Article

Flame Recognition in Video Images with Color and Dynamic Features of Flames

Jiaqing Chen, Xiaohui Mu, Yinglei Song*, Menghong Yu, Bing Zhang

School of Electronics and Information Sciences, Jiangsu University of Science and Technology, Jiangsu, Zhenjiang 212003, China

ABSTRACT

Recently, video based flame detection has become an important approach for early detection of fire under complex circumstances. However, the detection accuracy of most existing methods remains unsatisfactory. In this paper, we develop a new algorithm that can significantly improve the accuracy of flame detection in video images. The algorithm segments a video image and obtains areas that may contain flames by combining a two-step clustering based approach with the RGB color model. A few new dynamic and hierarchical features associated with the suspected regions, including the flicker frequency of flames, are then extracted and analyzed. The algorithm determines whether a suspected region contains flames or not by processing the color and dynamic features of the area altogether with a classifier, which can be a BP neural network, a k nearest neighbor classifier or a support vector machine. Testing results show that this algorithm is robust and efficient, and is able to significantly reduce the probability of false alarms.

Keywords: Fire Detection; RGB Color Model; Dynamic Features; Hierarchical Features; Feature Fusion

*CORRESPONDING AUTHOR

Yinglei Song, School of Electronics and Information Sciences, Jiangsu University of Science and Technology, Jiangsu, Zhenjiang 212003, China; syinglei2013@163.com;

1. Introduction

Real-time detection and early warning of fire is an important approach to alleviating the threats from fire hazards. Since fire often occurs randomly and a fire scene is usually complicated, traditional fire detection methods are often unable to detect fires and issue warnings in the early stages of fires. Recently, with the abundance of surveillance video cameras, fire detection technology based on video has become an important approach for the early detection of fire. Such detection methods analyze the features of video images and recognize potential occurrences of flames, fires can thus be recognized and under control before they develop into disasters[1]. Due to its ability to detect fires in their early stages, video based fire detection technology has attracted the attention of many researchers in the areas of fire safety and a large number of methods have been developed to detect the occurrence of flames by analyzing sequences of video images.

Color features are the most important features of flames and have been extensively used in methods for video based fire detection[2-9]. For example, a method of flame recognition based on RGB-HSI hybrid color model is developed in[2]. Based on a color model, regions that contain flames can be recognized based on the color features of flames. In[3], an approach that combines a statistical model with a background subtraction algorithm is developed to identify the regions that may contain flames. A HSI-YC_bC_r color model is used in[4] to analyze the color features of moving targets in a video and flame regions can then be recognized based on the results of analysis. Similarly, a generic color model[5], based on YC_bC_r color space with several rules, is proposed for flame pixels' classification. And in[6], potential fire regions are detected by modeling the

fire color with Gaussian mixture model. Most of these approaches develop color models to accurately describe the color features of flames and determine whether a suspected region contains flames or not by comparing its color features with those of flames. Dynamic features of flames provide another important criterion for the recognition of fires[10-15]. In[10], a Hidden Markov Model based approach is developed to detect the flicker dynamic features of a suspected region, interference from background can be effectively eliminated with this approach and an accurate description of the flicker dynamics of the suspected region can be obtained. In[11], an algorithm is developed to detect flames by combining the dynamic features extracted from the motion of moving targets with their color features described based on a CIE *L*a*b** color model. In[12], the dynamic features of flames are closely analyzed using the HSI color model. Several dynamic features, combining with color features, are used in[13] to train the dynamic SVM classifier, which is applied for fire alarm. In [14], analyses of growth rate and optical flow rate are applied to improve the accuracy of fire detection. However, due to the difficulty of accurately extracting and analyzing the dynamic features of moving objects in a video, dynamic features remains less popular than color features in flame detection.

In addition to color and dynamic features, features associated with the geometric shape of a segmented region are also used in the recognition of regions that may contain flames[16-20]. For example, in[16], the color, texture and contour features of a suspected region are obtained based on the HSI color model, and a neural network based classifier analyzes the features and identifies regions that contain flames. In[17], a Support Vector Machine based approach analyzes the static features related to the circular degree of a suspected region and the color features extracted with a RGB-HSI mixture model are combined with these features to recognize flame areas. Similarly, the color, texture and scintillation features are combined in[18] with an information fusion based algorithm to detect flames. Testing results show that the flame regions can be efficiently recognized with this information fusion based approach.

So far, a large number of flame detection methods have been developed and the accuracy and reliability of recognition have been significantly improved. However, since the dynamic features of flames have not been closely analyzed in most of these methods, the recognition accuracy is still not completely satisfactory under complex circumstances. In addition, the color features related to the spatial distribution of the color of pixels in a flame region have not been fully studied and utilized in the recognition of flames.

In this paper, we develop a new approach that combines the color, dynamic and hierarchical features of flames to improve the recognition accuracy. Firstly, the areas that may contain flames are accurately obtained with a new clustering based two-step approach. A few new dynamic features related to the change rates, flame similarity, and flicker frequencies of flames are extracted by analyzing the regions obtained in the first step. In addition, the spatial distribution of the color of pixels in a flame region is also analyzed and the hierarchical features of a flame region are then obtained based on the analysis. Flame regions are finally determined by processing all features with a classifier, which can be a BP neural network based classifier, a k nearest neighbor classifier or a support vector machine.

In order to test the recognition accuracy, we test the recognition accuracy of our approach on 10 videos. Five of these videos contain flames of forest fires; two of them do not contain flames and the remaining two contain flames that are under control and in safe conditions. In addition, we also use the approaches developed in[2,3,22], and[23] to process the testing videos and compare their recognition results with those of our approach. Our testing results show that, compared with four other approaches, our approach achieves significantly improved recognition rates in videos that contain flames. On the other hand, the false alarm rates of our approach on videos that do not contain flames or flames in safe conditions are significantly lower than those of the four other approaches.

2. The Recognition of Flame Regions

Since the effective identification of suspected regions that may contain flames is the basis for the

subsequent recognition of flames, an algorithm is needed to accurately determine the regions that are likely to contain flames in a sequence of video images. Color features of flames are used as the criteria for the identification of these regions.

2.1 The Recognition of Flame Pixels

The colors of pixels in flame regions in a video image follow a special distribution in the RGB color space[2,3]. In RGB color space, the color of a pixel can be described by the values of three channel components, namely R, G and B. In most cases, a high value of the R channel component is characterized by high brightness and high saturation. Therefore, for any flame pixel in (x, y) in an image, the following conditions should be satisfied:

$$R(x, y) \geq R_t \quad (1)$$

$$R(x, y) > G(x, y) > B(x, y) \quad (2)$$

where R_t is a threshold value that is often set to be 180. And $R(x, y)$, $G(x, y)$, $B(x, y)$ are the values of red, green and blue channel components for the pixel at (x, y).

However, the above screening criteria based on the RGB color model are only preliminary and may result in a large number of false positives. It is well known that the color information of each pixel is described by its R, G, and B channel components, which can be used to convert the representation of color in RGB space to that in other color spaces, such as HSI space, where the brightness and saturation information can be well separated. However, the conversion may also lead to the loss of image information. In this paper, we consider pairs of R-G, R-B, and G-B channel components in the RGB space, a statistical model is constructed based on these pairs to identify the potential flame pixels. According to the data provided in[3], we use the criterion as shown in equation (3) to process the pixels that satisfy equations (1) and (2) and only those that also satisfy the criterion are considered to be a flame pixel.

$$\begin{cases} 0.25 \leq G(x, y)/(R(x, y)+1) \leq 0.65 \\ 0.05 \leq B(x, y)/(R(x, y)+1) \leq 0.45 \\ 0.20 \leq B(x, y)/(G(x, y)+1) \leq 0.60 \end{cases} \quad (3)$$

To simplify the notation, we use *firep*(x, y) in the rest of the paper to characterize whether the pixel at (x, y)

is a flame pixel or not. *firep*(x, y) is 1 if the pixel is a flame pixel, otherwise it is 0. To evaluate the recognition accuracy of this approach, we compare the recognition result of this approach with those of the approaches developed in[2,10], and[17-19]. Most of these approaches recognize flame pixels with RGB and HSI based models. The comparison is made among our approach, the original model and RGB-HSI mixed model. The approach based on the RGB-HSI mixed model uses a constraint condition that depends on a threshold, which is computed based on the RGB model. Details of the method can be found in[3].
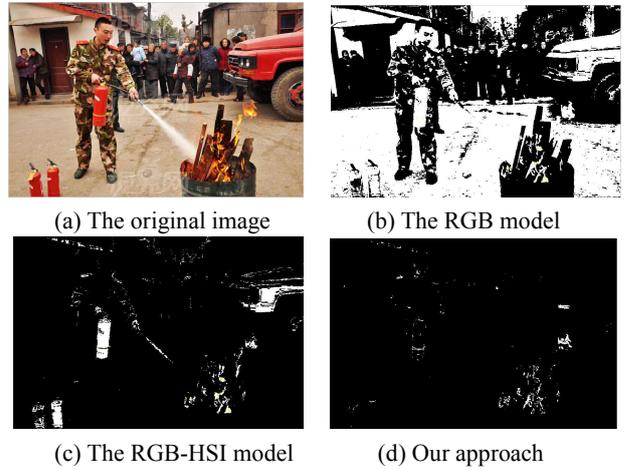

(a) The original image   (b) The RGB model

(c) The RGB-HSI model (d) Our approach

Figure 1. A comparison of the recogntion results with three different approaches

Figure 1 (**b**), (**c**) and (**d**) show the recognition results obtained with the three approaches respectively on the same image shown in **Figure 1(a)**. It is evident from the figure that the background in the original image is complex and contains many objects whose color is similar to the color of flames. It can be seen from the figure that the recognition result obtained with the original RGB model is the worst. In addition, our approach achieves the highest recognition accuracy. However, it is also clear that our approach fails to identify the center of the flame.

2.2 Refinements of recognized flame regions

It can be seen from the results shown in **Figure 1** that our approach is unable to accurately recognize the center of a flame region. The recognized flame regions thus need to be refined to further improve the recognition accuracy. We use a method often used in clustering algorithms to find pixels that are close to the fire pixels

that have been found in color and these pixels are very likely to be part of a flame.

Firstly, the mean value of the channel components of all the flame pixels that have been detected can be calculated with equation (4) as follows.

$$\begin{cases} \mu_R = \dfrac{1}{K}\sum_{i=1}^{K} R_f(x,y) \\[2mm] \mu_G = \dfrac{1}{K}\sum_{i=1}^{K} G_f(x,y) \\[2mm] \mu_B = \dfrac{1}{K}\sum_{i=1}^{K} B_f(x,y) \end{cases} \tag{4}$$

where $R_f(x,y)$, $G_f(x,y)$ and $B_f(x,y)$ are the values of R, G, and B channel components for the flame pixel at (x,y), K is the total number of flame pixels that have been detected.

Each pixel that is not determined to be a flame pixel is then processed by comparing the values of its three channel components with those of μ_R, μ_G and μ_B. The $dev(x, y)$ computed with Equation (5) is used to characterize the difference between the three channel components of such a pixel at (x, y) and the flame pixels that have been discovered, where $R_{nf}(x, y)$, $G_{nf}(x, y)$ and $B_{nf}(x, y)$ represent the R, G, and B channel components of the pixel respectively.

$$dev(x,y) = \left| R_{nf}(x,y) - \mu_R \right| + \left| G_{nf}(x,y) - \mu_G \right| + \left| B_{nf}(x,y) - \mu_B \right| \tag{5}$$

As shown in Equation (6), the value of $R_{nf}(x, y)$ is compared with μ_R, the pixel is set to be a flame pixel if the $R_{nf}(x, y)$ is larger than or equal to μ_R since such a pixel has a brightness value larger than the average brightness value of all flame pixels that have been detected. Otherwise, $dev(x, y)$ is compared with a threshold value τ and the pixel is considered to be a flame pixel if $dev(x, y)$ is less than τ.

$$\begin{cases} if\ R_{nf} \geq \mu_R \quad firep(x,y) = 1 \\[2mm] if\ R_{nf} < \mu_R \quad firep(x,y) = \begin{cases} 1 & if\ dev < \tau \\ 0 & others \end{cases} \end{cases} \tag{6}$$

Our testing results show that the recognition accuracy can be optimized in practice when the value of τ is around 30.

The approach described above is able to find pixels that are not detected to be a flame pixel based on equation (3) but are likely to be part of flames in the image. The detected flame regions can thus be refined after the approach is applied to an image. Since the flame region in the original image in **Figure 1**(**a**) is small, the effect of refinements created by our approach is not obvious. The image in **Figure 2** (**a**) is thus used to test the actual effect of the refinements. **Figure 2** (**b**) shows the flame regions recognized by our approach before the steps of refinements are applied and the flame regions obtained after the steps are applied to the image are shown in **Figure 2** (**c**). It can be seen clearly from **Figure 2** that the above approach for refinements can successfully detect the flame pixels that are missing in the results generated by preliminary recognition.
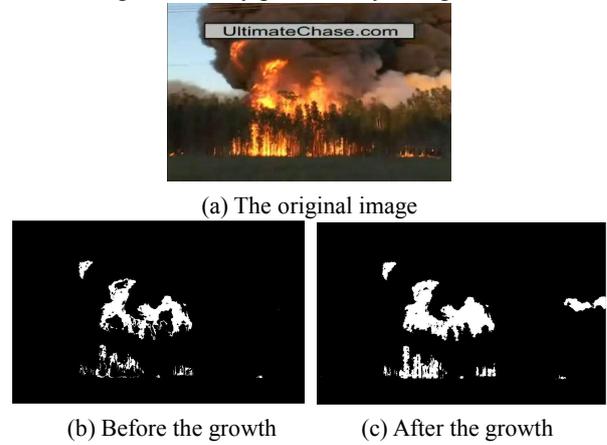

(a) The original image

(b) Before the growth   (c) After the growth

Figure 2. The effect of flame region refinements

3. Feature Extraction and Analysis

Flame recognition that only uses the color features of flame pixels is simple and fast, it can achieve satisfactory recognition results on some video images. However, since the background of a scene monitored by video cameras is generally more complex. Different types of interference sources thus may appear in the video images. These interference sources may cause a high false alarm rate if we only consider the color features in flame detection. Therefore, other types of features of flames are also needed to be extracted and analyzed to improve the recognition accuracy.

For example, many researchers have focused on the study of features related to the geometric shapes of flame regions, such as roundness, cusp number[17-19]. However, similar to the color features, these features are also static features and do not contain the information on the

motion of flames. In this section, we develop new features to describe the dynamics and hierarchical structures of flames, which can be combined with the color features to significantly improve the recognition accuracy. The dynamic features we consider include the change rate of flame areas, flame similarity and flame flicker frequency. The hierarchical features characterize the distribution of the difference between the R channel components of two adjacent flame pixels.

3.1 The Change Rate of Flame Areas

When a fire breaks out, it generally experiences a process from scratch to extinction, which is in accordance with the Gauss distribution. In the early stage of fire, the fire flames demonstrate a trend of continuous growth. On the other hand, the areas of many interference sources, such as candles, lights etc remain relatively stable. The change rate of flame areas can thus be used as an important criterion to recognize flames that are expanding with a certain rate.

The flame area change rate can be computed based on the change of the number of flame pixels in two frames that are adjacent in the video. Specifically, if $I1$, $I2$ represent the two frame images adjacent in the video, N_{I1} and N_{I2} represent the total number of flame pixels that have been detected in $I1$ and $I2$, the change rate of flame areas at $I2$ can be computed as follows.

$$A_r = \frac{\left|N_{I2} - N_{I1}\right|}{N_{I1}} \quad (7)$$

(a) Screenshot of fire   (b) Screenshot of sunset

**Figure 3.** Screenshots of videos of a fire scene and a sunset scene

To compare the change rates of flames and other interference sources, we select two videos and compute the change rates of flame areas for each of them. One of the two videos records a fire scene while a sunset scene is recorded in the other one. The frame rates of both videos are 25 frames per second and the change rate of flame areas for the first 75 frames of each video are computed for the comparison. **Figure 3** (**a**) and (**b**) show screenshots of the two videos respectively while **Figure 4** (**a**) and (**b**) show the areas of flame regions computed for the first 75 frames in both videos respectively.

It can be seen clearly from **Figure 4** that the area of a flame region changes in a fashion completely different from that of a suspected region without real flames. In **Figure 4** (**a**), the fire spreads, the area of the flame region gradually becomes larger. Finally, the area of the flame region tends to be stable. On the other hand, **Figure 4** (**b**) shows that the area of the suspected region in the sunset video undergoes two different stages. Although the area of the suspected region in each stage remains stable, a sudden change in the area of the suspected region is observed when a switch between two stages occurs.
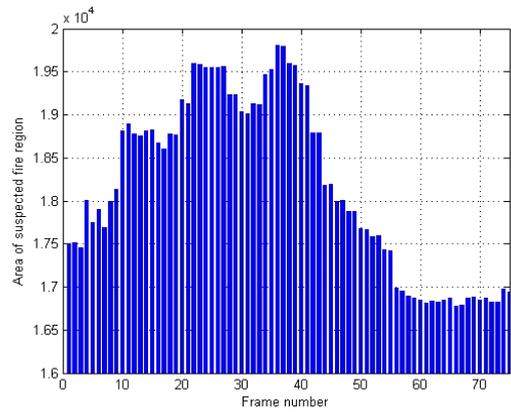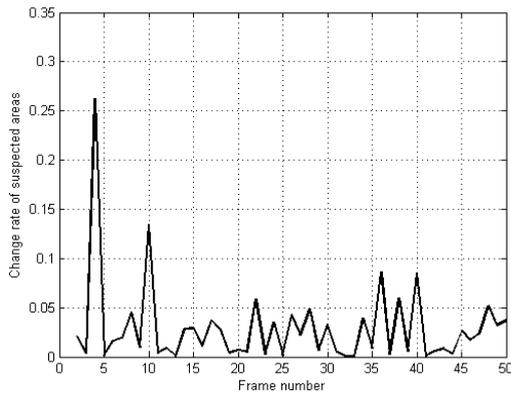

(a) The fire scene
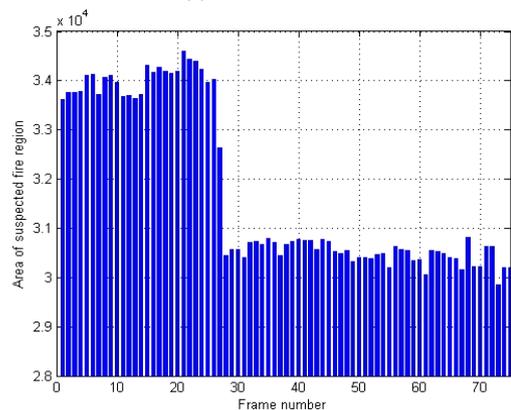
(b) The suset scene

Figure 4. The areas of the suspected regions in the selected 75 frames of the videos of a fire scene and a sunset scene.

Figure 5 compares the change rates of flame areas computed for the selected 75 frames in both videos. It is

evident from **Figure 5** (**a**) that the area change rate is in the range of [0.01,0.27] during the observation period. The value of the area change rate is relatively larger when the fire starts and spreads. However, its value is gradually stabilized and maintains at around 0.05 after the fire stops its expansion. On the other hand, as shown in **Figure 5** (**b**), the area change rate for the sunset scene is a small value in the range of [0.001,0.04]. In addition, its value remains relatively stable and is less than 0.015 in most of the observation time.

(a) The fire scene
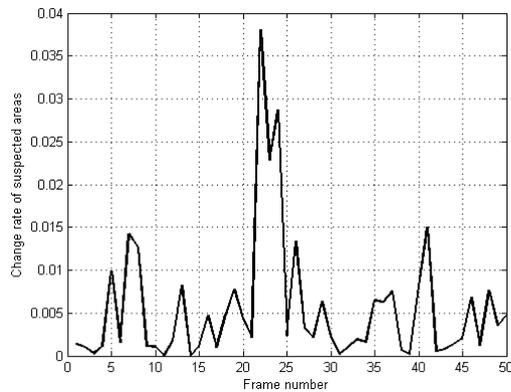
(b) The sunset scene

Figure 5. A comparison of the change rate of suspected areas in the videos shown in Figure 3
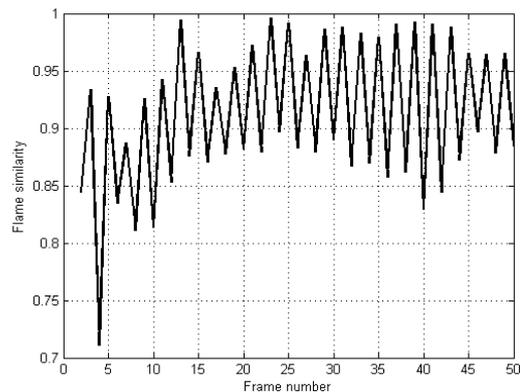
3.2 Flame Similarity

The shape similarity of flame in the adjacent two frames is also found to follow certain principles. Our experiments on objects such as sun, lamp, street lamp etc show that the value of the shape similarity varies in different ranges for different objects. We use $I_{i-1}(x, y)$, $I_i(x, y)$ to denote the binary images obtained from the images of frames i-1 and i respectively, where $I_{i-1}(x, y)$ or $I_i(x, y)$ is 1 if the pixel at (x, y) is determined to be a flame pixel in frame i-1 or i and their values are 0 if otherwise. The flame similarity between I_{i-1} and I_i can be computed with Equation (8) as follows.

$$\alpha = \frac{\sum\limits_{(x,y)\in\Omega} \overline{I_{i-1}}(x,y) \cap \overline{I_i}(x,y)}{\sum\limits_{(x,y)\in\Omega} \overline{I_{i-1}}(x,y) \cup \overline{I_i}(x,y)}$$

(8)

Where Ω is the set of all pixels in I_{i-1} or I_i, $\cap$ is an operator for "and" and $\cup$ is an operator for "or". It is not difficult to see that the numerator in the above equation is the number of pixels that are recognized to be flame pixels in both I_{i-1} and I_i, while the denominator is the number of pixels that are recognized to be flame pixels in one of I_{i-1} and I_i.

To show how the flame similarity varies in videos that contain flames and compare the results with those obtained on videos that do not, the two sets of videos that have been shown in **Figure 3** are tested, and the first 50 frames of each video are analyzed. **Figure 6** shows the values of the flame similarity obtained on these frames in the two videos. As shown in **Figure 6**, the flame similarity of adjacent frames in a fire video may vary in a wider range. **Figure 6** (**a**) shows that the values of the flame similarity between adjacent frames varies in a range of [0.8,1]. By contrast, **Figure 6** (**b**) shows that and the similarity values between two adjacent frames in the video of the sunset scene is mainly concentrated in the range of [0.95,1].

(a) The fire scene
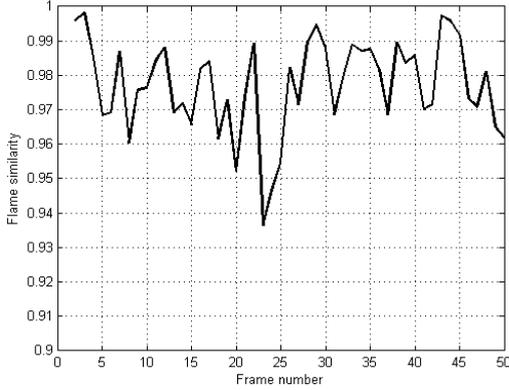
(b) The sunset scene

Figure 6. The flame similarity values obtained on the two videos shown in Figure 3

3.3 Flame Flicker Frequency

The flicker of flame appears to be random and irregular. But it in fact has special spectrum characteristics. Recent work has revealed that the flame flicker frequency is mainly concentrated in the range of 7-12 Hz[18,21], and has nothing to do with the flame size. In a short period of time, when the influence of external factors such as wind power can be neglected, the motion of flames is relatively stable, and the distribution of flame pixels is determined by the flickering motion of flames. Therefore, we start the analysis with the distribution of the difference between flame pixels in two adjacent frames, and we show later that the flame flicker frequency can be characterized based on this distribution.

Firstly, the change in the distribution of flame pixels in two adjacent frames i-1 and i can be computed with Equation (9), we use $P_r(i)$ to denote the change. The variables in equation (9) are defined the same way as they are defined in section 2.2. In the equation, the numerator represents the number of the pixels that are determined to be flame pixels in frame i but not in frame i-1, while the denominator is the total number of flame pixels in frame i-1. **Figure 7** shows the values of $P_r(i)$ computed for the selected 50 frames in the fire video in **Figure 3 (a)**.

$$P_r(i) = \frac{\sum\limits_{(x,y)\in\Omega} \overline{I}_i(x,y) - \sum\limits_{(x,y)\in\Omega} \overline{I}_{i-1}(x,y) \bigcap \overline{I}_i(x,y)}{\sum\limits_{(x,y)\in\Omega} \overline{I}_{i-1}(x,y)}, \; i \geq 2$$
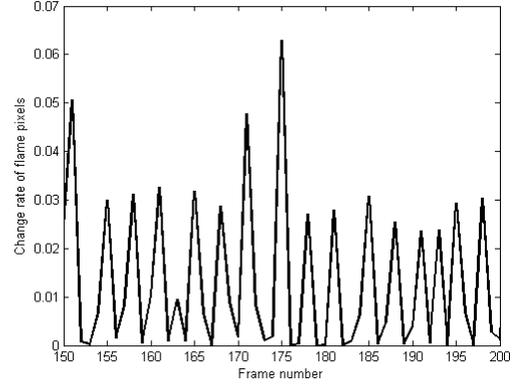
(9)

Figure 7. The computed values of $P_r(i)$ for the selected 50 frames in the fire video shown in Figure 3(a)

We then use the zero detection method to compute the flicker frequency of flames, the method needs to determine the zero point of $P_r(i)$. Specifically, the mean value μ_P of all $P_r(i)$s can be computed with Equation (10). We choose the mean value μ_P as the zero point, a statistics related to the flame flicker frequency can then be computed with Equation (11). It is not difficult to see that Equation (11) computes the number of times where the value of $P_r(i)$ turns from a negative number into a positive one. The value of sgn(x) is 1 if x is positive and -1 otherwise. Note that the value of $P_r(1)$ is not defined, we set $P_r(1) = 0$ for the convenience of computation.

$$\mu_P = \frac{1}{N-1}\sum_{i=2}^{N} P_r(i)$$

(10)

$$\begin{cases} N_f = \frac{1}{2}\sum_{i=2}^{N} \left| \text{sgn}(P_r(i) - \mu_P) - \text{sgn}(P_r(i-1) - \mu_P) \right| \\ \text{sgn}(x) = \begin{cases} +1 & (x \geq 0) \\ -1 & (x < 0) \end{cases} \\ P_r(1) = 0 \end{cases}$$

(11)

After N_f is computed, the flicker frequency f of flames in the video observation time can be obtained with Equation (12):

$$f = \frac{F_{FPS} * N_f}{N-1}$$

(12)

where F_{FPS} represents the video frame rate, N is the total number of frames where the analysis is performed.

To assess the effectiveness of the above method for

36

flicker frequency calculation, we use the method to evaluate the flicker frequency of the potential flame regions in six different videos as shown in **Figure 8.** Videos shown in Figure 8 (a), (b), (c), (d) contain flames while the remaining two only contain mobile headlights and a sunset scene respectively. In addition, due to the fact that the flicker frequency of flames is around 7-12 Hz, the frame rate of the videos must be a value of at least 24 frames per second from the Nyquist sampling theorem. Therefore we set the video frame rate to be 25 frames per second for all six videos. The obtained experimental results are shown in **Tables 1** and **2**.

(a) Video ①   (b) Video ②

(c) Video ③   (d) Video ④

(e) Video ⑤   (f) Video ⑥

Figure 8. Screenshots of videos tested for flame flicker frequency

Video Number	Frame Rate(frames/s)	Flicker Frequency(Hz)
Video ①	25	8.5
Video ②	25	7.8
Video ③	25	7.5
Video ④	25	7.8

Table 1. The flicker frequencies of flame regions in videos that contain flames

Video Number	Frame Rate(frames/s)	Flicker Frequency(Hz)
Video ⑤	25	4.8
Video ⑥	25	6.0

**Table 2.** The flicker frequencies of suspected regions in videos that do not contain flames

It is clear from **Table 1** that the flame flicker frequency is in the range of 7-12 Hz. By contrast, **Table 2** shows that the flashing frequencies of the automobile headlights and the setting sun are 4.8 Hz, 6.0 Hz respectively, and are not in the range of the flame's.

To further demonstrate the difference in the flicker frequencies of fire flames and other interference sources, we select Video ① in **Figure 8 (a)** and Video ⑤ in **Figure 8 (e)** for a closer analysis. For comparison, each group that contains 25 frames (i.e. 1s) is chosen as a sampling interval, a flicker frequency is then calculated for each group. **Figure 9 (a)**, (**b**) show the flicker frequencies computed for a few groups from Videos ① and ⑤ respectively. To add that, in practice, regions with high brightness may appear suddenly in a video of moving vehicles, which may generate a significant amount of interference for fire recognition, so we choose Video ⑤ for further testing. We are able to see clearly from **Figure 9** that the flicker frequencies remain relatively stable. Due to the disturbance from the smoke in the video, flicker frequencies fluctuate around a value of 8 Hz. By contrast, the flicker frequencies of vehicle headlights in Video ⑤ are always lower than 6 Hz (since Video ⑤ lasts only for 5 seconds, **Figure 9 (b)** has less sampling points).
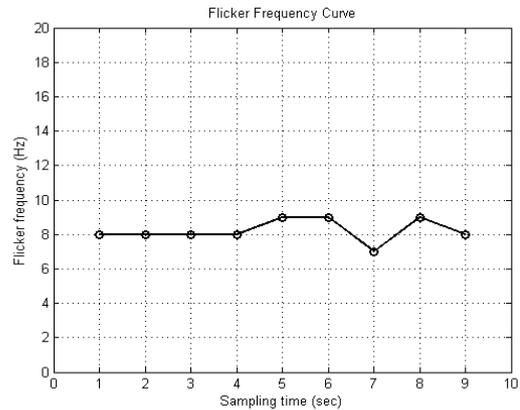
(a) Flicker frequencies obtained on video ①
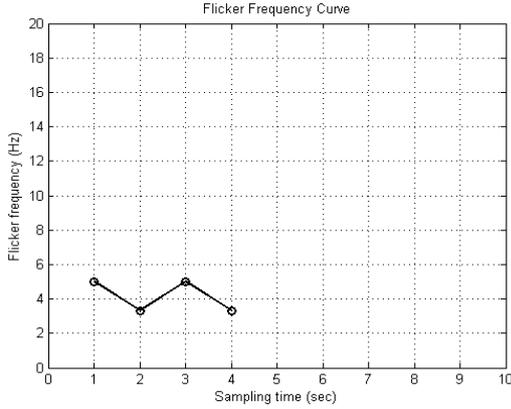
(b) Flicker frequencies obtained on video ⑤

Figure 9. A comparison between the flicker frequencies obtained on video ① and video ⑤

3.4 The Statistics of Gray Difference of R Channel Component

The layering features of flames are used to describe the percentages of flame pixels in different gray levels, they are often used as a criterion to strengthen the flame recognition results, so that the interference from other sources can be eliminated and the recognition accuracy can be improved.

For a given pixel p at (x, y), we consider the four pixels in the neighborhood of p and compute the difference between the R channel component of p and that of each of the four pixels. The distribution of these difference values for all pixels in a flame region is then used to obtain the layering features of the flame region. **Figure 10** shows a pixel at (x, y) and the surrounding 4 neighborhood pixels $(x-1, y)$, $(x+1, y)$, $(x, y-1)$ and $(x, y+1)$ to compute the differences of R channel components of adjacent flame pixels.
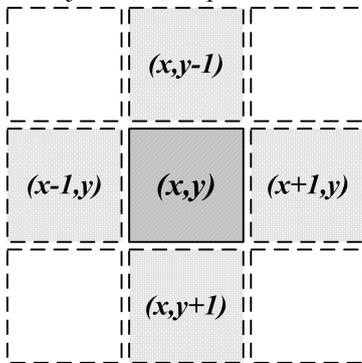

Figure 10. The selection of pixels to compute the differences of R channel components in the surrounding neighbor of the pixel at (x, y)

Since the inner regions of the flame are generally connected, a single pixel that is determined to be a flame pixel is often due to interference. We therefore simply omit these pixels and only analyze the regions that contain a large number of flame pixels. For a flame pixel at (x, y), the four pixels that are adjacent to (x, y) are investigated and the flame pixels in them are grouped into a connected region (i.e., a flame block). The difference between the R channel components of each pair of adjacent pixels in the group is then computed. Finally, the mean of all these difference values is computed by using Equation (13).

$$\mu_D = \frac{1}{M} \sum_{(x,y) \in \Lambda} \left| R(x, y) - R^{'} \right|$$

(13)

where μ_D is the mean of the differences of channel R components of all adjacent pairs of flame pixels, M denotes the total number of the connected regions as described above, Λ represents all the remaining flame regions after single flame pixels are removed, $R^{'}$ is the R channel component of the flame pixel that is adjacent to (x, y). $R^{'}$ is computed with Equation (14).

$$R^{'} = \begin{cases} R(x-1, y) & if \ firep(x-1, y) = 1 \\ R(x+1, y) & if \ firep(x+1, y) = 1 \\ R(x, y-1) & if \ firep(x, y-1) = 1 \\ R(x, y+1) & if \ firep(x, y+1) = 1 \end{cases}$$

(14)

where the value of *firep*(x, y) is 1 if the pixel at (x, y) is determined to be a flame pixel. Otherwise, its value is 0.

We now use the two videos shown in **Figure 3** to test whether the layering features of flame regions are different from those of suspected regions that do not contain flames. **Figure 11** (**a**) and (**b**) show the values of μ_D computed on frames in the videos that record a fire scene and a sunset scene respectively. **Figure 12** (**a**) and (**b**) show the histograms that describe the distributions of the difference values in both videos. Since a threshold R_t of value 180 is used to obtain the suspected flame regions in both videos and the gray value of the R component is in the range of [0, 255], the horizontal coordinates in **Figure 12** (**a**) and (**b**) vary in a rang of [0, 75].

It can be seen from **Figure 11** (**a**) that the value of μ_D is larger when the fire starts and becomes stable afterwards, it fluctuates around 5 in **Figure 11** (**a**). Our experiments on other available videos of fire scenes also

38

show that although the stable value of μ_D is not a constant and may vary in different fire scenes, the values of μ_D change in a fashion similar to the one shown in **Figure 11** (a) for all these videos. Therefore, this property is highly likely to be an important layering feature for flame regions. By contrast, **Figure 11** (b) shows that the value of μ_D is small and varies in the range of [1.2,1.45], which is quite different from what we have observed in videos of fire scenes, the sunset scene can thus be excluded from fire flames based on its layering feature.

From **Figure 12** (a) and (b), we are able to see that the difference values in a flame region are distributed in a larger range than that of a region that does not contain flames. The difference value in a flame region could be up to 75. The counts on values in interval [0, 75] gradually decrease and the rate of decline is smaller in **Figure 12 (a)**. For R gray level difference score of 1, its count in the fire scene is 2.4×10^4, which accounts for about 60 percent of the total counts. On the other hand, the count in the sunset scene is 7.2×10^4, which accounts for 80 percent of the total counts. This shows further that fire flames and sunshine interference sources are also different in this particular layering feature.
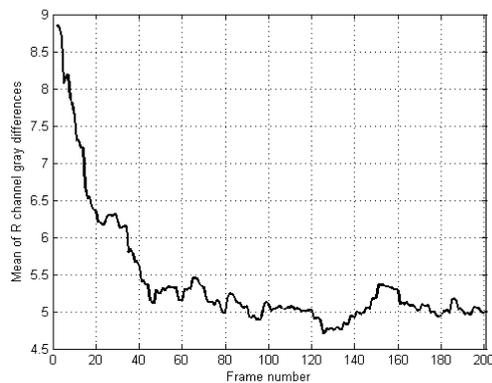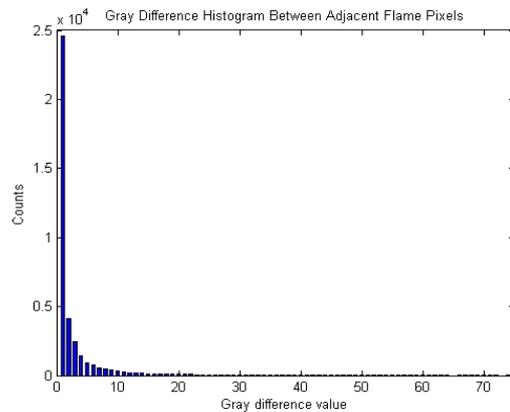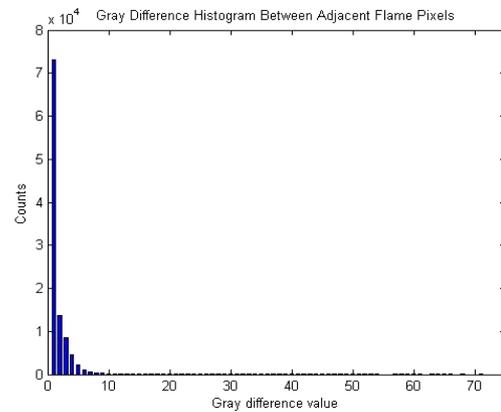

(a) The fire scene
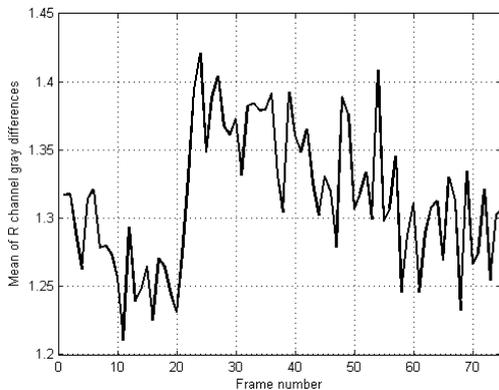
(b) The sunset scene

Figure 11. A comparison of the values of μ_D computed in videos shown in Figure 3

(a) The fire scene
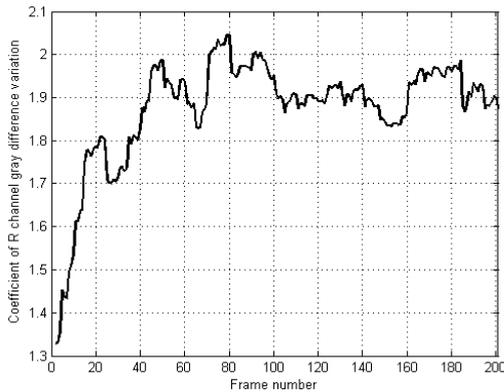

(b) The sunset scene

Figure 12. Histograms on the distribution of the difference of R channel components between adjacent flame pixels in the videos shown in Figure 3

3.5 The coefficient of R channel gray difference variation

Since the Coefficient of Variance (CV) provides a better measure of the degree of data dispersion than the variation itself, the differences between the R channel components of adjacent flame pixels are used to obtain CV, which is another characteristic parameter that can be used to recognize flames. Given the distribution of a random variable, the value of its CV is computed by dividing the mean value of the variable by its standard deviation. Here, for the convenience of calculation, we compute the square of CV as the value of the measure, in other words, we divide the square of the mean by the variance, as shown in Equation (15):

$$CV = \frac{\frac{1}{M} \sum_{(x,y) \in \Lambda} (\left| R(x,y) - R' \right| - \mu_D)^2}{\mu_D^2}$$

$$(15)$$

where all variables are as defined in Section 2.4. Here, we use the data obtained in Section 2.4 and select the videos in **Figure 3** to obtain the values of CVs for videos of a fire scene and a sunset scene. **Figure 13 (a)** and **(b)** show the values of CV obtained on all frames in the two videos respectively. From **Figure 13 (a)**, it is clear that the value of CV increases sharply when the fire starts and becomes stable at around 1.9 after 40 frames. In addition, the CV of fire flames is mainly concentrated in the range of [1, 2.5] and the amplitude of fluctuations is around 0.1. By contrast, **Figure 13 (b)** shows that the values of CV in the video of a sunset scene vary in the range of [3.8, 4.8] and the amplitude of the fluctuations is much larger. The value of CV can thus be used as another important criterion for the identification of fire.

(a) The fire scene
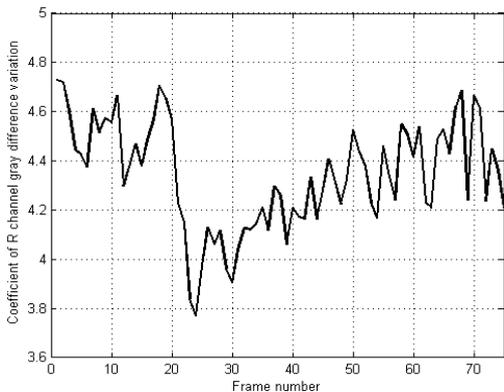
(b) The sunset scene

Figure 13. A comparison between the values of CV obtained on the videos in Figure 3

4. Putting all features together

In view of the fact that using a single feature for flame recognition may lead to high false positive rates and may suffer from other interference sources, we use a k-nearest neighbor (kNN) classifier, a support vector machine (SVM) and a BP neural network (BPNN) to process the five features that we have described in section 3 and identify flames respectively. We denote the five features with A_r, α, f, μ_D and CV, where A_r represents the flame area change rate; α is flame similarity; f denotes flame flicker frequency; μ_D denotes the mean value of gray difference of R channel component and CV is the variation coefficient of R channel gray difference.

As an example, **Figure 14** shows the flowchart of the flame recognition algorithm where a BP neural network is used for classification. The BP neural network consists of three layers. The input of the BP neural network is the five features A_r, α, f, μ_D and CV as described above. The decision of recognition is made based on the value of its output. Specifically, a fire alarm is triggered only when the output value is in the range of [0.5,1]. The network has five nodes in the input layer, 10 nodes in the intermediate layer and one node in the output layer. The standard gradient based algorithm is used to train the BP neural network. The flowcharts of the algorithms based on a kNN classifier and a SVM differs from the one in **Figure 14** only in the step of classification, where a kNN classifier or a SVM is used to determine whether a flame exists in the video sequence. The parameter *k* in the kNN classifier is selected to be 5 and all three classifiers are trained with the same training dataset. Specifically, the data in 250 frames is selected as the training set. 150 frames of the 250 training frames contain flames and the remaining 100 frames contain other interference sources.
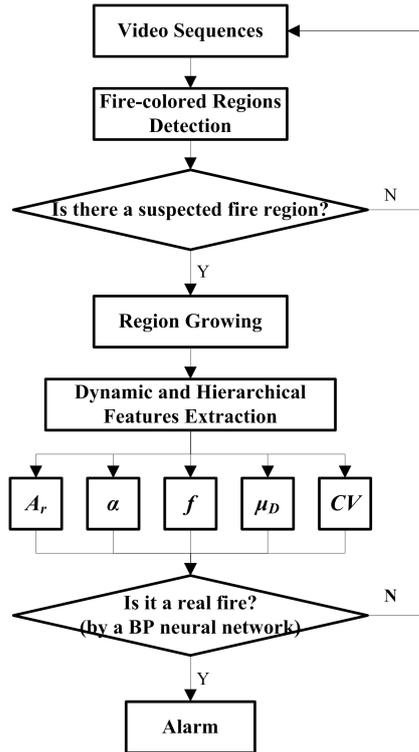
Figure 14. Flowchart of the flame recognition algorithm

5. Testing results and analysis

To test the recognition accuracy of the proposed flame detection algorithm, we have implemented it in Matlab 2011b with three different classifiers and tested the recognition accuracy on ten videos shown in **Figure 15.** And a description of each testing video is shown in **Table 3**. The testing is performed in a computer system that has window 8 operating system installed. The CPU in the system is Intel (R) Core (TM) Duo CPU 1.80 GHz and the main memory of the system is 4 GB.
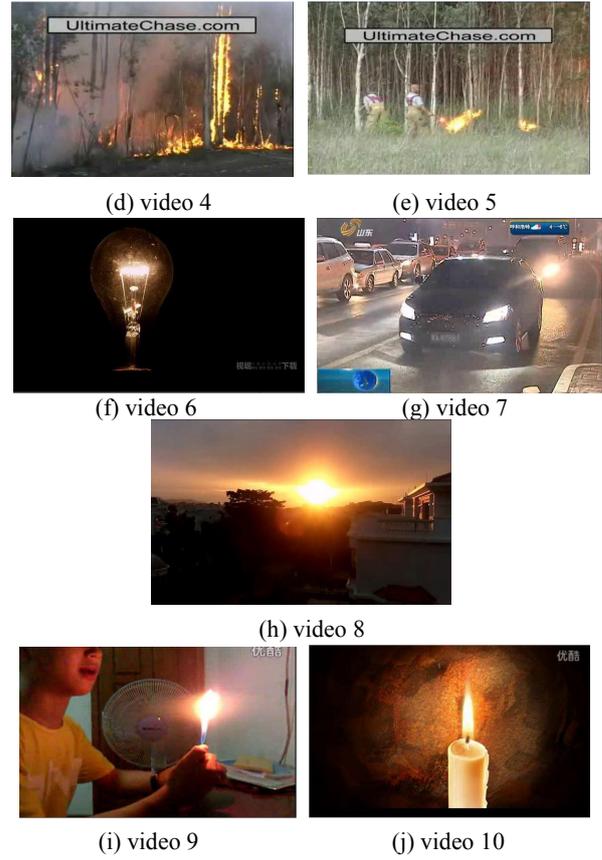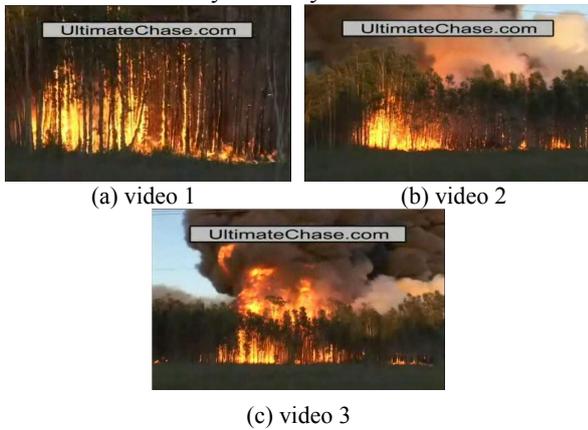

(a) video 1   (b) video 2

(c) video 3

(d) video 4   (e) video 5

(f) video 6   (g) video 7

(h) video 8

(i) video 9   (j) video 10

Figure 15. Screenshots of testing videos

From **Table 3**, we are able to see that videos 1 through 5 are videos of fire flames, interference sources for these videos include smoke, wind and moving objects. Videos 6 through 8 do not contain flames. Videos 9 through 10 contain flames that are under control and in safe conditions, a fire alarm should not be generated for these videos. The recognition results on videos that contain flames are shown in **Table 4**. **Table 5** shows the recognition results on videos that do not contain flames. **Table 6** shows the recognition results on videos of flames that are under control and in safe conditions.

Video Number	A description of the video
video 1	Forest fire, many flame regions and little interference in background
video 2	Forest fires, many flame regions, smoke contains small flame regions
video 3	Forest fires, many flame regions accompanied by smoke
video 4	Forest fires, many flame regions, there is smoke and wind disturbance

video 5	In the early stage of the fire, two men took a torch and walked back and forth in a red sign.
video 6	An incandescent lamp is on in an indoor dark environment
video 7	Night highway video, many cars passing, the ground is illuminated by the lights
video 8	Evening sunset video, the sky has a larger area of color interference
video 9	Indoor lighter flame
video 10	Indoor burning candle

Table 3. Descriptions of the testing videos

We also use the approaches developed in[2,3,22], and[23] to detect flames in all the testing videos and compare their recognition results with those of our approach. The approach in[2] uses a method that combines the RGB-HSI color model and motion detection for fire recognition. The approach in[3] uses a statistical color model of RGB and a similar background subtraction to detect flames. The algorithm in[22] recognizes flames with an approach based on YCrCb color space and K-means clustering. The algorithm developed in[23] uses a combination of radiation domain features to identify flames.

It can be seen from **Table 4** that all tested approaches are able to achieve good recognition results in videos 1 through 4, where the recognition rates of our approach are even higher than those of two other approaches when a BPNN is used as the classifier. However, the performance of our approach with all three classifiers on video 5 is not ideal and the recognition rates achieved on this video are slightly lower than those of the other four approaches. However, a closer analysis of the recognition results on this video shows that most of the frames where our approach fails to recognize flames are located within the first 100 frames. This is possibly due to the fact that the video starts to collect information from the early stages of the fire and the size of the flames is small in the beginning. In addition, the video also contains color interferences from two objects that move back and forth and the occlusion caused by the two moving objects interferes with the flame feature extraction, which results in a high false alarm rate. The testing result on this video shows that the process of image segmentation and feature extraction in our approach needs to be further improved.

Video Number	Total Number of Frames	Number of Frames with Flames	Recognition Rates						
			Approach in [2]	Approach in [3]	Approach in [22]	Approach in [23]	Our Approach (BPNN)	Our Approach (kNN)	Our Approach (SVM)
Video 1	250	250	98.4%	100%	99.6%	100%	100%	97.2%	97.6%
Video 2	250	250	97.2%	97.6%	99.2%	98.8%	100%	97.6%	98.4%
Video 3	250	250	95.2%	95.6%	96.0%	96.4%	97.6%	96.0%	97.2%
Video 4	250	250	97.6%	96.0%	98.0%	98.4%	99.6%	99.2%	98.8%
Video 5	375	375	90.7%	88.5%	86.1%	86.4%	85.9%	86.7%	86.4%

Table 4. Testing results on videos that contain flames

Indoor lighting, moving lights and sunlight are some common bright colors, and the testing results show that our approach can effectively eliminate interferences from these sources. **Table 5** clearly shows that our approach significantly outperforms the other two approaches on false alarm rates in videos 6 through 8. Of course, since video 7 contains slow moving lights and many sources of interference, our approach generates false alarms on 3.2%, 6.5%, and 4.8% of the frames respectively when three different classifiers are used for classification. This also suggests that our approach needs to be further improved on the elimination of interferences from sources other than flames.

Video Number	Total Number of Frames	Number of Frames with Flames	False Alarm Rates						
			Approach in [2]	Approach in [3]	Approach in [22]	Approach in [23]	Our Approach (BPNN)	Our Approach (kNN)	Our Approach (SVM)
Video 6	250	0	6%	0%	2%	1.2%	0%	0.8%	1.2%
Video 7	124	0	75.8%	43.5%	25.8%	19.4%	3.2%	6.5%	4.8%
Video 8	123	0	19.4%	27.4%	16.3%	13.8%	0%	8.1%	4.9%

Table 5. Testing results on videos that do not contain flames

Since fire alarms do not need to be issued for flames in candles, lighters and other types of flames that are under control and in safe conditions, it is a challenging problem to accurately recognize these types of flames from flames that may lead to fire hazards. It is clear from **Table 6** that our approach is able to achieve significantly improved false alarm rates on videos that contain these types of flames. Specifically, the person in video 9 ignites the lighter in his hand and the flame of the lighter is not stable since the hand that holds the lighter is shaking and the flame region is also large. All these may generate adverse effects on the identification process and the false alarm rates thus reach 12.1%, 15.4%, and 13.4% respectively for three different classifiers. On the other hand, since the location of the burning candle in video 10 remains fixed throughout the video, the false alarm rates are only 4.8%, 8.7%, and 7.1% respectively. Table 6 shows that our approach is effective on recognizing flames that are under control and in safe conditions and can achieve significantly reduced false rates in the presence of these types of flames. Testing results in Tables 4-6 also suggest that our approach can be combined with a BP neural network to achieve the best overall recognition accuracy.

Video Number	Total Number of Frames	Number of Frames with Interferences	False Alarm Rates						
			Approach in [2]	Approach in [3]	Approach in [22]	Approach in [23]	Our Approach (BPNN)	Our Approach (kNN)	Our Approach (SVM)
Video 9	149	149	79.9%	80.5%	53.7%	36.9%	12.1%	15.4%	13.4%
Video 10	126	126	77.0%	87.3%	65.1%	31.7%	4.8%	8.7%	7.1%

Table 6. Testing results on videos that contain flames in safe conditions

5. Conclusions

In this paper, we develop a new approach for video based flame recognition, the approach processes video images and recognizes regions that may contain flames with the RGB color model and a clustering based two-step approach. A few new dynamic and hierarchical features of the suspected flame regions are then extracted for further analysis. The final decision on whether fire occurs is made by a BP neural network that processes the features as its input. Our testing results show that, the approach is robust and able to identify the presence of flames under complex circumstances where other interference sources may also exist. In addition, our approach is able to accurately identify flames that are under control and in safe conditions. A comparison of our approach with four other existing flame recognition methods show that our approach can achieve improved recognition rates in most of the testing videos that contain flames and significantly improve on false alarm rates in testing videos that do not contain flames or only contain flames that are in safe conditions.

However, our testing results also show that the performance of the approach is still not satisfactory in cases where occlusions of flames exist or flame regions

are small. The false positive rates are especially high in the initial stage of a fire. New techniques thus need to be developed to further improve the segmentation of the suspected flame regions and the extraction of relevant features. In addition, since more computation time is needed to analyze the dynamics of suspected regions and obtain their layering features, our approach is computationally less efficient than approaches that only use static features of flames. Additional work is thus needed to improve the real-time performance of our approach.

The five features adopted by this approach for classification are based upon previous research results. So far, we have no knowledge on whether these features are indeed the best features that can be utilized for flame recognition or not. However, we believe that the recently developed methods in deep learning can be used to extract optimal or near optimal features for classification and the recognition accuracy can probably be significantly improved. This also constitutes an important part of our future work.

References

1. G. Zang, C. Huang, Y. Wang *et al.*, "Multi-criterion indentification technology of fire with video and its application," Computer Applications and Software, vol. 30, no. 2, pp. 65-67, February 2013.
2. T. Chen, P Wu, and Y Chiou, "An early fire-detection method based on image processing," in Proceedings of International Conference on Image Processing, Singapore, October 2004, pp. 1707-1710.
3. T. Celik *et al.*, "Fire detection using statistical color model in video sequences," Visual Communication & Image Representation, vol. 18, no. 2, pp. 176-185, January 2007.
4. P. Gomes, P. Santana, and J. Barata, "A vision-based approach to fire detection," International Journal of Advanced Robotic Systems, vol. 11, no. 149, pp. 1-12, 2014.
5. T. Celik, H. Demirel, "Fire detection in video sequences using a generic color model," Fire Safety Journal, vol. 44, no. 2, pp. 147-158, 2009.
6. L. Chen, W. Huang, "Fire detection using spatial-temporal analysis," in Proceedings of the world Congress on Engineering, London, U.K., July 2013, pp. 1-4.
7. W. Horng, J. Peng, and C. Chen, "A new image-based real-time flame detection method using color analysis," in Proceedings of IEEE Networking, Sensing and Control, March 2005, pp. 100-105.
8. G. Marbach, M. Loepfe, and T. Brupbacher, "An image processing technique for fire detection in video images," Fire Safety Journal, vol. 41, no. 4, pp. 285-289, 2006.
9. T. Truong, Y. Kim, and J. Kim, "Fire detection in video using genetic-based neural networks," IEEE International Conference on Information Science and Applications, April 2011, pp. 445-449.
10. B. Toreyin, Y. Dedeoglu, and A. Cetin, "Flame detection in video using hidden markov models," in Proceedings of International Conference on Image Processing, Genoa, Italy, September 2005, pp. 1230-1233.
11. T. Celik, "Fast and efficient method for fire detection using image processing," ETRI Journal, vol. 32, no. 6, pp. 881-890, December 2010.
12. Y. Wang, D. Xu Dafang, X. Chen *et al.*, "Fire detection method based on flame dynamic features," Measurement & Control Technology, vol. 26, no. 5, pp. 7-9, 2007.
13. J. Zhao, Z. Zhang, S. Han *et al.*, "SVM based forest fire detection using statistic and dynamic features," Computer Science and Information Systems, vol. 8, no. 3, pp. 821-840, June 2011.
14. S. Rinsurongkawong, M. Ekpanyapong, and M. N. Dailey, "Fire detection for early fire alarm based on optical flow video processing," IEEE International Conference on ECTI-CON, May 2012, pp. 1-4.
15. B. Toreyin, Y. Dedeoglu, and A.E. Cetin, "Computer vision based method for real-time fire and flame detection," Pattern Recognition Letters, vol. 27, no. 1, pp. 49-58, 2006.
16. F. Yuan, G. Liao, Y. Zhang, "Feature extraction for computer vision based fire detection," Journal of University of Science and Technology of China, vol. 36, no. 1, pp. 39-43, January 2006.
17. X. Wu, Y. Yan, J. Du *et al.*, "Fire detection based on fusion of multiple features," CAAI Transactions on Intelligent Systems, vol. 10, no. 2, pp. 240-247, April 2015.
18. Y. Yang, J. Du, S. Gao *et al.*, "Video flame detection based on fusion of multi-feature," Journal of Computer-Aided Design & Computer Graphics, vol. 27, no. 3, pp. 433-440, March 2015.
19. G. Yadav, V. Gupta, V. Gaur *et al.*, "Optimized flame detection using image processing based techniques," Indian Journal of Computer Science and Engineering, vol. 3, no. 2, pp. 202-211, April 2012.
20. M. Kandil, M. Salama, "A new hybrid algorithm for fire-vision recognition," in IEEE Eurocon, May 2009, pp. 1460-1466.
21. L. Wang, A. Li, C. Hao, "A fire detection method using the flame image jumping feature," in Proceedings of the 33rd Chinese Control Conference, Nanjing, China, July 2014, pp. 7421-7425.

22. Z. Wang, P. Liu, and T. Cui, "Research on forest flame recognition algorithm based on image feature," The International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, vol. XLII-2/W7, pp. 925-928, September 2017.

23. Z. Wu, T. Song, W. Wu, X. Shao, and Y. Liu, "Spectral spatio-temporal fire model for video fire detection," International Journal of Pattern Recognition and Artificial Intelligence, vol. 32, no. 5, 1850013, May 2018.