVM, we select 30 000 non-smoke block samples (negative samples) and 10 000 smoke block samples (positive) as the training samples, and select 90 000 negative samples and 30 000 positive samples as the testing samples. In addition, we choose the radial basis function (RBF): $K_r(x, x_i) = \exp\left(-\frac{\| x - x_i \|}{2\sigma^2}\right)$ as the kernel function of SVM. Then, we train the SVM and get the block classifier by

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{N} \alpha_i K_r(x, x_i) - b\right) \tag{10}$$

where x_i are the support vectors, α_i is the parameter of hyperplane.

The training and testing results are in Tab.1. As shown in Tab.1, the detection rate of training smoke samples is 91.3% and the detection rate of testing smoke samples is 75.8%. Because of the two assumptions in synthesizing, the synthetic smoke images cannot completely reflect the information of real smoke images, and this leads the overfitting phenomenon in training SVM.

Tab.1 Training and testing results of SVM

	Training Samples	Testing Samples
Positive samples/ smoke samples	10 000	30 000
Detection numbers	9 130	22 740
Detection rate	91.3%	75.8%
Negative samples/ non-smoke sample	30 000	90 000
Elimination numbers	24 480	69 495
False rate	18.4%	22.8%

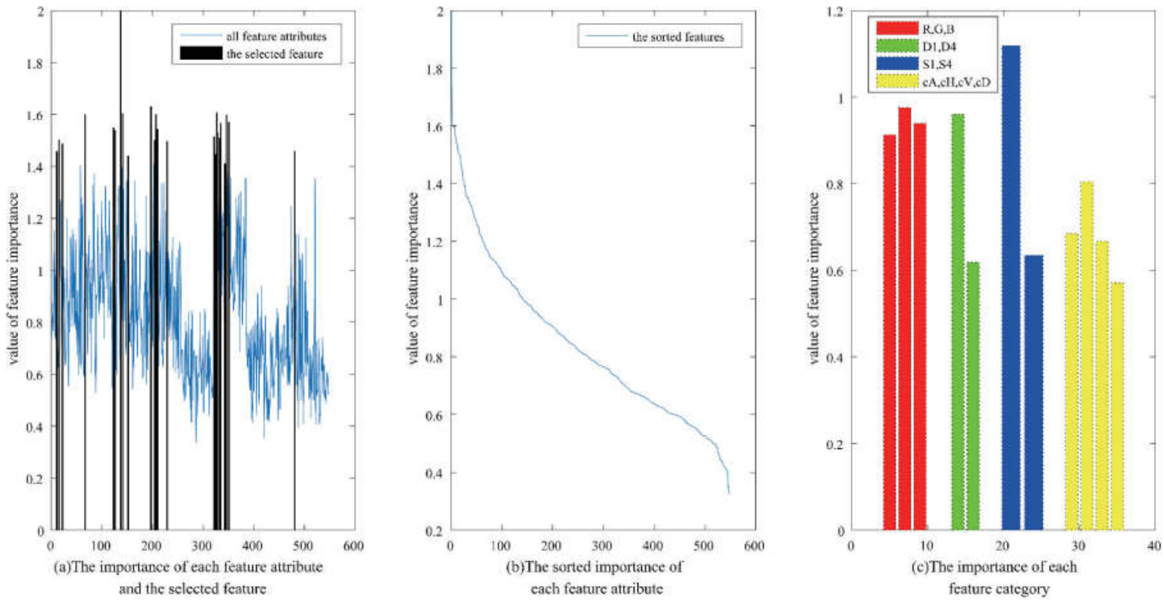

Fig.8　The features selected by Random Forest

3　The smoke regions extracted by classifier

According to the classifier obtained from training SVM, we can detect the smoke regions in video images. And the extracted smoke regions are shown in Fig.9.

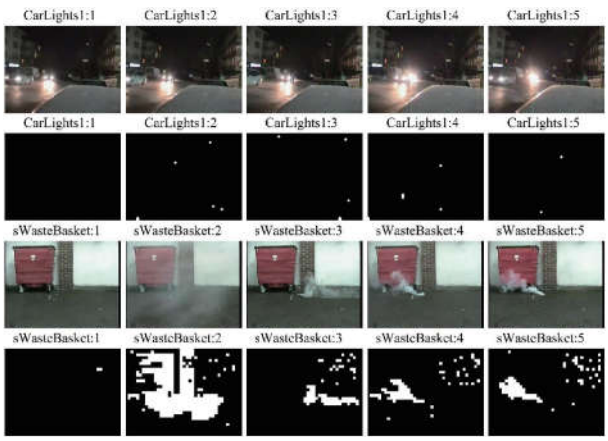

Fig.9　Smoke regions extracted by SVM classifier

As shown in Fig. 9, for non-smoke video CarLights1, hardly any none smoke regions are extracted, and the classifier shows good performance on this video smoke detection. But for the smoke video sWasteBasket, although the smoke regions can be extracted by the classifier in some way, some non-smoke regions in the video are also extracted as smoke regions, and there are always the white-wall areas being misjudged in the video. Actually, there are two reasons for the misjudgement between smoke regions and white walls. The first one is in training SVM, the Dark Channel features are involved to get the classifier, and the smoke regions and white walls have similar characteristics in video images about dark channel. The second one is that the smoke block samples used in training are synthesized under a very strong assumption which assumes the parameter $A=[1,1,1]$, and this makes the smoke block samples seem to be as white as the white wall areas in the video images. Therefore, the misjudgement is hardly avoidable. Fortunately, we can solve this problem below by analyzing the Convex Degrees and Growth Rate of extracted smoke regions.

4　The convex degrees and growth rate of the smoke regions

Airflows make the smoke diffused and disorder the smoke motion, which leads to the irregularity and the growth of smoke regions. We can define two parameters to describe these two characteristics. According to the Iso-perimetric Theorem, when the perimetric is given, the circle has the maximum area of all shapes. So, we can define the convex degrees in Eq.(11) to describe the contours of the smoke region. And obviously, the growth rate defined in Eq.(12) can reflect the growth of the smoke region.

$$p^a = \frac{p^2}{4\pi a},\ p^a \geqslant 1 \tag{11}$$

where p^a is Convex Degrees, p is the perimetric, a is the area, i. e. p^a is the square of perimetric to area ratio, and obviously, $p^a \geqslant 1$. We can utilize p^a.

$$g^r = \frac{s_t + \Delta t - s_t}{(t + \Delta t) - t} \tag{12}$$

where g^r denotes the smoke growth rate, s_t, $s_{t+\Delta t}$ is the area of the smoke region in time t and $t + \Delta t$, respectively.

Utilizing the classifier training from SVM, we can extract the smoke region $M_{ti}(x, y)$ from the $(t_i)^{\text{th}}$ frame image. And with the dilation operation of $M_{ti}(x, y)$, we can easily get the outer boundary of the corresponding smoke region. Using M_{ti}^{out} to denote the outer boundary, defining N_{ti} as the number of smoke pixels of smoke region $M_{ti}(x, y)$, and N_{ti}^{out} as the number of smoke pixels of the outer boundary $M_{ti}^{\text{out}}(x, y)$, we can calculate the convex degrees and the smoke growth rate as follows:

$$\left.\begin{aligned} p_{ti}^a &= \frac{(N_{ti}^{\text{out}})^2}{4\pi N_{ti}} \\ g_{ti}^r &= \frac{1}{t_{i+1} - t_i} \times (N_{ti+1} - N_{ti}) \end{aligned}\right\} \tag{13}$$

where p_{ti}^a denotes the Convex Degrees of the $(t_i)^{(\text{th})}$ frame smoke regions, g_{ti}^r is the smoke growth rate from the $(t_i)^{(\text{th})}$ frame to the $(t_{i+1})^{(\text{th})}$ frame, and $t_1 \leqslant t_2 \leqslant t_3 \leqslant, \cdots, \leqslant t_n$.

A single frame Convex Degrees and two successive sampling frames growth rate of the smoke region is haphazard for smoke detection, and we cannot confirm the existence of smoke in the video just with the results of p_{ti}^a and g_{ti}^r. To detect the smoke accurately, we sample the video frames in a certain sampling period, i.e. we sample the $(t_1)^{\text{th}}$, $(t_2)^{\text{th}}, \cdots$, $(t_n)^{\text{th}}$ frames in the video. Here, the value of n is depended on the frequency of $p_{ti}^a > \text{th}_{c1}$ and $g_{ti}^r > \text{th}_{d1}$, the process is shown as follows:

$$\begin{aligned} &s1:\ c_p = \text{count}(p_{ti}^a > \text{th}_{c1}),\ i = (1, 2, \cdots, n) \\ &s2:\ c_g = \text{count}(g_{ti}^r > \text{th}_{d1}),\ i = (1, 2, \cdots, n) \\ &s3:\ \begin{cases} \text{if } (c_p > c) \text{ and } (c_g > c) \text{ and } (t_n < t_{\text{video}}) \\ \text{save } t_n,\ \text{stop sampling, smoke exist;} \\ \text{else} \\ \text{no smoke, } t_n = t_{\text{video}}; \end{cases} \end{aligned} \tag{14}$$

where th_{c1}, th_{d1}, c are threshold values and depend on the statistical data of experiments. t_{video} is the duration of the video. c_p is the frequency of $p_{ti}^a > \text{th}_{c1}$, and c_g is the frequency of $g_{ti}^r > \text{th}_{d1}$.

In Eq. (14), we define t_n as the sampling termination parameter (or sampling termination frame). When there is smoke in the video, getting t_n means the end of sampling, and correct detection of smoke. When there is no smoke in the video, t_n lasts until the video is over. The defined sampling termination parameter is an important parameter for giving a fire alarm, especially in real-time situations.

Detecting smoke successfully doesn't mean getting a fire alarm. A sudden disappear once smoke or flashes of artificial light may cause false alarms. So we can use t_n obtained in Eq.(14) to calculate the average convex degrees and the average growth rate as follows:

$$\left.\begin{aligned} a_p &= \frac{1}{t_n - 1} \times \sum_{i=1}^{n-1} p_{ti}^a \\ a_g &= \frac{1}{t_n - 1} \times \sum_{i=1}^{n-1} g_{ti}^r \end{aligned}\right\} \tag{15}$$

where a_p and a_g are the average of convex degrees and growth rate, respectively.

$$\left.\begin{aligned} &\text{if: } (a_p > \text{th}_c) \text{ and } (a_g > \text{th}_D) \\ &\text{get fire-alarm;} \\ &\text{else} \\ &\text{no alarm.} \end{aligned}\right\} \tag{16}$$

where th_c and th_D are threshold values and depend on the statistical data of experiments. Based on the above conditions in Eqs.(13)～(16), we can detect smoke in the video and get a fire alarm.

5 Experimental results

The proposed method was implemented on a personal computer with an Intel(R)Core(TM)i3-3220, 3.30 GHz processor and detected smoke in offline video from http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SampleClips.html. In our experiment, the threshold values of c, th_c and th_D range from 4 to 10, 2 to 8 and 0.5 to 0.8, respectively. And the smoke detection and fire-alarm results are in Tab.2 and Tab.3.

In Tabs.2 and 3 where t_n denotes the sampling termination parameters, a_p is the average of convex degrees, a_g is the average of growth rate, f_{smoke} denotes detecting smoke or not, f_{fire} denotes fire-alarm or not, and t_{alarm} is the time(or frame) getting fire-alarm.

Tab.2 The values of the parameters in smoke detection and fire alarming

Video No.	Video Names	a_p	a_g	t_n
1	CarLights1	0.028 9	0.001 0	155
2	CarLights2	0.013 1	0.000 8	160
3	sBehindtheFence	3.277 1	0.62 50	32
4	sBtFence2	4.990 3	0.663 6	11
5	sEmptyR1	3.277 1	0.778 6	81
6	sEmptyR2	2.092 9	0.530 6	98
7	ShorterIsyamNight	0.628 6	0.026 0	255
8	sMoky	8.220 1	−0.062 9	5
9	sWasteBasket	2.462 0	0.749 6	14
10	sWindow	2.117 6	0.561 5	20

Tab.3 The results of smoke detection and fire alarming of video images

Video No.	Video Names	f_{smoke}	f_{fire}	t_{alarm}
1	CarLights1	no	no	——
2	CarLights2	no	no	——
3	sBehindtheFence	yes	yes	32
4	sBtFence2	yes	yes	11
5	sEmptyR1	yes	yes	81
6	sEmptyR2	yes	yes	98
7	ShorterIsyamNight	no	no	——
8	sMoky	yes	no	——
9	sWasteBasket	yes	yes	14
10	sWindow	yes	yes	20

As shown in Tabs. 2 and 3: ① The detection results of videos of CarLights1, CarLights2 and ShorterIsyamNight show no smoke detected in the videos, as is the case of the actual situation. And the sampling termination parameters of these videos are either their total number of frames or the maximum of the sampling frame(In our smoke detection, we set the maximum sampling frame at 255 for a real time detecting situation and when this maximum is exceeded, the sampling frame is re-initiated and smoke detection is restarted). Theoretically, the values of a_p and a_g are 0 for these non-smoke videos, but because of the assumption in $A=[1,1,1]$ results in some errors in the classifier and renders a_p and a_g close to 0 but not exactly 0. Fortunately, this has no big impact on the video smoke detection. ② The smoke dection result of the sWoky video shows that there is smoke in the video but no fire-alarm is given. Because there is smoke existing in the video at the very being, smoke is detected in the 5th frame(at about 0.33 s, the frame rate of the video being 15 frames per second), but since the value a_g is so close to zero and also a negative value, this indicates that the smoke regions are not growing but shrinking, so there is no fire-alarm given in the video smoke detection. ③ The videos that give both smoke and fire alarms include videos to 6 and 9 to 10. In these videos, smoke is detected and a fire-alarm is issued timely, the fire-alarm sampling frame being the same as the sampling termination parameters t_n. For instance, the video of sWasteBasket detects smoke in the 14th frame(about 1.4 s, the frame rate of the video being 10 frames per second) and issues a fire alarm. This well meets the requirements for real-time smoke detection.

In addition, in offline video smoke detection, the sampling termination parameters t_n is the sign of early fire alarm. The smaller the value of t_n, the earlier a fire-alarm is issued in the videos. In real time video smoke detection, t_n could be the sign to initialize sampling time. If t_n is too large, there could not be smoke in the videos from time 0 to t_n, and then, we initialize the original sampling time to 0 and restart smoke detection.

6 Conclusion

In this paper, a video smoke detection method with RF features selection was proposed. By selecting the features and analyzing the convex degrees and growth rate, we were finally obtained the smoke detection and fire-alarm results. And the experimental results show that: ①The proposed method can detect smoke in video images and get a fire alarm in an effective way, and the defined sampling termination parameter can reduce the false-alarm rate and provide support for real-time video smoke detection; ② The method of synthesizing smoke images decreases the difficulty of getting smoke samples; ③Analyzing the original extracted features automatically by training RF simplifies the SVM training process. However, the method we used to synthesize smoke images has two strong assumptions and this leads to the overfitting phenomenon in SVM training. And we cannot completely get rid of the dependence on experimental statistics because of the pre-set thresholds in analyzing smoke region Convex Degrees and growth rate. Therefore, the future work we should do is to find a better synthesized smoke image model and use the adaptive method to analyze the Convex Degrees and growth rate.

References

[1] CHENT H, YIN Y H, HUANG S F, et al. The smoke detection for early fire-alarming system base on video processing [C]// Proceedings of the International Conference on Intelligent Information Hiding and Multimedia. Pasadena, USA: IEEE Computer Society, 2006: 427-430.

[2] TÖREYIN B U, DEDEOǦLU Y, ÇETIN A E. Wavelet based real-time smoke detection in video[C]// European Signal Processing Conference. Antalya, Turkey: IEEE Press, 2005: 1-4.

[3] ALEJANDRO O B, LEONARDO M G, GABRIEL S P, et al. Improvement of a video smoke detection based on accumulative motion orientation model[C]// Electronics, Robotics and Automotive Mechanics Conference. Cuernavaca, Morelos, Mexico: IEEE Press, 2011: 126-130.

[4] YUAN F. A fast accumulative motion orientation model based on integral image for video smoke detection[J]. Pattern Recognition Letters, 2008, 29(7): 925-932.

[5] XU Z G, XU J L. Automatic fire smoke detection based on image visual features[C]// International Conference on Computational Intelligence and Security Workshops. Harbin, China: IEEE Press, 2007: 316-319.

[6] YANG J, CHEN F, ZHANG W D. Visual-based smoke detection using support vector machine[C]//Proceedings of the 4th International Conference on Natural Computation. Jinan, China: IEEE Press, 2008, 4: 301-305.

[7] HORNG W B, PENG J W. Image-based fire detection using neural networks[C]// Proceedings of the 9th Joint International Conference on Information Sciences. Kaohsiung, Taiwan, China: Atlantis Press, 2006, [2016-10-01]http://xueshu.baidu.com/s? wd=paperuri%3A%2890d519bf26ad093adab4b2c913f00ada%29&filter=sc_long_sign&tn=SE_xueshusource_2kduw22v&sc_vurl=http%3A%2F%2Fdx.doi.org%2F10.2991%2Fjcis.2006.301&ie=utf-8&sc_us=13120745438069482226.

[8] TUNG T X, KIM J M. An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems[J]. Fire Safety Journal, 2011, 46(5): 276-282.

[9] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.

[10] TANG K, YANG J, WANG J. Investigating haze-relevant features in a learning framework for image dehazing[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE Press, 2014: 2995-3002.

[11] CRIMINISI A, SHOTTON J, KONUKOGLU E. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning[J]. Foundations and Trends © in Computer Graphics and Vision, 2012, 7(2-3): 81-227.

[12] GENUER R, POGGI J M, TULEAU-MALOT C. Variable selection using random forests[J]. Pattern Recognition Letters, 2010, 31(14): 2225-2236.