Dual-Encoded Features from Both Spatial and Curvelet Domains for Image Smoke Recognition

Feiniu Yuan[1, 2], Tiantian Tang[3], Xue Xia[2*], Jinting Shi[4*], and Shuying Li[5]
[1] College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, China.
[e-mail: yfn@ustc.edu.cn]
[2] School of Information Technology, Jiangxi University of Finance and Economics, China.
[e-mail: yeziandkuma@qq.com]
[3] School of Communications and Electronics, Jiangxi Science and Technology Normal University, China.
[4] Vocational School of Teachers and Technology, Jiangxi Agricultural University, China.
[e-mail: icanflysjt@126.com]
[5] School of Automation, Xi'an University of Posts & Telecommunications, China.
*Corresponding author : Xue Xia, Jinting Shi

Received March 5, 2018; revised October 22, 2018; accepted November 7, 2018; published April 30, 2019

Abstract

Visual smoke recognition is a challenging task due to large variations in shape, texture and color of smoke. To improve performance, we propose a novel smoke recognition method by combining dual-encoded features that are extracted from both spatial and Curvelet domains. A Curvelet transform is used to filter an image to generate fifty sub-images of Curvelet coefficients. Then we extract Local Binary Pattern (LBP) maps from these coefficient maps and aggregate histograms of these LBP maps to produce a histogram map. Afterwards, we encode the histogram map again to generate Dual-encoded Local Binary Patterns (Dual-LBP). Histograms of Dual-LBPs from Curvelet domain and Completed Local Binary Patterns (CLBP) from spatial domain are concatenated to form the feature for smoke recognition. Finally, we adopt Gaussian Kernel Optimization (GKO) algorithm to search the optimal kernel parameters of Support Vector Machine (SVM) for further improvement of classification accuracy. Experimental results demonstrate that our method can extract effective and reasonable features of smoke images, and achieve good classification accuracy.

Keywords: Curvelet Transform, Dual-encoded Local Binary Pattern (Dual-LBP), Completed Local Binary Pattern (CLBP), Gaussian Kernel Optimization (GKO), Smoke Recognition

This research was supported by National Natural Science Foundation of China (Grant No. 61862029, Grant No. 61562031), Science Technology Application Project of Jiangxi Province (No. GJJ170317).

1. Introduction

Generally, fire causes significant economic losses and probably lead to severe death. In order to avoid fire occurrence, many traditional fire detection technologies have been widely used. These methods are usually based on temperature sensors, humidity sensors, and traditional ultraviolet and infrared fire detectors. Since traditional methods need to sample combustion products for analysis, they are required to be placed in the vicinity of fire. In addition, traditional detectors are susceptible to external environment influences, such as airflow, dust. Traditional methods cannot provide us with detailed information about burning situation. Therefore, traditional smoke detectors are unreliable in open, large and special spaces.

In most cases, fire will be initially accompanied by the emergence of smoke, and smoke often lasts for a few minutes before flames emerge. According to this observation, visual smoke detection methods detect smoke from videos or images, and they are able to give early alarms of fire.

Early smoke has special visual features, such as color, texture, and shape, which play an important role in fire detection. There are many texture feature extraction methods that have been proposed. Gray-level co-occurrence matrices [1] is a way to describe texture by exploring spatial correlation between gray values of neighboring pixels. LBP [2] provides a binary-coding feature extraction manner by encoding the relationship between central pixels and their neighboring pixels. HOG [3] extracts features of edges and gradients.

Many methods can achieve excellent performance by capturing multi-scale and multi-direction information in transform or frequency domains. Compared with other transforms, Curvelet transform is strongly anisotropic and its needle-shaped elements provide a high directional sensitivity to represent curved singularities in images. In contrast, wavelet transform shows a good representation only at point singularities because it has a poor directional sensitivity. Additional directional-based transforms, such as Dual-Tree Complex Wavelet Transform (DTCWT) and Gabor Wavelets, provide more multi-direction information than Wavelets, but they still have limited directional selectivity. Ridgelet is suitable for representing line singularities in objects, so it's rarely found in practical applications [4].

To extract discriminative features, we propose a novel feature extraction based on spatial and Curvelet domains. The main contributions of this paper are listed as follows:

1) We use Curvelet transform to extract discriminative features from original images, and then encode these images consisting of discriminative Curvelet coefficients to generate LBP codes based on Curvelet domains.

2) We first aggregate histograms of LBP maps from Curvelet domains to produce a histogram image of size 256×50, and then encode the histogram image again to generate novel codes, which are called Dual-encoded Local Binary Patterns (Dual-LBP).

3) We concatenate histograms of Dual-LBPs from Curvelet domain and Completed Local Binary Patterns (CLBP) from spatial domain to generate dual-encoded features for smoke classification. Finally, we adopt Gaussian Kernel Optimization (GKO) algorithm to search the optimal kernel parameters of Support Vector Machine (SVM) for further improvement of classification accuracy.

2. Related Work

There are many methods proposed for smoke detection. Chenebert et al. [5] presented a flame pixel detection method in video images or still images using a non-temporal texture driven approach. The method did not use any time information. Chen et al. [6] used a color model based on RGB for fire smoke detection. However, there are many objects having the same color distribution as fire, so the method gives a false alarm inevitably for these fire-like object. Celik et al. [7] proposed a universal color model for fire pixel detection, and the algorithm used the YCbCr color space to separate chrominance and luminance components more effectively than other color spaces (such as RGB). Yuan et al. [8] proposed an accumulative motion model based on integral image techniques. The model estimated movement directions of objects in real-time for analysis of smoke. Zhang et al. [9] proposed a real-time forest fire detection algorithm using artificial neural networks based on dynamic characteristics of fire regions segmented from video images. Yu et al. [10]presented a method by using color and motion features for video smoke detection. The method could distinguish smoke from objects with similar color distribution by involving motion features and color information, which greatly improved the reliability of video smoke detection. Toreyin et al. [11] achieved smoke detection based on edge magnitude differences, in which the characteristics of smoke such as movement, flashing, edge blur and color were used. Once the scene lacks obvious edges or cluttered objects, the method raises false alarms.

Texture feature features play a key role in smoke detection, Ojala et al. [2] firstly proposed Local Binary Pattern (LBP) for texture classification. It is an efficient and simple gray-scale texture descriptor, which captures spatial characteristics of texture. LBP features have demonstrated very powerful discriminative capability, low computational complexity, and low sensitivity to illumination variations.

To further improve the discriminative capability of LBP, many variants of LBP were proposed in the past decade. Yuan et al. [12] proposed an effective smoke detection method, in which features were extracted by concatenating histograms of local binary patterns (LBP) and local binary pattern variances (LBPV) from image pyramids, and an BP neural network was used for classification. Yuan et al. [13] presented sub-oriented histograms of LBP for image smoke classification. Gubbi et al. [14] proposed a video smoke detection algorithm based on wavelet and Support Vector Machines (SVM) classification. Liao et al. [15] proposed Dominant Local Binary Patterns (DLBP) for texture classification by regarding the more frequently occurred patterns as dominant features. Guo et al. [16] proposed a Completed LBP (CLBP) approach, which encoded the magnitudes and signs of differences between a center pixel and its neighbors. CLBP provides excellent classification performance.

Above-mentioned methods extract features in spatial domains. Many methods achieve robust features from transform or frequency domains. Elaiwat et al. [17] proposed a multimodel Curvelet-based method for textured 3D face recognition. Each keypoint was detected across number of frequency bands and angles on 3D faces. Ucar et al. [18] presented an algorithm that was for facial expression recognition by integrating Curvelet transform and online sequential extreme learning machine (OSELM) with radial basis function (RBF) hidden node having optimal network architecture.

Although Curvelet transform provides a powerful multi-scale capability to extract discriminative smoke features, Curvelet-based image classification methods are limited to features, since the Curvelet coefficients are regarded as a holistic features extracted from the whole images [19]. To this end, we propose a duplex feature coding approach based on Curvelet transform to extract features from interpolated smoke images.

Many papers have been proposed to optimize kernel functions. Chapelle et al. [20] devised a gradient-based algorithm, which optimized a kernel function with multiple unconstrained parameters for SVM. Ghiasi-Shirazi et al. [21] considered the problem of optimizing a kernel function over translation invariant kernels for the task of binary classification. Wu et al. [22] proposed a direct method to build sparse kernel learning algorithms by adding one more constraint to the original convex optimization problem for sparse large margin classifiers. Ye et al. [23] considered the problem of multiple kernel learning (MKL) for regularized kernel discriminant analysis (RKDA), in which the optimal kernel matrix was obtained as a linear combination of pre-specified kernel matrices. All above methods formulated the kernel learning problem as an optimization problem based on a special task, such as SVM.

3. Our Algorithm

The framework of our method is shown in **Fig. 1**. Our method consists of four main steps: Curvelet transform of original images, extraction of Dual-encoded Local Binary Patterns (Dual-LBP) on Curvelet coefficient sub-images, concatenation of histograms of Dual-LBP and Completed Local Binary Patterns (CLBP), and Gaussian Kernel Optimization (GKO) of SVM classification.

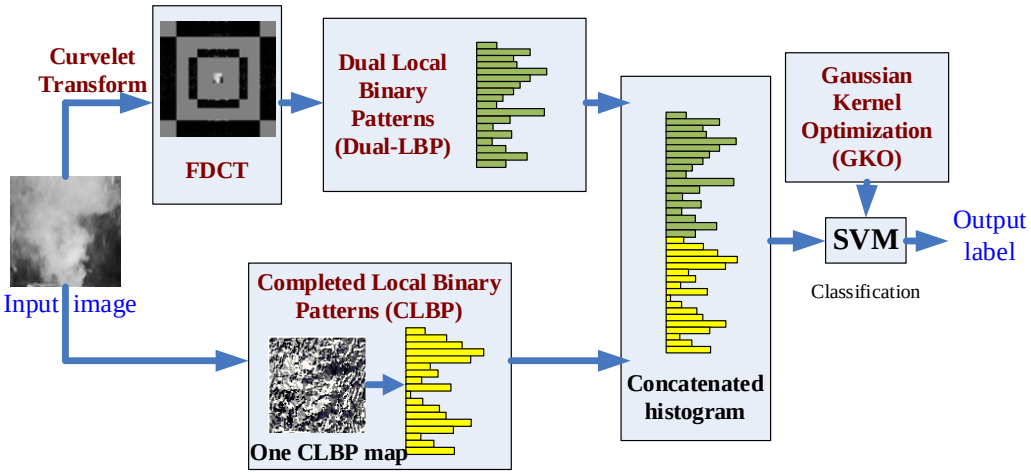


Fig. 1. The proposed smoke recognition framework

3.1 Curvelet transform

Curvelet transform was first proposed and structured with the tight frame by Candes and Donoho in 1999 [24]. Motivated by the need of image analysis, the second generation Curvelet transform [25] was introduced in 2005. It is not only simpler, but also faster and less redundant. Curvelets exhibit highly anisotropy and commendable directionality, which are beneficial for image edge representation. Smoke image edges are always curved, so Curvelet is almost the optimal representation of a singular smooth curve.

A pair of window functions, which are called “radial window” and “angular window”, are defined as $W(r)$ and $V(t)$. These windows meet the following admissibility conditions:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, r \in (\frac{3}{4}, \frac{3}{2}) \tag{1}$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, t \in (-\frac{1}{2}, \frac{1}{2}) \tag{2}$$

Then, the frequency window U_j is defined:

$$U_j(r,\theta) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{2^{\lfloor j/2 \rfloor}\theta}{2\pi}\right) \tag{3}$$

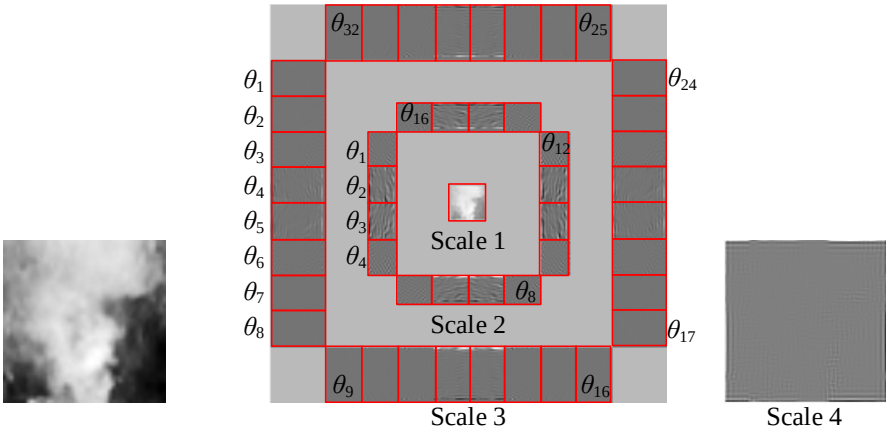
where $\lfloor j/2 \rfloor$ represents the integer part of $j/2$. Hence, the support of U_j is a polar "wedge" that is defined by the support of W and V. Varying scales j and directions U produce multi-scale and multi-direction transform.

These digital transforms are linear. We take a Cartesian array $f[t_1, t_2]$ ($0 \le t_1, t_2 < n$) as input and get an output of digital coefficients from the digital Curvelet transform. The digital Curvelet coefficients are defined:

$$c^D(j,\theta_l,k) = \sum_{0 \le t_1, t_2 < n} f[t_1,t_2] \overline{\varphi^D_{j,\theta_l,k}[t_1,t_2]} \tag{4}$$

where each $\varphi^D_{j,\theta_l,k}$ is a digital mother Curvelet (the superscript D represents "digital"), t_1 and t_2 are spatial variable, and j, θ_l and k are scale, orientation, and position index, respectively.

In the first step of our method, bilinear interpolation is needed to generate normalized images of size 128×128 from original images with different sizes. The scale number of subbands is set to $log_2(\min(w, h)-3)$, where w and h are width and height of input images, respectively. Hence, the scale is 4 for a 128×128 image. Digital Curvelet coefficients are real-valued. The multi-resolution Curvelet transform of different scales have different characteristics. Lower scales, denoted as coarser scales, contain low frequency information whereas higher scales, known as detailed and finer scales, consist of high frequency information.



(a) An interpolated image (b) Curvelet coefficients at 4 scales

Fig. 2. The four-scale Curvelet coefficients of a smoke image

To implement Curvelet transforms, we first perform a 2D FFT on the interpolated 128×128 image. Then the 2D Fourier frequency plane of the image is divided into many parabolic wedges. Finally, an inverse FFT of each wedge is applied to find the Curvelet coefficients at each scale j (j=1,2,3,4) and angle θ_l, and the range of l varies at different scales. An example of Curvelet coefficients at each scale is shown in **Fig. 2**. A red rectangular box stands for the coefficient map at one scale on one direction. The coefficient at scale 1 is displayed in the center. The coefficients at scale 2 on 8 directions and those at scale 3 on 16 directions are displayed in two loops around scale 1. Each block is equivalent to the pseudo polar tiling of the frequency plane with trapezoids.

There are two different digital implementations of Fast Digital Curvelet Transform (FDCT), which are based on Unequally Spaced Fast Fourier Transform (USFFT) and Wrapping Transform, respectively. In this paper, we use wrapping based Curvelet for feature extraction. The procedure of Curvelet based on wrapping is as follow:

(1). Apply the 2D FFT and obtain Fourier samples

$\hat{f}[n_1, n_2], -n/2 \le n_1, n_2 \le n/2$, where n_1 and n_2 are frequency-domain variable.

(2). For each scale j and angle θ_l, form the product

$$\tilde{U}_{j,\theta_l}[n_1, n_2]\hat{f}[n_1, n_2]$$

(3). Wrap this product around the origin and obtain

$\tilde{f}_{j,l}[n_1, n_2] = W(\tilde{U}_{j,\theta_l}\hat{f})[n_1, n_2]$, where $0 \le n_1 < L_{1,j}$ and $0 \le n_2 < L_{2,j}$ for $\theta_l \in (-\pi/4, \pi/4)$.

(4). By applying the inverse 2D FFT to each $\tilde{f}_{j,l}$, discrete coefficients $c^D(j,l,k)$ are obtained.

According to the above process, as shown in **Fig. 2**, we obtain a set of coefficient maps with four scales and sixteen directions from a normalized smoke image. Thus, we obtain coefficient maps containing coarse-to-fine and multi-directional texture information. The first and fourth scales contain only one coefficient map, and the second and third scales contain sixteen and thirty-two coefficient maps, respectively. It is worth noting that coefficient maps are in different sizes.

Being different from other traditional multi-scale transforms like wavelet transform, the coefficient map generated by Curvelet contains directional information of smoke, elevates ability to represent smoke textures and singularities along smoke edges.

To extract features from these coefficient maps, we propose Dual-encoded Local Binary Patterns (Dual-LBP) to get information on each coefficient map.

3.2 Dual-encoded Local Binary Patterns

LBP is a gray-scale texture descriptor and can achieve rotation invariance after being mapped to RI (Rotation Invariant) pattern [2]. LBP captures spatial structures of textures in an image by encoding differences between one central pixel and its local neighborhood. However, structural frequency information is not involved in LBP codes. To solve this problem, we use LBPs to extract frequency structures of images from Curvelet coefficient maps with different scales, orientations and locations.

In the Curvelet coefficients, the first scale contains only one coefficient map $c(1,1)$, the second one contains sixteen coefficient maps $c(2,l)$ (l=1, 2,…, 16), and the third one contains thirty-two coefficient maps $c(3,l)$ (l=1, 2,…, 32), and the fourth scale also contains only one coefficient map $c(4,1)$. We compute LBP maps from coefficient maps of all scales. These LBP maps on coefficient maps can capture variations of coefficients in a local region for all scales. To avoid interpolation of coefficients, we employ a 3×3 rectangular neighborhoods instead of circular neighborhoods to compute an LBP codes as follows:

$$map_{m,n}(j,l) = \sum_{p=0}^{P-1} s(c_{m,n}^{p}(j,l) - c_{m,n}(j,l))2^{p} \tag{5}$$

where $c_{m,n}(j,l)$ denotes the value of a central point (m, n) in a coefficient map $c(j, l)$, $c_{m,n}^{p}(j,l)$ is the value of the pth neighbor of the center point (m, n), P is the number of neighbors, $s(x)$ is a binarization function that returns 0 for negative values and 1 otherwise, and $map_{m,n}(j, l)$ is just an original LBP code at pixel (m, n) for the coefficient map $c(j, l)$.

Since we have a set of coefficient maps $c(j, l)$, we obtain 50 LBP maps $map(j, l)$ from these coefficient maps. LBP codes of each coefficent contain contrast information in local regions.

We compute the histogram of each coefficient LBP code map $map(j, l)$, formulated as follows:

$$\mathbf{H}_{j,l}(b) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\delta(map_{m,n}(j,l)-b) \tag{6}$$

where $\delta(x)$ is a function that returns 1 for x=0 and 0 otherwise.

We can obtain a lot of histograms $\mathbf{H}_{j,l}\in\mathbf{R}^{256\times1}$ from a set of coefficient LBP maps $map(j, l)$. In our implementation, the first scale has only one histogram $\mathbf{H}_{1,1}$, the second one generates 16 histograms $\mathbf{H}_{2,l}$ (l=1,…,16), the third one obtains 32 histograms $\mathbf{H}_{3,l}$ (l=1,…,32), and the fourth scale also has only one histogram $\mathbf{H}_{4,1}$. Therefore, we have 50 histograms. To combine information from different scales and orientations, we aggregate all these histograms together to form an LBP histogram map of size 256×50, formulated as follows:

$$\mathbf{M}=[\mathbf{H}_{1,1}, \mathbf{H}_{2,1},\ldots, \mathbf{H}_{2,16}, \mathbf{H}_{3,1},\ldots, \mathbf{H}_{3,32}, \mathbf{H}_{4,1}] \tag{7}$$

where $\mathbf{H}_{j,l}$ is just a column vector of the histogram map $\mathbf{M}$, which represents the histogram of each LBP map $c(j,l)$.

Hence, we obtain the new map $\mathbf{M}$ that is aggregated by normalized LBP histograms with 256 bins from fifty coefficient maps. Apparently, the size of the aggregated histogram map $\mathbf{M}$ is equal to 256×50.

In the second step of encoding, we apply the LBP encoding method again on the histogram map $\mathbf{M}$ to generate another LBP map, defined as follows:

$$\mathbf{E}_{m,n} = \sum_{p=0}^{P-1} s(\mathbf{M}^{p}_{m,n} - \mathbf{M}_{m,n})2^{p} \tag{8}$$

where $\mathbf{E}_{m,n}$ is an Dual-LBP code at a center point (m, n), $\mathbf{M}_{m,n}$ is the value of the aggregated histogram map at the center point, and $\mathbf{M}^{p}_{m,n}$ is the pth neighboring value of the center point.

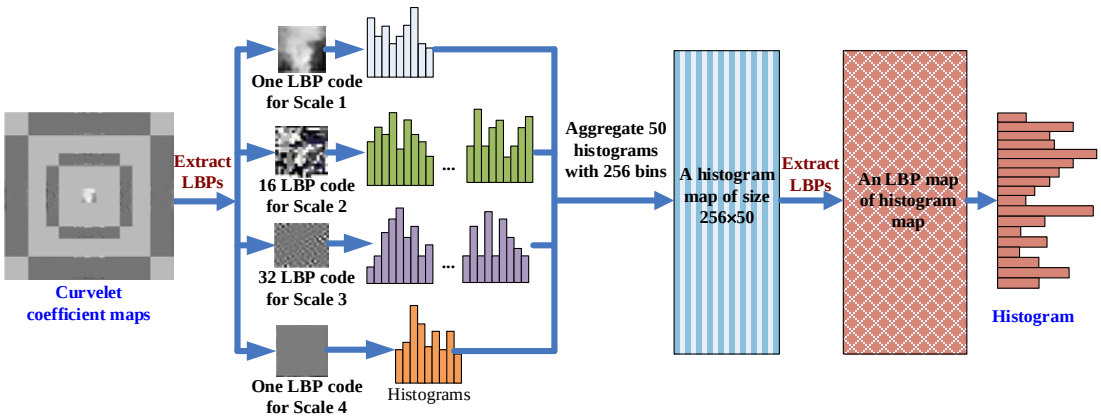


Fig. 3. Dual-Encoded features from Curvelet coefficient maps

Thus, we obtain another new LBP map from the histogram map $\mathbf{M}$. Then we compute the histogram of the histogram map $\mathbf{M}$. Dual-LBP extracts more details about frequent features from smoke images. The framework of Dual-LBP Encoding method is shown in **Fig. 3**.

3.3 Completed Local Binary Patterns

The original LBP is a computationally simple and efficient operator, but it only computes differences between a center pixel value and its corresponding neighbors' gray values. The original LBP operator discards the magnitudes of differences by encoding the signs of differences in a 3×3 rectangular neighborhood.

Guo et al. [16] proposed CLBP, an extension of the original LBP operator. The CLBP operator contains three operators, which are denoted as CLBP_S, CLBP_C and CLBP_M,

respectively. The CLBP_S operator is just the same as the original LBP operator, which encodes the sign of local differences to reflect directions of local gradients. While CLBP_M involves the magnitudes to preserve variance information. CLBP_C encodes the differences between local center pixels and the global one to represent whole image gray levels. CLBP_C and CLBP_M are defined as follows:

$$CLBP\_C_{P,R} = s(g_c - c_1) \tag{9}$$

$$CLBP\_M_{P,R} = \sum_{p=0}^{P-1} s(m_p - c_2)2^p \tag{10}$$

where c_1 is the average gray level of the whole image, c_2 is the mean difference magnitude of local neighborhood, g_c is the gray level of the center point, and m_p is the magnitude of the pth local difference, $s(x)$ is defined as a binarization function, which is the same as Eq. (5).

3.4 Final features

We use CLBP to obtain three kinds of codes, which are CLBP_S, CLBP_C and CLBP_M, for each pixel. We compute a joint 2D histogram of CLBP_M and CLBP_C, and then reshape the 2D histogram to a 1D histogram. Finally, we concatenate the 1D histogram with the histogram of CLBP_S to obtain the histogram $\mathbf{H}_{CLBP}$ of CLBP. The histogram $\mathbf{E}_{m,n}$ of Dual-LBP method captures frequency features of images in Curvelet domains. The CLBP method is encoded textures of images in spatial domains. We think that visual characteristics of smoke can be better captured if we combine spatial and frequency information. Hence, the final histogram $\mathbf{H}$ is obtained by combining $\mathbf{E}_{m,n}$ and $\mathbf{H}_{CLBP}$, formulated as follows:

$$\mathbf{H}=[\mathbf{E}_{m,n}, \mathbf{H}_{CLBP}] \tag{11}$$

After extracting features, we will consider the issue of features classification. We input the obtained histogram $\mathbf{H}$ into SVM for training and testing.

Since Curvelet coefficients contain components of different frequency, which correspond to different spatial distribution, they reflect spatial texture structure. Dual LBP models the relations between different coefficients to intrinsically captures co-occurrence texture structure. In other words, the proposed Dual LBP describes smoke textures in a macroscopic view [26].

3.5 Classification using SVM with GKO

We used Support Vector Machines (SVM) [27] to solve the image smoke classification problem. SVM is widely used in different fields such as clustering, classification, and dimensionality reduction. SVM is divided into two forms, which are linearly separable and linearly inseparable, respectively. Here we involve kernel trick to deal with linear inseparable features. Kernel trick is thus a way to implicitly transform linear inseparable features of data onto a new space where the data becomes linearly separable [28]. The implicit new space is always higher-dimensional (possibly infinite) [29]. In general, the Gaussian kernel function, also known as Radial Basis Function (RBF) [30], to describe the relationship between every two feature vectors, as shown in Eq. (12),

$$\mathbf{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2}{2\sigma^2}\right) \tag{12}$$

where $\mathbf{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the correlation or similarity between each two features $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ that are histograms $\mathbf{H}$.

The earliest method of optimizing $\beta=2\sigma^2$ is to use cross-validation or grid search. One of the most well known methods is leave-one-out, which leaves only one sample as the test set and

the remaining samples as the training set. Because each sample is repeatedly used during iterations, the method consumes a large amount of computation time. Hence, we use Gaussian Kernel Optimization (GKO) [31] to optimize β in our experiments.

GKO is a kernel optimizing method for unsupervised learning, which is different from optimized methods of other supervised learning. The GKO method does not need any constraints, and the β value obtained by the GKO method can be used as a starting point for further optimization. Hence, we use the GKO method to calculate the optimal value of β in eq. (12). We define random variable $Y_{ij}=X_{ij}^2/\sigma^2$ that satisfies the non-central Chi-square distribution with a degree of freedom one, where $X_{ij}=\|\boldsymbol{x}_i-\boldsymbol{x}_j\|$ and the variance $\sigma^2=\frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{j=1}^{n-1}(X_{ij}-\mu)$. The optimal value of β can be obtained by the following equation:

$$\beta \approx \begin{cases} \sigma^2/2.6 & \lambda \le 0.01 \\ \sigma^2/L(\lambda) & 0.01 < \lambda < 100 \\ \lambda\sigma^2 = \mu^2 & \lambda \ge 100 \end{cases} \tag{13}$$

where $L(\lambda)$ represents a function of λ. The relationship between λ and $L(\lambda)$ is shown in **Fig. 4**, where $\lambda=(\mu/\sigma)^2$, and $\mu=\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}X_{ij}$ represents the mean of the data set. The detailed proof of Eq. (13) is provided in [31].

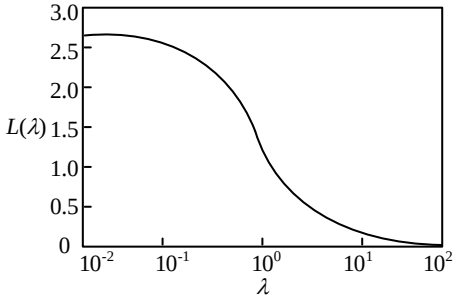


Fig. 4. Relationship between λ and $L(\lambda)$ when $\lambda \in (0.01, 100)$

According to the above method, β is optimized. In our smoke recognition experiments, *gamma* $=1/\beta=4.6283$ in the kernel function and the cost c=35 in the loss function.

4. Experimental Results and Analysis

4.1 Data sets

Several experiments were conducted on four data sets, each of which has an imbalanced number of smoke and non-smoke images. All images were manually cropped, resized and labeled as smoke images or non-smoke images. Smoke images of the data sets are easily distinguished by human eyes. The data sets are available at http://staff.ustc.edu.cn/~yfn/index.html. Smoke images of all datasets were resized to the size of 48×48 and converted to grayscale images for feature extraction. **Table 1** lists the details of the data sets. We used Set1 for training, and Set2, Set3, and Set4 for testing. Some samples are shown in **Fig. 5**. It can be seen that both intra-class and inter-class variances of smoke and non-smoke images are very large.

Table 1. The image datasets

Data sets	Total number of images	Number of smoke images	Number of non-smoke images	Usage
Set1	1383	552	831	Training
Set2	1505	688	817	Testing
Set3	10712	2201	8511	Testing
Set4	10617	2254	8363	Testing

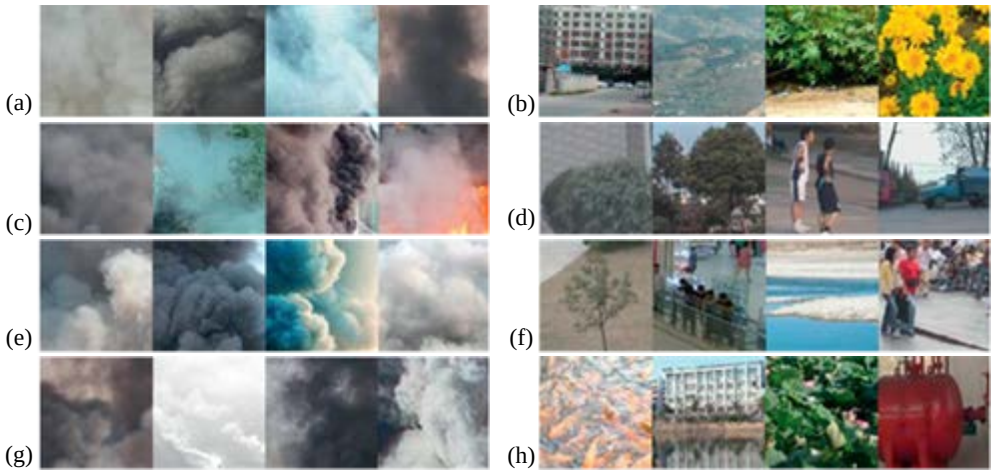


Fig. 5. Samples from the four data sets. (a) Smoke and (b) non-smoke images from Set 1. (c) Smoke and (d) non-smoke images from Set 2. (e) Smoke and (f) non-smoke images from Set 3. (g) Smoke and (h) non-smoke images from Set 4.

4.2 Implementation of compared methods

In order to verify the effectiveness of our method, we compared our method with some state-of-the-art algorithms by the three evaluation criteria in [32], which are Detection Rate (DR), False Alarm Rate (FAR) and Error Rate (ERR). They are defined as follows:

$$\begin{aligned} DR &= P_p / Q_p \times 100\% \\ FAR &= N_p / Q_n \times 100\% \\ ERR &= \left(Q_p - P_p + N_p\right) / \left(Q_p + Q_n\right) \times 100\% \end{aligned} \tag{14}$$

where P_p and N_p respectively denote the numbers of accurately detected true positive samples and negative samples mistakenly classified as positive samples, and Q_p and Q_n are the numbers of positive and negative samples, respectively.

4.3 Analysis of results

In our experiments, we used several feature extraction methods to validate the ability of our method to distinguish between smoke and non-smoke images on the three test sets. These compared methods are DRLBP [33], CLBP [16], LDBP [34], PLBP [35], PRICoLBP [36], MDLBP [37] LTrP [38] and DFD [39]. The compared LBP variants are all un-mapped for fair comparisons. The threshold for LTrP is set to 0.1 to demonstrate better performance, and g for RBF in SVM is set to 1/1383 for all other comparison features. For DFD, default setting is adopted to extract features.

We involve LBP and CLBP in our feature extraction step. Dual-LBP features based on the Curvelet domain and CLBP features on spatial domain are combined to form the final feature. In our CLBP, histograms of sign component and joint histograms of magnitude and center pixel maps are cocatenated to form CLBP_S_M/C. Finally, we aggregate dual LBP and CLBP features (denoted as Dual-LBP + CLBP) as our final feature vector, whose dimension is 256+768=1024.

Table 2. Experimental results for smoke detection

Methods	Set2(%)			Set3(%)			Set4(%)			Year
	DR	FAR	ERR	DR	FAR	ERR	DR	FAR	ERR	
Our method	96.7	**1.47**	**2.33**	95.5	**2.81**	**3.15**	95.7	**2.64**	**3.00**	**N.A.**
DRLBP	95.8	2.82	3.46	95.0	3.30	3.65	95.1	2.80	3.25	2016
CLBP	96.4	2.33	2.92	95.6	3.54	3.71	94.5	2.98	3.50	2010
PLBP	95.1	2.69	3.72	94.0	6.98	6.79	93.4	6.07	6.19	2011
LDBP	97.5	4.65	3.65	97.2	6.07	5.40	97.8	5.48	4.78	2017
PRICoLBP	95.5	2.57	3.46	94.0	3.69	4.17	94.5	3.23	3.71	2014
MDLBP	**97.7**	4.16	3.32	**97.3**	3.83	3.60	**97.2**	3.29	3.18	2016
LTrP	97.4	3.18	2.92	94.9	3.61	3.91	94.8	3.72	4.04	2012
DFD	92.0	4.28	5.98	92.1	11.90	11.10	91.4	11.30	12.00	2014

From **Table 2**, we find that our method achieves lower FARs than other methods on three testing data sets. MDLBP involves information across RGB channels, so it obtains the best DRs among all the methods. While all the other LBP variants are conducted on grayscale images. So it does not provide fair comparisons.

At the same time, the DRs got by our method are not obviously higher than other methods. Hence, ROC (Receiver Operating Characteristic Curve) is adopted to present a more comprehensive comparison, as shown in **Fig. 6**. By varying classification threshold t from -1 to 1 at step 0.1, DR and FAR pairs are obtained at every step to plot ROC.

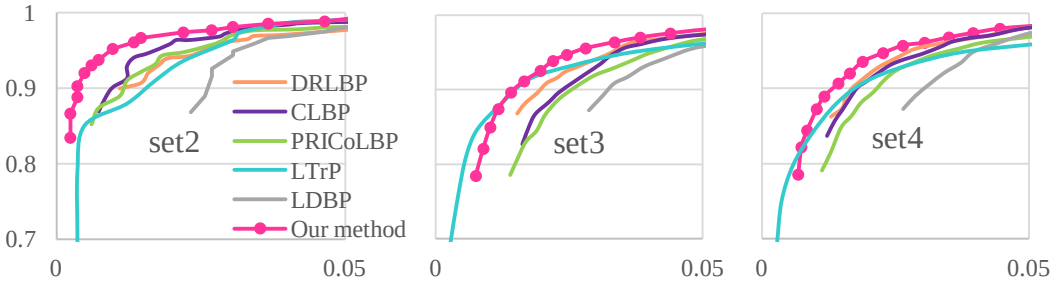


Fig. 6. ROCs of comparison methods on Set2, Set3 and Set4.

Although the DRs of our method do not exceed the ones of other methods obviously, the ROCs illustrate that our method outperforms others, which means that the best classification planes are not always at t=0.

The encoding step in our method can be replaced by any LBP-based methods. For instance, in **Table 2** and **Fig. 6**, Dual-LBP + CLBP is adopted. Similarly, the other three combinations are Dual-LBP + LBP, Dual-CLBP + CLBP, Dual-CLBP + LBP. The experimental results of the 4 combinations are shown in **Table. 3**.

Although the FAR of our method is not the lowest on Set3 and Set4, the DR of our method is highest and ERR is lowest. Overall, our Dual-LBP +CLBP performs best among all the

combinations.

Table 3. Comparisons of 4 combinations for our method

Methods	Set2(%)			Set3(%)			Set4(%)		
	DR	FAR	ERR	DR	FAR	ERR	DR	FAR	ERR
Dual-LBP+CLBP	**96.7**	**1.47**	**2.33**	**95.5**	2.81	**3.15**	**95.7**	2.64	**3.00**
Dual-LBP+LBP	96.1	2.45	3.12	93.9	2.60	3.32	94.4	**2.43**	3.11
Dual-CLBP+LBP	94.3	2.57	3.99	93.1	**2.84**	3.69	93.3	2.64	3.51
Dual-CLBP+CLBP	95.8	2.08	3.06	94.7	3.47	3.84	94.6	3.04	3.54

It is notable that Dual-CLBP + CLBP performs worse than others on Set3 and Set4. The reasons may be: 1) After Curvelet transform, an original image is decomposed into sub-bands. Low-frequency ones correspond to flat regions, in which the sign of gradient can better capture the invariance than magnitudes do. 2) There are correlations between Curvelet coefficients. Hence, the M and C components in CLBPs bring redundancy rather than improvement.

Lower FAR means lower accidental false alarm, which is of great significance for smoke classification, and it can reduce the serious consequences of false alarms. Therefore, our method is of great practical application value.

Table 4. Performance comparisons on GKO and other versions of SVM

Methods	DR(%)	FAR(%)	ERR(%)	Accuracy(%)	Time(s)	Total SVs
RBF-SVM with GKO	96.7	**1.47**	**2.32**	**97.67**	19.81 (18.55s GKO+1.26s SVM)	**207**
RBF-SVM with Grid Search	97.4	3.67	3.19	96.87	214.1	298
RBF-SVM without GKO	97.8	3.79	3.06	96.94	1.47	244

As shown in **Table 4**, we employ different parameter optimization methods to demonstrate the performance of GKO. We also compare our approach with the grid search, which is proposed in [30]. According to the experimental results, grid search method is proved not suitable for parameters optimization for different datasets. The GKO algorithm improves the accuracy of SVM.

Although the GKO step is time-consuming, it provides better classification performance and shorten the classifying time. Meanwhile, grid search consumes 214.1 seconds. Hence the GKO algorithm yields better performance than the grid search one. The computation time and the number of support vectors by the GKO algorithm are less than that of grid search on Set2.

5. Conclusion

In order to improve the performance of the smoke classification, we present a novel feature extraction method termed Dual-LBP, and we combine the proposed Dual-LBP and CLBP to improve the discriminative ability of features. The Dual-LBP method first adopts Curvelet transform to decompose smoke textures into coarse-to-fine components. Then LBP histograms are extracted from the decomposed components, *i.e.*, Curvelet coefficients, to generate a histogram map to describe local distributions of coarse-to-fine smoke textures. Third, LBP encoding is applied to the histogram map to capture texture distribution relations between different frequencies.

The advanced feature encoding is explored, which connected Curvelet domains and spatial domains. Furthermore, our method discovers the potential relationship between each scale of the Curvelet coefficients and improves the smoke classification performances. Extensive experiments show that our method achieves improvements in smoke recognition over some state-of-the-art methods.

References

[1] C. Palm, "Color texture classification by integrative co-occurrence matrices," *Pattern Recognition*, vol. 37, no. 5, pp. 965-976, May, 2004. Article (CrossRef Link).

[2] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, July, 2002. Article (CrossRef Link).

[3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 886-893, June 20-25, 2005.

[4] J. Ma and G. Plonka, "The Curvelet transform: A review of recent applications," *IEEE Signal Processing Magazine*, vol. 27, no. 2, pp. 118-133, March 2010. Article (CrossRef Link).

[5] A. Chenebert, T. P. Breckon, and A. Gaszczak, "A non-temporal texture driven approach to real-time fire detection," *in Proc. of 18th IEEE International Conf. on Image Processing*, pp. 1741-1744, September 11-14, 2011.

[6] T. H. Chen, P. H. Wu, and Y. C. Chiou, "An early fire-detection method based on image processing," in *Proc. of Int. Conf. on Image Processing*, pp. 1707-1710, October 24-27, 2004. Article (CrossRef Link).

[7] T. Çelik and H. Demirel, "Fire detection in video sequences using a generic color model," *Fire Safety Journal*, vol. 44, no. 2, pp. 147-158, February, 2009. Article (CrossRef Link).

[8] F. Yuan, "A fast accumulative motion orientation model based on integral image for video smoke detection," *Pattern Recognition Letters*, vol. 29, no. 7, pp. 925-932, May, 2008. Article (CrossRef Link).

[9] D. Zhang, S. Han, J. Zhao, Z. Zhang, C. Qu, Y. Ke, et al., "Image Based Forest Fire Detection Using Dynamic Characteristics with Artificial Neural Networks," in *Proc. of Int. Joint Conf. on Artificial Intelligence*, pp. 290-293, April 25-26, 2009. Article (CrossRef Link).

[10] C. Yu, J. Fang, J. Wang, and Y. Zhang, "Erratum to: Video fire smoke detection using motion and color features," *Fire Technology*, vol. 46, no. 3, pp. 651-663, July, 2010. Article (CrossRef Link).

[11] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Wavelet based real-time smoke detection in video," in *Proc. of 13th European Signal Processing Conf.*, pp. 1-4, September 4-8, 2005. Article (CrossRef Link).

[12] F. Yuan, "Video-based smoke detection with histogram sequence of LBP and LBPV pyramids," *Fire Safety Journal*, vol. 46, no. 3, pp. 132-139, April, 2011. Article (CrossRef Link).

[13] F. Yuan, J. Shi, X. Xia, Y. Yang, Y. Fang, and R. Wang, "Sub oriented histograms of Local Binary Patterns for smoke detection and texture classification," *Ksii Transactions on Internet & Information Systems*, vol. 10, no. 4, pp. 1807-1823, April, 2016. Article (CrossRef Link).

[14] J. Gubbi, S. Marusic, and M. Palaniswami, "Smoke detection in video using wavelets and support vector machines," *Fire Safety Journal*, vol. 44, pp. 1110-1115, November, 2009. Article (CrossRef Link).

[15] L. S, L. MW, and C. AC, "Dominant local binary patterns for texture classification," *IEEE Transactions on Image Processing*, vol. 18, no. 5, pp. 1107-1118, March, 2009. Article (CrossRef Link).

[16] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of Local Binary Pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, March, 2010. Article (CrossRef Link).

[17] S. Elaiwat, M. Bennamoun, F. Boussaid, and A. El-Sallam, "A Curvelet-based approach for textured 3D face recognition," *Pattern Recognition*, vol. 48, no. 4, pp. 1235-1246, April, 2015. Article (CrossRef Link).
[18] A. Uçar, Y. Demir, and C. Güzeliş, "A new facial expression recognition based on Curvelet transform and online sequential extreme learning machine initialized with spherical clustering," *Neural Computing & Applications*, vol. 27, no. 1, pp. 131-142, April, 2016. Article (CrossRef Link).
[19] T. Mandal, Q. M. J. Wu, and Y. Yuan, "Curvelet based face recognition via dimension reduction," *Signal Processing*, vol. 89, no. 12, pp. 2345-2353, December, 2009. Article (CrossRef Link).
[20] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for Support Vector Machines," *Machine Learning*, vol. 46, no. 1-3, pp. 131-159, May, 2002. Article (CrossRef Link).
[21] K. Ghiasi, R. Safabakhsh, and M. Shamsi, "Learning Translation Invariant Kernels for Classification," *Journal of Machine Learning Research*, vol. 11, pp. 1353-1390, April, 2010. Article (CrossRef Link).
[22] M. Wu, B. Schölkopf and G. Baktr, "A direct method for building sparse kernel learning algorithms," *Journal of Machine Learning Research*, vol. 7, pp.603-624, April, 2006. Article (CrossRef Link).
[23] J. Ye, S. Ji, and J. Chen, "Multi-class discriminant kernel learning via convex programming," *Journal of Machine Learning Research*, vol. 9, pp. 719-758, June, 2008.
[24] E. J. Candès and D. L. Donoho, "Recovering Edges in Ill-Posed Inverse Problems: Optimality of Curvelet Frames," *Annals of Statistics*, vol. 30, no. 3, pp. 784-842, June, 2002. Article (CrossRef Link).
[25] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete Curvelet transforms," *Multiscale Modeling & Simulation*, vol. 5, no. 3, pp. 861-899, September, 2006. Article (CrossRef Link).
[26] Q. Wang, M. Chen, F. Nie and X. Li, "Detecting coherent groups in crowd scenes by multiview clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.1-1, 2018. (Online) Article (CrossRef Link).
[27] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems & Their Applications*, vol. 13, no.4, pp. 18-28, July-August, 1998. Article (CrossRef Link).
[28] A. Temko, C. Nadue, "Classification of acoustic events using SVM-based clustering schemes," *Pattern Recognition*, vol. 39, pp. 682-694, April, 2006. Article (CrossRef Link).
[29] S. Yin, J. Yin, "Tuning kernel parameters for SVM based on expected square distance ratio," *Information Sciences*, vol. 370-371, pp. 92-102, November, 2016. Article (CrossRef Link).
[30] B. Scholkopf, K. Sung, C.J.C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing support vector machines with Gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2758-2765, November, 1997. Article (CrossRef Link).
[31] J. B. Yin, T. Li, and H. B. Shen, "Gaussian kernel optimization: Complex problem and a simple solution," *Neurocomputing*, vol. 74, pp. 3816-3822, November, 2011. Article (CrossRef Link).
[32] F. Yuan, J. Shi, X. Xia, Y. Fang, Z. Fang, and T. Mei, "High-order local ternary patterns with locality preserving projection for smoke detection and image classification," *Information Sciences*, vol. 372, pp. 225-240, December, 2016. Article (CrossRef Link).
[33] R. Mehta, K. Egiazarian, "Dominant Rotated Local Binary Patterns (DRLBP) for texture classification," *Pattern Recognition Letters*, vol. 71, pp. 16-22, February, 2016. Article (CrossRef Link).
[34] P. S. Hiremath, R. A. Bhusnurmath, "Multiresolution LDBP descriptors for texture classification using anisotropic diffusion with an application to wood texture analysis," *Pattern Recognition Letters*, vol. 89, pp. 8-17, April, 2017. Article (CrossRef Link).
[35] X. Qian, X. S. Hua, P. Chen, and L. Ke, "PLBP: An effective local binary patterns texture descriptor with pyramid representation," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2502-2515, October-November, 2011. Article (CrossRef Link).

[36] X. Qi, R. Xiao, C-G. Li, et al., "Pairwise rotation invariant co-occurrence local binary pattern," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2199-2213, April, 2014. Article (CrossRef Link).
[37] S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel decoded local binary patterns for content-based Image retrieval," *IEEE Transactions on Image Processing*, vol. 25, no.9, pp. 4018-4032, June, 2016. Article (CrossRef Link).
[38] S Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2874 - 2886, April, 2012. Article (CrossRef Link).
[39] Lei Z, Pietikainen M, Li S Z, "Learning discriminant face descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no.2, pp. 289-302, February, 2014. Article (CrossRef Link).

Feiniu Yuan received his B.Eng. and M.E. degrees in Mechanical Engineering from Hefei University of Technology, Hefei, China, in 1998 and 2001, respectively, and his Ph.D. degree in Pattern Recognition and Intelligence System from University of Science and Technology of China (USTC), Hefei, in 2004. From 2004 to 2006, he worked as a post-doctor with State Key Lab of Fire Science, USTC. From 2010 to 2012, he was a Senior Research Fellow with Singapore Bioimaging Consortium, Agency for Science, Technology and Research, Singapore. He is currently a professor and a PhD supervisor with Jiangxi University of Finance and Economics. His research interests include 3D modeling, image processing and pattern recognition.

Tiantian Tang received her B.E. degree in Communication Engineering from Institute of Technology, East China Jiaotong University, Nanchang, China, in 2011, and her M.E. degree in Signal and Information Processing from Jiangxi Science and Technology Normal University, Nanchang, China, in 2015. Her research interests include image processing and pattern recognition.

Xue Xia received her B.E. degree in Film & TV Arts and Technology and M.E. degree in Communication and Information Engineering from Shanghai University, Shanghai, China, in 2011 and 2014, respectively. She is currently a PhD candidate with School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. Her research interests include 3D display technology, image processing and pattern recognition.

Jinting Shi received her B.E. degree in Computer Science and Technology from Jiangxi Normal University, Nanchang, China, in 2003, M.S. degree in Computer Science and Technology from Jiangxi Agricultural University, Nanchang, China, in 2008, and Ph.D. degree in Management Science and Engineering from Jiangxi University of Finance and Economics, Nanchang, China. Her research interests include image processing and pattern recognition.

Shuying Li received the B.E.degree from University of Science and Technology of China, Hefei, China, and the Ph.D degree from Chinese Academy of Sciences. She is currently a professor with School of Automation, Xi'an University of Posts & Telecommunications, Xi'an, Shaanxi, China. Her research interest includes remote sensing, computer vision and pattern recognition.