Video Flame Detection Method Based on Improved Fast Robust Feature

Zhang Lingkai\textsuperscript{1,*} and Lu Li\textsuperscript{1}
\textsuperscript{1} Shanghai Dianji University, Shanghai, 201306, China

\textsuperscript{*}E-mail: zhanglingkai1997@163.com

Abstract. Flame detection has important practical significance. Based on the uniqueness of the color information of the flame, using the flame's color model to initially extract the suspected flame area to improve the accuracy of flame detection in complex environments. In order to reduce the amount of unnecessary algorithm calculation, an improved feature extraction method combining color information with acceleration robust combining features (SURF) is proposed. On this basis, the morphological features of flame, such as roundness and rectangularity, are added as auxiliary classification features, and the flame region features extracted from the image are input to support vector machine (SVM) for training and learning. The experimental results show that the flame detection method proposed in this paper is applicable to a wider range of scenes, with the advantages of high accuracy and robustness, high reliability, small calculation amount, short detection time, and still has good robustness in complex scenes with more interferences.

1. Introduction

Fires are common natural and man-made disasters that adversely affect normal production activities, cause huge economic losses, and threaten personal safety. Therefore, the implementation of the flame detection method technology is of great significance. The current fire detection system is designed by fire characteristics such as smoke, temperature or heat radiation\textsuperscript{[1]}, and the detection effect is good only in a small enclosed space or in a specific environment, and it is prone to false alarms and leaks in a complicated environment. The phenomenon of reporting cannot cope with the real-time requirements of fire detection.

Flames have very unique and diverse features, such as color, brightness, and disorderly beating. These features are very intuitive and can be directly perceived and reacted quickly. Computer vision technology can extract reliable information from video or images, resulting in fairly good processing results for the recorded environment.

Based on the rapid development of computer vision technology, many scholars have done research on video flame detection in recent years. Toreyin\textsuperscript{[2]} et al. used the wavelet frequency domain to represent the edge of the flame, and the high-frequency features at the edge of the fire zone were used to detect the characteristics of the flame; however, other researchers used the various features of the flame to detect better. Wu Yuyin\textsuperscript{[3]} improved the color model of Chen et al., established a new flame color model, and combined a variety of flame contours and structural features to obtain a higher detection rate. Tang Yanyan\textsuperscript{[4]} used GMM to detect moving objects in the flame scene, and extracted color features and area variation features to identify the flame in the scene. The method has lower accuracy in complex environments. Zhu Teng\textsuperscript{[5]} combined fire and non-fire conversion with fire
motion information to establish a hidden Markov model, and judged whether the flame was detected by judging the similarity between the test sequence and the two models. In recent years, some scholars have also trained the neural network to identify the flame. For example, Zhen[6] extracted the suspected flame region through the RGB model, trained CNN to classify the normalized candidate fire feature map, and achieved good results. The richness of training data may increase its detection rate. In this paper, the flame candidate region in the scene is extracted by the color information of the flame, and the color information is combined with the SURF feature to extract the SURF feature of the flame region. On this basis, the morphological features of the flame are added to eliminate the regular interference, so that the extracted features are more With flame characteristics, it can better describe the flame. The flame detection method proposed in this paper improves the accuracy of flame detection and has better robust performance in different complex scenes.

2. Flame detection method

2.1. Flame suspected area extraction

The flame is constrained by factors such as the combustion temperature and the type of material of the combustion material, and has a unique color distribution. According to the study by Chen[7], the flame color has a special distribution law in the RGB and HIS color spaces, and the following flames are proposed. The color rule is as in formula

\[
\begin{align*}
\text{rule1}: & R > R_T \\
\text{rule2}: & R \geq G > B \\
\text{rule3}: & S \geq (255 - R) * S_T / R_T
\end{align*}
\]

(1)

In the above formula, R, G, B represent the color components of the flame in each channel of the RGB color space, Rt is the lower threshold of the red component, and St is the saturation threshold of the S component in the HIS color model. Research shows that \( R_t \in [115, 135], S_t \in [45, 60] \). After Chen, relevant scholars improved the color model of the flame, such as adding the color distribution law in the YCbCr color space; proposing new pixel point color constraints; and new color-based methods such as color clustering analysis. Excessively harsh color region screening conditions will perform well in some scenarios, but will ignore a lot of information in complex scenes. This paper uses Chen's proposed flame color model to extract flame candidate regions as shown in Figure 1.

**Figure 1.** Extracting suspected flame regions according to the distribution law of flame color in RGB and HSV color spaces

2.2. Improved SURF feature extraction

SURF (Speeded Up Robust Features) is a robust local feature point detection and description algorithm proposed by Herbert Bay et al. in 2006[8]. It is improved according to the SIFT algorithm proposed by David Lowe et al in 1999. The large lifting makes the algorithm available in real-time computer vision systems. The SURF feature of the flame in the extracted image excludes some interferents that are similar in flame characteristics.
2.2.1 SURF feature extraction. SURF feature is to process the integral image. The value of each pixel in the integrated image is the sum of all the elements in the upper left corner of the corresponding position on the original image. The definition is as shown in equation (2):

$$I_E(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(x, y)$$  \((2)\)

The Gaussian scale space is established. The kernel function used in the SURF algorithm is a Gaussian kernel function, which satisfies the requirements of scale transformation. The Gaussian kernel function \(G(t)\) is as shown in equation (3):

$$G(t) = \frac{\partial^2 g(t)}{\partial x^2}$$  \((3)\)

Where \(g(t)\) is a Gaussian function and \(t\) is a Gaussian variance. In order to achieve the same image at different scales, the convolution of the Gaussian kernel \(G(t)\) with the image function \(I(x)\) at point \(x\) is achieved. Its representation is as shown in equation (4):

$$L(x, t) = G(t) \cdot I(x, t)$$  \((4)\)

The Gaussian kernel obeys a normal distribution, and its coefficients are getting lower and lower from the center point. The Surf algorithm uses a box filter to approximate the Gaussian filter to improve the operation speed. Boxfilter (Boxfilter) through a few simple search for the integral map, you can complete the calculation of pixel addition and subtraction between different areas of the image, that is, the image filtering problem.

The Hessian matrix is a square matrix composed of second-order partial derivatives of multivariate functions. This is the core part of the SURF algorithm. It describes the local curvature of the function. By constructing the Hessian matrix to generate image-stable edge points, feature extraction is performed. basis. Find the Hessian matrix of equation (5) for each pixel:

$$H(f(x, y)) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2$$  \((5)\)

The position of the key point is determined according to the discriminant of the Hessian matrix, which is based on whether the current point is brighter or darker than other points in the surrounding neighborhood. Its discriminant is as shown in equation (6):

$$det(H) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2$$  \((6)\)

The H matrix calculates the second-order partial derivative of the binary function by the convolution between specific cores shown in equation (4), so that the three elements of the H matrix can be obtained, and the H matrix is as in equation (7):

$$H(x, \sigma) = \begin{pmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{yx}(x, \sigma) & L_{yy}(x, \sigma) \end{pmatrix}$$  \((7)\)

Introduce the error between the approximation of the weight balance and the exact value. The scale change will also change the weight, then the H matrix discriminant can be expressed as follows:

$$det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2$$  \((8)\)

Comparing each pixel point processed by the Hessian matrix with 26 points in the scale space and the original image space neighborhood, the key points are initially determined, and the key points with weak energy and the key points of incorrect positioning are removed, and finally Stable feature points. Then, the feature points of the sub-pixel level are obtained by the 3-dimensional linear interpolation method, and those points smaller than a certain threshold value are removed, and the extreme value is
reduced to reduce the number of detected feature points, and finally only the strongest feature points are detected.

In order to obtain the final eigenvector, the Surf algorithm calculates the sum of the horizontal and vertical harr wavelet features of all points in the 60-degree sector by statistically analyzing the harr wavelet features in the circular neighborhood of the feature points, i.e., within the circular neighborhood of the feature points. After the fan shape rotates at a certain interval and the harr wavelet feature value in the region is again counted, the direction of the sector with the largest value is finally taken as the main direction of the feature point, and finally the feature vector is formed.

2.2.2 Improved SURF features. The traditional SURF texture feature is a static two-dimensional texture feature. In order to make the extracted features more consistent with the flame characteristics, an improved fast robust feature extraction method for video flame detection is proposed. The color information of the flame is added to the flame. In the SURF feature detection process, the global detection of the feature points is limited to the flame candidate regions extracted above, which can effectively reduce the interference points of the complex background.

The specific process is as follows:
Step 1: Perform preprocessing on the input image for Gaussian blurring;
Step 2: traversing the selected points using a color model for each pixel;
Step 3: Construct a Hessian matrix on the extracted pixel points, and calculate a feature value;
Step 4: Construct an image pyramid;
Step 5: Locating the feature points, determining the main direction of the feature points, and extracting the SURF features in the screening area;

For each pixel, the color information is combined with the color information to determine whether it is a flame point, and then the SURF feature is calculated, processed sequentially, and input into the SVM classifier for training until the entire image is traversed. Perform feature point detection on the flame in the same scene as shown in Figure 2.

![Figure 2](a). Outdoor flame image extraction SIFT features. (b): Outdoor flame image extraction SURF features. (c): Outdoor flame image extraction improved SURF features.

| Table 1. Computational Quantity of Different Algorithm Feature Points and Comparison of Algorithm Detection Time |
|---------------------------------------------------------------|
| Feature Points | Number of feature points | time consuming(s) |
|----------------|----------------------------|-------------------|
| SIFT           | 5943                       | 16.63             |
| SURF           | 1352                       | 2.439             |
It can be seen from Figure 2 and Table 1 that the flame detection method proposed in this paper adds SURF features only to the pixel points of the flame candidate area after adding the color information, which is more representative of the flame characteristics, and reduces the interference of the background features during training. The calculation amount is reduced to 70%-75% of the original algorithm, which is less and more real-time.

2.2.3 Adding auxiliary features. Based on the color information of the flame and the SURF characteristics of the image, other information of the flame cannot be ignored, and the morphological characteristics of the flame also have a high degree of recognition. Therefore, the morphological feature of the flame is added as an auxiliary classification feature. The SURF features are added and merged together as a discriminating basis for the classifier. The flame is used as a non-rigid target, and its shape is not highly regular and can be used to exclude some regular interferences such as lights[9].

In order not to increase the calculation amount of the algorithm, the circularity and rectangularity are selected as the auxiliary features. The circularity is used to measure the complexity of the edge contour of the object and represents the similarity between the edge of the object and the circle. Calculate the area of the object area and the length of the edge of the object to obtain the circularity value C:

$$C = \frac{L^2}{4\pi S}$$ (9)

Where: S is the area of the object area and L is the length of the edge of the object. The circularity value represents the complexity of the edge contour of the object. The more the edge contour of the object is meandered, the larger the value is, the minimum value is 1. Objects such as lights with similar flame colors have lower edge complexity, a circularity value of approximately 1, and an edge contour of the flame is irregular, and the circularity value is usually much larger than 1, so Degrees can eliminate interference with low edge complexity.

The degree of rectangle measures the degree of similarity between the edge contour of the object and the rectangle. The rectangular value R is defined as equation (10):

$$R = \frac{S}{SR}$$ (10)

Where SR is the area containing the smallest rectangle of the area where the object is located. The degree of squareness indicates how much the object area fills its smallest circumscribed rectangle. The value of the rectangular value is $R \in [0, 1]$, and the squareness of the flame region is usually about 0.5. By the degree of squareness, it is possible to distinguish between objects resembling a rectangle (rectangle is close to 1) and objects with a thin shape (rectangle is close to 0).

2.2.4 SVM classifier. Support. Vector Machines (SVM) is a machine learning method developed according to statistical learning theory. It was first proposed by Cortes and Vapnik et al.[10] in 1995. The basic idea is to pass the input space through the inner product function. Linear transformation into a high-dimensional feature space, and then obtain the linear optimal classification surface, that is, construct a linear function in high-dimensional space to achieve the nonlinear discriminant function in the original space.

Constructing the optimal linear function of the two types of problems correctly separates the two categories in the sample set (training error rate is 0), making the classification interval the largest. Normalize the classification line equation to make training samples $(x_i, y_i), i = 1, 2, \cdots, n, x \in \mathbb{R}^d, y \in \{+1, -1\}$ (where $x_i$ is the n-dimensional training sample, which is the category information of the sample) that satisfies the formula (11):

$$y_i[(\omega \cdot x_i) + b] - 1 \geq 0, i = 1, 2, \cdots, n$$ (11)
The classification interval is $2 \frac{\| \omega \|^2}{\| \omega \|^2}$, so that the interval is maximized, that is, $\| \omega \|^2$ is minimized. The classification surface that satisfies the above formula and minimizes $2 \frac{\| \omega \|^2}{\| \omega \|^2}$ is the optimal classification surface. Video flame detection is obviously a nonlinear classification problem. The nonlinear transformation is used to transform the problem into a linear problem in high-dimensional space, and the optimal classification surface is obtained in the transformation space. In the optimal classification plane, the inner product can be used to complete the linear classification after nonlinear transformation, but the computational complexity does not increase. The common inner product functions have polynomial kernel function, Gaussian radial basis kernel function and Sigmoid kernel function. The three kernel functions are defined as follows:

$$k(x \cdot x_i) = [\delta \cdot (x, x_i) + 1]^q$$

(12)

$$k(x \cdot x_i) = \exp\left(-\frac{\|x-x_i\|^2}{\delta^2}\right)$$

(13)

$$k(x \cdot x_i) = \tanh(\delta \cdot (x, x_i) + c)$$

(14)

Where $x_i$ is the eigenvector in the input classifier, and $\delta$ and $c$ are the kernel function parameters.

The experimental results are as shown in Table 2. The Gaussian radial basis kernel function is the best, and the classification accuracy of the research content is higher.

| Kernel function                          | Number of test samples | Recognition rate(%) |
|------------------------------------------|------------------------|---------------------|
| Polynomial                               | Polynomial             | 88.9                |
| Gaussian radial basis                    | Gaussian radial basis  | 93.5                |
| Sigmoid                                  | Sigmoid                | 81.9                |

3. Experimental results and analysis

3.1. Experimental

The experiments described in this paper were carried out under the environment of Intel(R) Core(TM)i5-5200U//CPU@2.20GHz4.00GB memory Windows10 system Matlab 2018a. Most of the selected video data and image data came from the flame of Bilkent University. The video library (http://signal.ee.bilkent.edu.tr/VisiFire/index.html), the rest comes from the Internet. The data content is described in Table 1. The text of your paper should be formatted as follows:

| Serial number | Video data description                                      |
|---------------|------------------------------------------------------------|
| 1             | Empty scene, with large areas of yellow grass and blue sky, with fire spots in the center of the scene. |
| 2             | In the wild scene, two workers are walking around, the background is the forest, and the fire is getting bigger. |
| 3             | Highway scene, bright road, there is a car passing by, there is a fire point in the lower right corner |
| 4             | In the wild scene, the background is the forest, the fire point is concentrated, and the area is larger. |
Wild scenes, forest fires, large flame areas, and a large number of thick smoke layers.

Tunnel scene, with strong light, the vehicle quickly passed.

Highway scene, vehicle driving, car lights illuminate the ground, similar to flame.

### 3.2. Result analysis

Common video detection performance indicators [11], TP /% flame frame accuracy, FP% for flame frame miss detection rate; TN /% non-flame frame accuracy, FN /% for non-flame frame false detection rate.

| Serial number | Total number of frames | Number of flame frames | Literature [10] | Literature [7] | This article |
|---------------|------------------------|------------------------|-----------------|----------------|--------------|
|               |                        |                        | TP/%            | FP/%            | TP/%         | FP/%         | TP/%         | FP/%         |
| 1             | 140                    | 140                    | 75.7            | 24.3            | 78.6         | 21.4         | 84.3         | 15.7         |
| 2             | 260                    | 260                    | 86.2            | 13.8            | 81.2         | 18.8         | 93.1         | 6.9          |
| 3             | 1201                   | 1201                   | 93.0            | 7.0             | 94.4         | 5.6          | 97.3         | 2.7          |
| 4             | 246                    | 246                    | 93.9            | 6.1             | 94.3         | 5.7          | 100          | 0            |
| 5             | 245                    | 245                    | 95.5            | 4.5             | 91.8         | 8.2          | 100          | 0            |
| 6             | 151                    | 0                      | 68.9            | 31.1            | 55.0         | 45.0         | 98.0         | 2.0          |
| 7             | 146                    | 0                      | 61.7            | 38.3            | 78.1         | 21.9         | 94.5         | 5.5          |

From the comparison of the results of Table 3 and Table 4, the literature [10] uses the HSI color model to extract the flame candidate region, selects five flame features such as flame sharp angle and flame image shape similarity, and inputs into the SVM to train and establish the classification model. The method utilizes the dynamic and static characteristics of the flame, and the detection effect is better. Based on the RGB color model, the literature [7] detects whether the test sample is a flame by iterating the dynamic growth rate of the measured video flame, and the process mainly depends on the color information. There are large misjudgments for interferences with similar colors such as headlights and hay. From the detection effect, the performance of the video 1 comparison algorithm is not ideal, because the flame area in the video data is very small, only a few points can be detected in some frames, and because there are large pieces in the background. In the yellow grassland, the literature [7] showed a high false detection rate; in video 2, the fire gradually became larger, and the literature [7] performed better than the literature [10] because the dynamic growth of the flame was adopted in the literature [7]. The rate is used as a basis for classification, so there is even better performance in such scenarios. Video 3-5, the flame area is more obvious, both the literature [10] and the literature [7] have shown good results, in the scene of large flame area and no significant change in the combustion area, the method of the literature [10] performance is slightly Better than the literature [7].

For non-flame video data, the algorithm of [7] relies on a single flame feature, which cannot be effectively excluded in the face of complex objects such as lights, hay, bright roads, etc. in complex situations. In the case of the method [10], the detection characteristics of the method in [10] are not obvious when the flame is relatively stable, and the bright area illuminated by the lamp has the same characteristics as the flame, which is likely to cause false detection. The algorithm proposed in this paper, based on the color information of the flame, determines the suspected area of the flame, combines the color information with the SURF feature, preliminarily excludes areas that do not conform to the color information, and retains the flame candidate area, in
the initially selected flame suspected area. Limiting the detection of SURF feature descriptors improves the efficiency and effectiveness of feature extraction, and reduces the computational complexity of feature detection. Therefore, the real-time performance of the detection method is improved, and the texture features extracted by color information are combined to eliminate a large number of complex backgrounds. Characteristic interference can better detect the flame in a variety of scenes and is not affected by regular interferences. For video 1, the flame effect detected by the applied color information is not significant due to the small flame area. For video 2-7, the proposed algorithm exhibits a high detection rate and a low accuracy. The detection rate can reach 100% under the condition of stable fire or sufficient combustion. Under the influence of similar interferences, the algorithm in this paper also has a fairly high robustness, which can eliminate most of the interference such as car lights and hay.

4. Conclusion
In this paper, based on the color characteristics of the flame, the suspected area of the flame is extracted. The SURF feature of the flame image is extracted by the color information combined with the SURF feature of the flame region. On the basis of this, the auxiliary classification feature of the flame is added to the circularity and the rectangle is used to eliminate the interference. To make the features more flame-like, the SVM classification method has great advantages in solving small sample and nonlinear problems, and a higher detection rate is obtained by training small sample data. The video flame detection method proposed in this paper improves the recognition accuracy of the flame, the reliability is high, the calculation amount is small, and the detection time is short. Under the complicated scene with more interferences, the robustness is still good. In some cases where the flame area is small or in the early stage of the fire, such as video 1, the detection rate of the method is low. This is because the color information is not prominent in the early stage of the flame, and it is impossible to effectively detect a sufficiently large candidate area. In the improvement work, it is necessary to add and combine the effective features such as motion features in the early stage of detecting the flame, so that the application range of the algorithm is wider.

Conference
[1] Baumberg and D. Hogg. An adaptive eigenshape model. In Proceedings of the British Machine Vision Conference, volume 1, pages 87–96, 1995
[2] Toreyin, B.U., Dedeoglu, Y., Gudukbay, U., Cetin, E.: Computer vision-based method for realtime fire and flame detection. Pattern Recognition Letters 27(4), 49–58 (2006)
[3] Wu Yuyin, Yan Yunyang, Du Jing, et al. Multi-feature fusion flame detection algorithm [J]. Journal of Intelligent Systems, 2015, 10 (2) : 240-247
[4] Tang Yanyan, Yan Yunyang, Liu Yi'an. Fast Flame Detection Using GMM[J]. Computer Science, 2012, 39(11): 283-285+297
[5] Zhu Teng (1), Jeong-Hyun Kim (1), Dong-Joong Kang (1). Fire detection based on hidden Markov models [J]. International Journal of Control, Automation and Systems, 2010, Vol.8 (No. 4): 822-830
[6] Zhong, Zhen; Wang, Minjuan; Shi, etc. A convolutional neural network-based flame detection method in video sequence[J]. SIGNAL IMAGE AND VIDEO PROCESSING, 2018, Vol.12 (No. 8): 1619- 1627
[7] CHEN TH,WU PH,CHIOU Y C. An early fire-detection method based on image processing[C]//Proceedings of the 2004 International Conference on Image Processing. Singapore City,Singapore:IEEE, 2004, 3: 1707 -1710
[8] H Bay, T Tuytelaars, LV Gool. SURF: speeded up robust features [J]. European Conference on Computer Vision, 2006, 110(3): 404-417
[9] Zeng Sitong, Wu Haibin, Shen Peihui.Research on Video Flame Detection Method Based on Multi-feature Fusion[J].Journal of Graphics,2017,38(04):549-557
[10] Zhang Zhengrong, Li Guogang. Fire Detection Technology Based on Support Vector Machine[J]. Microcomputer & Applications, 2010, 29(24): 70-72+76

[11] Yan Yunyang, Zhang Huizhen, Liu Yi'an, Gao Shangbing. Video Flame Detection Based on GMM and 3D LBP Texture[J]. Journal of Shandong University (Engineering Science), 2019, 49(01): 1-9