Research on flame recognition technology based on local complex features

Di Zhu 1, Kamarul Arifin Ahmad1* and Aolin Chen 1

1 Faculty of Engineering, University Putra Malaysia, 43400 UPM Serdang Selangor Darul Ehsan Malaysia

*Corresponding author’s e-mail: aekamarul@upm.edu.my

Abstract. Traditional flame recognition methods based on image features are difficult to extract flame image features effectively, resulting in low flame recognition accuracy, while such methods mostly perform specific flame recognition for specific scenes, and when the scene, flame color and other features change, it is difficult to perform flame recognition effectively. To address this problem, this paper proposes a flame recognition scheme based on local complex features. Its main purpose is to fuse multi-scene flame data, introduce the characteristics of flame in color space through the process of extracting feature descriptors in SIFT, so as to filter the extracted feature descriptors with noise interferers, and transform the extracted feature descriptors into feature vectors by using the key point bag-of-words method, and finally a general fast flame recognition model based on the limit learning machine. In this paper, we explore the upper limit of the capability of traditional image feature representation to pave the way for the application of deep learning to the flame recognition problem.

1. Introduction

Fire warning is of great significance to people's lives and industrial security. Fire has the characteristics of suddenness and rapid spread of flame burning, while fire poses a serious threat to people's life and property. Therefore, how to identify flames, smoke, and early warning and human intervention, is particularly important [1]. A large number of studies have shown that fires are often accompanied by a large amount of sulfur dioxide, carbon monoxide and nitrogen oxides and other diverse harmful gases, thus causing great damage to the environment and air and affecting the ecological balance [2]. Traditional fire identification methods based on physical sensors, most sampling smoke sensors, temperature sensors, ultraviolet sensors and other real-time monitoring of the physical signal state of the area, so as to achieve the effect of flame, smoke detection [3]. However, the dependence of the environment of this type of method is strong, it is difficult to effectively identify the flame, and at the same time requires a lot of human and material resources for deployment and maintenance, which poses a serious challenge to carry out fire prevention and control work [3]. On this basis, the flame recognition method based on image features proposed in this paper mainly extracts flame image features and constructs classifiers for flame recognition, which to a certain extent alleviates the dependence on the external environment and thus improves the flame recognition accuracy [4].
2. Local filtering SIFT feature extraction

2.1. Flame color spatial characteristics
Celik et al. proposed that in the R, G, and B channels of the color image, for the flame region pixels, the corresponding pixel points should satisfy that the red channel pixel value is greater than the mean value of the red channel component of the whole image, the red channel pixel value of the corresponding pixel point is greater than the blue channel pixel value, and the blue channel pixel value is greater than the green channel pixel value. As shown below:

\[
R(x,y), G(x,y), \text{ and } B(x,y) \text{ in equation 1-3 represent the pixel values of } (x, y) \text{ pixel points in the red, green, and blue channels, respectively, and } R_{\text{mean}} \text{ represents the average of all pixel values of the red channel component in the whole image. Celik et al. proposed that the ratio of the three color channel components R, G, and B can be used as another criterion to judge the flame pixels.}
\]

Equation 4 represents the component pixel ratios by pixels in R, G, and B channels, respectively, while setting thresholds to achieve detection of flame regions.

2.2. SIFT feature extraction
The scale-invariant feature transformation algorithm is proposed by Lowel, and the main process is as follows, firstly, SIFT is performed on each image frame [4]. The process is as follows: Firstly, the image pyramids at different scales are established and the Gaussian difference is performed, and the formula is as follows:

\[
* \text{ in Equation 5 represents the convolution operation, } G(x,y,) \text{ represents the two-dimensional normal distribution, and } D(x,y,) \text{ is generated by convolving the difference of pixel } (x,y) \text{ at different scales with the input image [5]. The extreme value point of the Gaussian difference is found to be the key point of the image.}
\]

The key points are further screened, and the key points with low contrast are removed by Eqs. 6-8. If \(>0.03\), the feature point will be retained, otherwise discarded, and the unstable edge response points in the key points are removed by Eqs. 9-10. When the key points satisfy Eq. 10, the feature points are retained, otherwise discarded [6].

The orientation parameters of the keypoints are calculated using the gradient orientation distribution properties of the keypoint neighborhood pixels, as shown in Equation 11. The final feature point descriptor is formed.

2.3. Introduction of flame color space properties and feature encoding
The characteristics of flame in color space are introduced to calculate the SIFT extracted feature points in RGB color space components, and the specific filtering conditions are shown in Equation 1-4, so as to filter out the noise.

Csurka et al. proposed the bag-of-keys (BOK) method, a vector quantization technique with invariant properties for image affine transformations and high robustness to images with cluttered backgrounds. The main processes include: (1) obtaining the extracted image feature descriptors; (2) constructing a feature point-based dictionary, which is a clustering process using k-means, with k representing the number of clustering centers obtained after the final, and each clustering center is called a vocabulary, which together form the dictionary of BOK; (3) computing all the feature descriptors in each sample's row vectors of each sample and the Euclidean distance of each vocabulary in the dictionary, so as to find the one with the smallest distance among them and update the feature vector. BOK achieves to map each feature point in the image to a certain lexicon in visual space, and finally converts the feature descriptors of the image into a feature vector with the same dimension, so as to put them into the model for training, thus better solving the problem that SIFT feature descriptors cannot be modeled directly [6]. Bag of keypoints encoding feature vectors have been widely used on various feature processing problems, Farquhar et al. incorporated generalized generative models on the basis of Bag of keypoints for corresponding improvements, and Koike et al. used Bag of keypoints method for biometric feature information encoding.
In this paper, after getting the filtered feature descriptors and then Bag Of Keypoints processing, the feature descriptors of the feature points are clustered to get the feature point dictionary based on SIFT and flame color space characteristics, and the final feature vector is formed by comparing the distance between the feature descriptors of each sample and each word in the dictionary [7].

3. ELM model construction

First proposed in 2004, the Extreme Learning Machine is a single hidden layer feedforward neural network characterized by randomly generated input layer parameters. This property of the extreme learning machine responds to a certain biological learning mechanism [8]. The Moore-Penrose generalization is utilized so as to obtain the output layer weights with the minimum L2 parametric number, and only the number of hidden neurons needs to be adjusted throughout the learning process. In addition, the limit learning machine has a simple structure and has the ability of extreme approximation to the objective function, so it has a very fast learning speed and excellent generalization ability [9]. The model of the extreme learning machine is shown in Figure 1.

![Figure 1. Extreme learning machine model diagram.](image)

The parameters between the hidden layer and the output layer represented by $B$ in Equation 12, $H$ represents the output of the hidden layer, and $T$ represents the label of the sample. The resulting weights of the output layer are $W_h$, where $H^+$ is the Moore-Penrose generalized inverse of $H$.

From the perspective of the experimental process, the method can be further divided into three parts: data pre-processing, feature extraction and model training [10]. Data preprocessing mainly converts the video in the image database into frame images, feature extraction mainly introduces the flame color space characteristics into SIFT feature extraction, and uses the key point bag-of-words method to process to get the feature vector. Model training is mainly to construct the extreme learning machine classification model.

4. Simulation Experiment

4.1. Experimental setup and flame recognition data set

Real fire image data were used for the experiments, which were conducted on the Matlab2014 platform with an experimental environment of: Intel Core i7, CPU 3.60GHz, and RAM 8.00GB. All flame images in the data used for the experiments, from both publicly available data and network acquisitions, were divided into a training set and a test set. A general overview of the dataset and the data composition of the training and test sets are given in Table 1. The flame images in the dataset contain different scenes and different types of flames, and both the training set and test set images are randomly selected from the dataset.
Table 1. Flame data description.

| Name                | Quantity | Remarks                                                                 |
|---------------------|----------|-------------------------------------------------------------------------|
| Total images        | 15301    | 13968 from public dataset + 1333 from Network                          |
| Training collection | 11000    | 7000 images with flames and 4000 images without flames                 |
| Test Collection     | 4301     | 2074 images with flames and 2227 images Images without flames          |

4.2. Experimental feature extraction

The image data are processed according to the above method. The feature representation of some flames after SIFT processing is described below.

It can be concluded that the SIFT feature can identify the position of the flame better for the ordinary image with a single background, but it cannot identify the position of the flame more accurately for the scene with a smaller flame area and a more complex background. To further filter the noise points, the characteristics of flame in color space are introduced into the above SIFT feature descriptors to get more accurate flame area [11]. It can be clearly seen that the effect of introducing the SIFT feature identifier filtered by color space information is better than the effect of using SIFT feature extraction alone.

After extracting the feature descriptors using the SIFT method based on color space information filtering, the Bag of Keypoints method is further used to transform the unequal number of feature point descriptors of each image into feature vectors of consistent dimensionality. It can be concluded that the non-flame feature vectors have higher peaks at around columns 120 and 330 of the vocabulary and lower peaks in the rest of the columns, and the maximum peak of the feature vectors does not exceed 5. The histogram of the flame features obtained has a wide distribution of peaks, and the maximum peak of each feature vector exceeds 5. This indicates that the histogram of the flame feature vectors extracted by the method differs more significantly from the non-flame feature vectors, thus ensuring that the features are more distinguishable, thus ensuring that the features have a large degree of differentiation.

4.3. Experimental results and analysis

To verify the effectiveness of the method proposed in this paper, three methods are introduced for comparison. Among them, method I uses texture features, information entropy, flame spread rate extraction features, and support vector machine (SVM) as recognition model; method II uses the SIFT feature extraction method based on color space feature filtering proposed in this paper and support vector machine (SVM) as recognition model; method III uses SIFT to extract feature descriptors and BOK feature vector transformation, and uses extreme learning machine as the recognition model. The SVM is implemented using Libsvm toolbox with RBF radial basis kernel function, the kernel parameter is set to 10 and the regularization parameter is set to 100; the ELM hidden layer node is set to 4000 and the regular term coefficient C is set to 1e-3. In order to eliminate the influence of random initial values of ELM, the results of this paper are the average of 100 experiments.

Table 2. Table of experimental results.

| Method Description                        | Training time (s) | Test time (s) | Training accuracy (%) | Test precision (%) | Recall rate(%) | Correctness (%) |
|-------------------------------------------|-------------------|---------------|-----------------------|--------------------|---------------|-----------------|
| Information entropy + texture features + flame spread rate +SVM SIFT+Flame color characteristic + BOK+SVM | 3.25              | 0.63          | 83.50                 | 64.70              | 38.44         | 69.52           |
| SIFT+BOK+ELM                              | 44.15             | 9.42          | 98.15                 | 95.28              | 96.83         | 96.88           |
| SIFT+BOK+ELM                              | 35.20             | 1.38          | 98.72                 | 96.44              | 97.76         | 94.50           |
| SIFT+flame color characteristic + BOK+ELM | 31.73             | 2.27          | 99.30                 | 97.31              | 97.43         | 96.52           |
The comparison results of the four methods on 11,000 training samples and 4301 test samples are given in Table 2. To avoid the volatility of the experimental results, the experimental results in the table are the average values obtained from 100 experiments. It can be seen that the methods in this paper, as well as methods two and three, are significantly better than algorithm one in terms of testing accuracy, which fully illustrates the obvious effect of using SIFT feature extraction in flame recognition, while features such as texture do not have good generality on different scenes and flame types; the training and testing time of method three is significantly better than method two, and the testing accuracy is slightly better than method two, which indicates that using the extreme learning machine constructing the model greatly improves the efficiency of model training and testing while ensuring accuracy [12]. In comparison with the other methods, the method of this paper is superior in training and testing time, testing accuracy and recall accuracy, which fully demonstrates that the proposed method of this paper can deal with the generalized fast flame recognition problem more accurately and effectively.

By analyzing the training and testing time of the extreme learning machine with different numbers of hidden layer nodes, it can be seen that the model testing time changes less in the process of increasing the number of hidden layer nodes from 100 to 5000, which is controlled within 3 seconds, which indicates that the method proposed in this paper has greater industrial application value and can perform flame determination quickly.

By analyzing the ROC curves of the four methods, it is obvious that the area of the ROC curves of the method in this paper and method two and method three is much larger than that of method one, and the average performance of the model. It is obviously better than method one in terms of the average performance of the model. This is consistent with the data comparison results in Table 2. Also comparing the ROC curves of method two and method three, it shows that the effect of the feature extraction method proposed in this paper is better than the direct use of SIFT feature extraction. Also, comparing the ROC curves of the four methods shows both the advantages of the feature extraction method described in this paper and the advantages of the extreme learning machine in model construction.

5. Conclusion
The core problem of flame image recognition lies in effectively capturing flame regions and features, and building fast recognition models. In this paper, by innovatively adding flame color space features to SIFT feature extraction, we effectively filter noise points and convert the obtained feature descriptors into feature vectors by using the key point bag-of-words method, and then build a fast general flame recognition model using an extreme learning mechanism. The experimental results show that the proposed method has a large improvement in feature representation, accuracy rate, and regression rate, compared with the traditional image feature flame recognition method. Meanwhile, the experiments in this paper explore the upper limit of traditional image feature representation in flame recognition by improving the feature representation, so as to pave the way for deep learning to be applied to flame recognition problems and deep learning feature representation for comparison.

References
[1] Yu Lichun, Liu Jinqing. Flame image recognition algorithm based on improved Mask R-CNN [J]. Computer Engineering and Applications, 2020, 56(21): 194-198.
[2] Zhu YH, Xia L, Wang JX. Improved BP neural network based ceramic shuttle kiln flame image recognition method [J]. China Ceramic Industry, 2020, 27(01): 16-20.
[3] Li Chuan-chen, Bai Yong, Chen Yimin. Research on flame image recognition technology based on migration learning [J]. High Technology Communication, 2019, 29(03): 274-282.
[4] Zhu YH, Yang J. Flame image recognition based on deep learning for ceramic shuttle kilns [J]. Western Leather, 2018, 40(22): 152-153.
[5] Liu G. Design of flame detection system for pulverized coal boiler based on embedded image recognition [J]. Modern manufacturing technology and equipment, 2018(06): 28-29.
[6] Mao W.T., Wang W.P., Jiang M.X., Ouyang J. A fast flame image recognition method based on local feature filtering [J]. Computer Applications, 2016, 36(10): 2907-2911.
[7] Yu Lu,Bu Le Ping. A review of common algorithms in flame image recognition[J]. Information Technology,2014(03):189-193.

[8] QIU Guoqing,JIANG Tianyue,FENG Hanqing,TU Lefe. Fire image recognition algorithm based on flame tip angle feature[J]. Computer Application and Software,2013,30(12):52-55+143.

[9] Feng Shicheng,Kong Dongdong,Lu Shuen. Research on fire flame criterion based on image recognition[J]. Automation Technology and Applications,2011,30(05):67-69+77.

[10] Wang Q. The use of flame image features in fire digital image recognition[J]. Modern Machinery,2008(05):60-61+64.

[11] Sun Peng,Chai Tianyou,Zhou Xiaojie,Yue Heng. Flame image recognition system for alumina rotary kiln firing zone[J]. Journal of Chemical Engineering,2008(07):1839-1842.

[12] Han Pu,Zhang Xin,Wang Bing,Pan Weihua. Neural network-based interactive furnace flame image recognition[J]. Chinese Journal of Electrical Engineering,2008(20):22-26.