A Noisy-Robust Approach for Facial Expression Recognition

Ying Tong[12], Yuehong Shen[1*] Bin Gao[1*], Fenggang Sun[1], Rui Chen[2] and Yefeng Xu[1]
[1] College of Communications Engineering, PLA University of Science and Technology
Nanjing, 210007, China
[e-mail: tong791015@gmail.com, chunfeng22259@126.com, feimaxiao123@gmail.com, sunfg@sdau.edu.cn，574645458@qq.com]
[2]Department of Communication Engineering, Nanjing Institute of Technology
Nanjing, 211167, China
[e-mail: tong791015@gmail.com，chenrui@njit.edu.cn]
*Corresponding author: Yuehong Shen and Bin Gao

Received May 29, 2016; revised October 20, 2016; accepted February 9, 2017; published April 30, 2017

Abstract

Accurate facial expression recognition (FER) requires reliable signal filtering and the effective feature extraction. Considering these requirements, this paper presents a novel approach for FER which is robust to noise. The main contributions of this work are: First, to preserve texture details in facial expression images and remove image noise, we improved the anisotropic diffusion filter by adjusting the diffusion coefficient according to two factors, namely, the gray value difference between the object and the background and the gradient magnitude of object. The improved filter can effectively distinguish facial muscle deformation and facial noise in face images. Second, to further improve robustness, we propose a new feature descriptor based on a combination of the Histogram of Oriented Gradients with the Canny operator (Canny-HOG) which can represent the precise deformation of eyes, eyebrows and lips for FER. Third, Canny-HOG's block and cell sizes are adjusted to reduce feature dimensionality and make the classifier less prone to overfitting. Our method was tested on images from the JAFFE and CK databases. Experimental results in L-O-Sam-O and L-O-Sub-O modes demonstrated the effectiveness of the proposed method. Meanwhile, the recognition rate of this method is not significantly affected in the presence of Gaussian noise and salt-and-pepper noise conditions.

Keywords: Facial expression recognition, anisotropic diffusion filter, histogram of oriented gradients

1. Introduction

Facial expressions are one of the most important human body language constituents, which can accurately express personal emotions, mental and psychological conditions. In recent years, computer analysis has been used to better understand human facial expressions and has presented an important prospect of application in human-machine interaction, which has drawn a lot of attention. For example, video cameras have become an integral part of many consumer devices and can be used for capturing facial images for recognition of people and their emotions. Generally, facial expression feature extraction and classification are the most critical aspects for accurate facial expression recognition (FER) [1]. Even though much work has already been done, higher FER accuracy at reasonable speeds remains a great challenge. In view of the categorization of different expression features affecting classification accuracy, in this paper we will focus on finding an efficient expression feature representation.

In the past decade, there has been a large amount of literature on developing feature extraction methods, which can be broadly divided into two categories: subspace-based methods [2-6,32-36] and local descriptor-based methods [7-13,30,37-40]. The main idea of the former is to seek the optical projection matrix with respect to a predefined criterion to project the original image onto a low-dimensional subspace. Principal Component Analysis (PCA) [2] and Independent Component Analysis (ICA) [3] are two traditional subspace methods for feature extraction, while many manifold learning algorithms for discovering intrinsic low-dimensional embedding of data were recently developed. Nikitidis et al. [4] proposed the Maximum Margin Projection for visual data analysis, while Gao et al. [5] presented an enhanced Fisher discriminant criterion to extract robust and efficient features. Furthermore, Lu et al. [6] proposed locality preserving projections for feature extraction. These methods can find meaningful embeddings for facial recognition, but prone to confront with the embarrassment of feature dimension, i.e., by using subspace-based methods, high-dimensional data spaces can be mapped to low-dimensional feature subspaces, while redundant features are removed to improve the efficiency of classification. However, such approaches need huge sample spaces. Therefore, for JAFFE-like databases, these methods are limited to mapping a valid low-dimensional subspace through a limited sample and thus the resulting feature description is not accurate, which leads to decreased recognition rates.

Compared with subspace-based methods, local descriptor-based methods seem to have also become popular in the literature. Among them, Gabor wavelets and Local Binary Patterns (LBP) are two representative methods. Gabor wavelets aim to encode multi-scale and multi-orientation information of facial expression images [7-8], which achieve good performance but at the cost of high computational loads. It would therefore be desirable to reduce dimensions and retain lower computational costs and feature sizes. LBP [9] describes facial expression texture information by encoding the gray-value differences between each pixel and its neighboring pixels into binary codes. Many variants of this approach have been

proposed [10] and applied to FER. For example, Zhao et al. [12] extracted the local texture feature by applying LBP to facial expressions, while Cao et al. [11] combined LBP with embedded Hidden Markov Models for facial expression recognition.

The Gabor wavelet transform can detect multi-scale and multi-directional expression texture information in local descriptor-based methods, but has a high computational cost and produces a large number of feature dimensions. The selection of the dimensionality reduction method will affect recognition accuracy. Compared with the Gabor wavelet transform, the LBP algorithm can extract facial features faster, as it has a strong texture discrimination ability and simpler calculations. However, the LBP operator is sensitive to non-monotonic illumination variation and also shows poor performance in the presence of random noise [13].

Despite the successful application of existing feature representation methods, any efficient such method is expected to satisfy the following three criteria: 1) It should be highly discriminative of inter-expression differences but robust to intra-personal variations. 2) It should be resistant to non-monotonic illumination variation and image noise. 3) The expression descriptors should have good performance for a reasonable feature size and low computational cost. Inspired by the Histograms of Oriented Gradients (HOG) method [14], we present a novel feature descriptor named Canny-HOG to address the limitations of existing techniques. HOG was originally developed to improve the performance of pedestrian detection systems [14], but has been used extensively in many applications. Tan et al. [15] utilized it as an eye state detector to determine whether or not the eyes are open or close. Minetto et al. [16] presented an effective gradient-based descriptor for single-line text regions. Gilanie et al. [17] extracted objects from T2 weighted brain MR images using gradient histograms. However, its application in the case of FER is limited. The key idea of HOG is to pool the local orientation information instead of the gray-value differences of small image patches, while expression characteristics are described by the deformation orientation of the eyes, the eyebrows, the lips and other facial components. Therefore, the HOG descriptor is more applicable to FER, as it is highly discriminative of inter-expression differences and robust to intra-personal variations compared with the traditional descriptors like LBP, Gabor and so on.

In this work, we present two variants of the HOG descriptor. First, we use the Canny operator instead of the original gradient operator for calculating the gradient magnitude and orientation in each pixel. As a result, it is more precise for pooling the edge orientation information of facial expression components. Then, we further readjust the size of small spatial regions called “cells”, and larger multi-cell regions called “blocks”. It would therefore be desirable to obtain a feature descriptor with superior performance that retains features of reasonable size and the implementation of which incurs a lower computational cost.

However, during image acquisition process, the image quality can be easily affected by system noise, light sources and other factors. Specifically in the case of FER, weaker muscle deformation textures and facial noise are not easily distinguished. Therefore, it is inadvisable

to extract features from the original images using the HOG descriptor directly, i.e. without pre-filtering. There are many image filtering methods, including averaging filters, median filters, Wiener filters etc. However they usually not only smooth-out the noise and blur the edges of the image, they also cannot retain the structural information of the original images [18]. In recent years, in order to protect the important structural characteristics from filtering operations, a nonlinear filter method has been developed by Perona and Malik, who proposed the anisotropic diffusion filter [19]. The principle is to use a large element size filter to reduce noise in flat areas with less gray value variations, and a small element size to make the feature prominent in edge areas with more gray-value variations. This approach has been widely used in many fields in recent years. Yan [20] proposed an enhancement processing on seismic image based on anisotropic diffusion. Chen [21] proposed a digital watermark algorithm combined with anisotropic diffusion filter and wavelet pyramid, which can ensure that the watermark is invisible and robust, while Fu [22] proposed a filtering method for medical images based on median filtering and anisotropic diffusion. Therefore, inspired by these articles, we present a novel filtering method based on an anisotropic diffusion filter for pre-processing facial expression images to reduce facial noise and retain facial muscle deformation textures.

To summarize the above arguments, the combination of reliable filtering and an effective feature descriptor is key to precise classification. In this paper, we propose a highly robust and discriminative two-step facial expression representation scheme, shown in **Fig. 1**. For the first step, the anisotropic diffusion filter is improved by adjusting the diffusion coefficient, which consists of two factors, namely, the gray-value difference between the object and the background and the gradient magnitude of the object. This allows us to overcome the disadvantages of the original anisotropic diffusion filter, and to both reduce facial noise and retain more complete facial muscles' wrinkle deformation textures, which contain significant expression information. For the second step, we adopt a Canny-HOG descriptor to extract facial expression features after filtering. Compared with the original HOG descriptor, the new descriptor is both precise and has lower computational cost. Thus, we can obtain an effective expression feature representation which has three merits: 1) high discrimination; 2) robustness to noise; and 3) low computational cost. Simulation results on images from the JAFFE and CK databases show that the combination of the improved anisotropic diffusion filter and Canny-HOG operator for FER can improve the recognition accuracy significantly compared with traditional methods like as LBP, Gabor and so on , as well as in Gaussian white noise and salt and pepper noise environments.

The remainder of this paper is organized as follows. Section 2 introduces the basic theory of the anisotropic diffusion (AD) filter and its improvement. Section 3 details the Canny-HOG descriptor compared with HOG descriptor. Section 4 presents the experimental results on the JAFFE and CK database images, while Section 5 concludes the paper.

Fig. 1. A scheme of our FER method

2 The original and improved AD filter

2.1 The original AD filter

When it comes to image pre-processing methods, some traditional filters, including Gaussian low-pass filters and Wiener filters, can remove noise but blur edges and lose crucial structure information of images in the process. To alleviate such problems, many researchers put forward new image filtering operators, such as the AD filtering method, which uses the directional diffusion coefficient instead of Gaussian convolution [19].

The AD equation (namely, P-M equation) is a heat conduction equation having the form of a partial differential equation, which is expressed as follows:

$$\frac{\partial f(x,y,t)}{\partial t} = div(c(\|\nabla f(x,y,t)\|)\cdot \nabla f(x,y,t)) \tag{1}$$

In this equation, $f(x,y,t)$ is the gray value of the input image $f(x,y)$ after the t-th iteration, $div(\cdot)$ is the divergence operation, while $c(\cdot)$ is the diffusion coefficient, calculated as follows:

$$c(\|\nabla f(x,y,t)\|) = \frac{1}{1+\left(\frac{\|\nabla f(x,y,t)\|}{K}\right)^2}, \tag{2}$$

where $\nabla f(\cdot)$ is the gradient operator, $\|\nabla f(x,y,t)\|$ is the gradient magnitude of $f(x,y,t)$, and K is the threshold value. As shown in (2), the diffusion coefficient $c(\cdot)$ is a monotonically decreasing function dependent on the gradient magnitude. When the gradient magnitude is larger, the diffusion coefficient is smaller, which allows the filter to retain the edge information by applying a weakly smoothing operator. On the contrary, when the gradient magnitude is smaller, the diffusion coefficient is larger, which corresponds to a stronger smoothing operator, helpful in reducing the image noise. Thus, the effect of the AD filter depends mostly on the diffusion coefficient.

2.2 The improved AD filter

We take into consideration that, apart from for the primary facial expression areas, like the eyes, the eyebrows and the mouth, different facial expressions will also cause various muscle deformations in non-primary facial expression areas, such as the chin, the forehead and the wings of the nose. Hence, texture details in these areas are one of the key factors of FER

[10,41]. However, the gradient magnitude of these non-primary facial areas is smaller, and thus the diffusion coefficient of traditional AD filter implementations is larger [42-44], which will result in loss of details and incomplete expression information. In view of this, we propose a new diffusion coefficient equation based on a regulatory factor and the gradient magnitude as follows:

$$c(\|\nabla f(x,y,t)\|) = \frac{1}{1+\left(\frac{\|\nabla f(x,y,t)\| + M(x,y,t)}{K}\right)^2} \tag{3}$$

where $\|\nabla f(x,y,t)\|$ is the gradient magnitude, $M(x,y,t)$ is a regulatory factor equal to $\alpha \cdot \Delta I(x,y,t)$, of which, α is the regulatory coefficient having an empirical value and $\Delta I(x,y,t)$ is the gray value difference between the object's value $f(x,y,t)$ and the background value $I_b(x,y,t)$, given by the following equation:

$$\Delta I(x,y,t) = f(x,y,t) - I_b(x,y,t) \tag{4}$$

More specifically, the background value $I_b(x,y,t)$ is not directly equal to the neighbouring value applied in the conventional approach, which is calculated using the following steps:

(1) Set (x,y) as the current pixel; its average value $f_m(x,y,t)$ is then calculated as follows:

$$f_m(x,y,t) = \frac{1}{9}\sum_{i=1}^{3}\sum_{j=1}^{3} f(x-2+i, y-2+j, t) \tag{5}$$

(2) By comparing the current value $f(x,y,t)$ and the average value $f_m(x,y,t)$ from two aspects, we calculate the background value $I_b(x,y,t)$. On one hand, if $f(x,y,t)$ is greater than $f_m(x,y,t)$, $I_b(x,y,t)$ is equal to the average value of the rest of the pixels in the 3×3 neighborhood whose gray values are less than $f_m(x,y,t)$; on the other hand, if $f(x,y,t)$ is less than $f_m(x,y,t)$, $I_b(x,y,t)$ is equal to the average value of the remaining pixels in the same neighborhood whose gray values are greater than $f_m(x,y,t)$. An example template and the corresponding process of calculating the background value $I_b(x,y,t)$ are shown in **Fig. 2**.

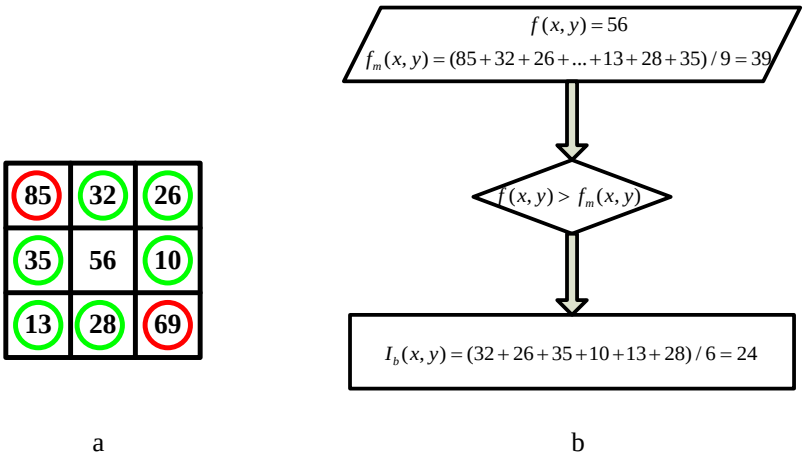


Fig. 2. Example template and its calculation process.
(a) Example template, (b) Background value calculation process

In **Fig. 2**, the current value $f(x,y)$ is 56 (namely the central pixel value) and the average value $f_m(x,y)$ is 39 (calculated through Equation (5)). Obviously, the current value $f(x,y)$ is greater than the average value $f_m(x,y)$, which means that the object is prominent and the background is non-prominent. Therefore, we choose some pixels whose values are less than $f_m(x,y)$, indicated by green circles in **Fig. 2(a)**, as the background pixels and calculate their average value as the result of $I_b(x,y)$ (shown in **Fig. 2(b)**). The advantage of this approach is that we can detect some odd pixels whose values are greater than $f_m(x,y)$, which are caused by noise, illumination and other factors in the 3×3 neighborhood. As shown in **Fig. 2(a)**, these odd pixels are indicated by red circles and are not consistent with the background pixels. Discarding these odd pixels, the background value $I_b(x,y)$ is obtained more accurately.

Thus, it can be seen that by combining the gray value difference $\Delta I(x,y,t)$ and the gradient magnitude $\|\nabla f(x,y,t)\|$, the diffusion coefficient is modified to distinguish the background noise and the muscle deformation texture correctly. This effect is analyzed from three aspects.

Firstly, in some primary facial expression areas, such as the eyes, mouth and eyebrows, the gradient magnitude $\|\nabla f(x,y,t)\|$ and the gray value difference $\Delta I(x,y,t)$ are larger, so the diffusion coefficient, as shown in Equation (3), is smaller, which results in the weakly smoothing effect of the improved anisotropic diffusion filter and protects important expression information.

Secondly, in the background areas with monotonous gray value variation, for example,

cheek, nose and hair, $\|\nabla f(x,y,t)\|$ and $\Delta I(x,y,t)$ are both smaller and the diffusion coefficient is larger, so the smoothing effect of the improved filter is stronger, and the filter can filter out background noise effectively.

Thirdly, in some non-primary facial expression areas, such as the chin, forehead and the wings of the nose, although the gradient magnitude $\|\nabla f(x,y,t)\|$ is not prominent, the gray value difference $\Delta I(x,y,t)$ can make up for the lack of gradient magnitude, and thus adjusts the diffusion coefficient properly and reduces the smoothing effect. In this manner, noise can be eliminated while retaining the expression details in the images, which is helpful for extracting the complete features of the facial expression image.

2.3 Comparison with two AD filters

In order to compare the advantages and disadvantages of the original AD filter (hereafter referred to as the P-M filter) and the improved AD filter (hereafter referred to as the improved P-M filter) more intuitively, we considered a facial expression image from the JAFFE database as an example. As is shown in **Fig. 3**, compared to the original face image (**Fig. 3 (a)**), the original P-M filter and the improved P-M filter can both remove the facial speckle noise effectively, as shown in **Fig. 3(b)** and **Fig. 3(c)**, but for the details' areas, such as the chin (outlined by a red box in the figures), the effect of the original P-M filter is not ideal. In the enlarged detail area of the original face image, shown in **Fig. 3(d)**, there are obvious muscle wrinkle deformations and facial noise. The original P-M filter not only removed the noise but also weakened the texture (as shown in **Fig. 3(e)**), while the improved P-M filter removed the noise effectively, yet preserving the expression details (**Fig. 3(f)**). This provides a foundation for the subsequent accurate extraction of expression features.

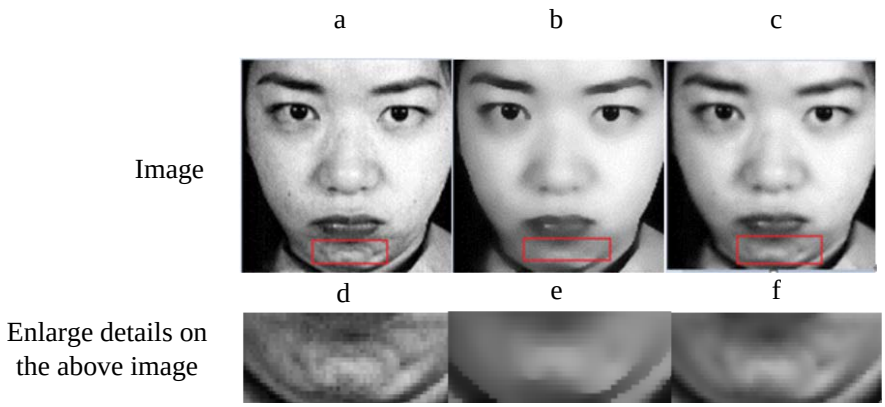


Fig. 3. Filtering results of the two AD filters. (a) original image, (b) Original P-M filtering result, (c) Improved P-M filtering result, (d)~(f) Enlarged details of corresponding images

3 HOG descriptor and its improvement

3.1 The original HOG descriptor

The Histogram of Oriented Gradients (HOG) is a very successful feature descriptor for object detection in computer vision. The key idea behind this descriptor is to pool the local orientation information instead of the magnitudes of small image patches. The image is divided into sets of small spatial regions called "cells", which form the basic units of this descriptor, and several neighboring cells constitute a larger local region called a "block". The HOG descriptor is formed as follows:

(1) Calculate the gradient magnitude and orientation of each pixel in the image.

(2) Divide the image into many cells in the same size.

(3) The local orientation histogram is extracted for every cell, and is accumulated with other histograms in other cells within the same block (but with different Gaussian-distributed weights according to the spatial distance). This is shown in **Fig. 4, (a)-(b)**.

(4) The histograms contained within each block are concatenated and normalized to form the final feature (**Fig 4. (c)**).

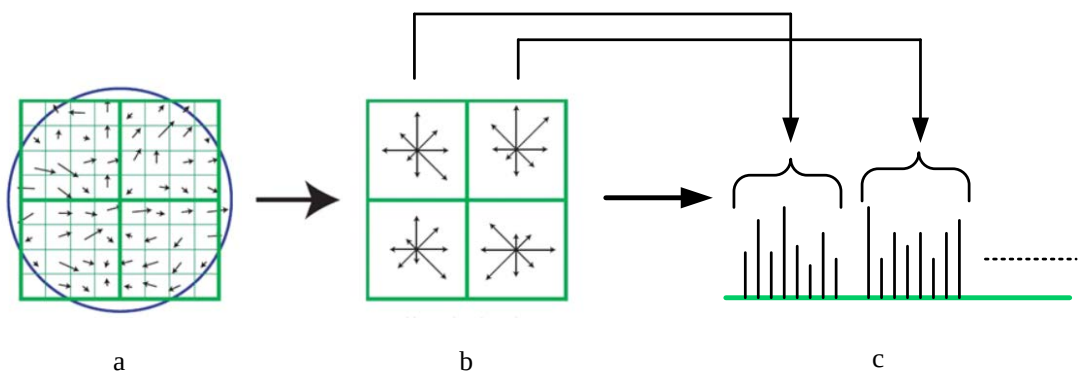


Fig. 4. Feature representation using the HOG descriptor.

3.2 Canny-HOG descriptor

In [14], the cell size and block size are set to 8×8 pixels and 16×16 pixels respectively, which means one block is made up of 2×2 cells, while the orientation space was equally divided into nine intervals. The gradient magnitude and orientation of each pixel was calculated using a 1-D central symmetric template [-1 0 1]. The corresponding equations are as follows:

$$G(x,y)=\sqrt{[f(x+1,y)-f(x-1,y)]^2+[f(x,y+1)-f(x,y-1)]^2}, \tag{6}$$

$$\theta(x,y) = \arctan \frac{f(x,y+1) - f(x,y-1)}{f(x+1,y) - f(x-1,y)}. \tag{7}$$

Due to the size of the database images being 128×128 pixels, each image was divided into 8×8 = 64 non-overlapping blocks and 2×2 = 4 cells per block, which is consistent with the parameter settings of the original study [14], so that the feature vector of the original HOG descriptor is 64 (blocks) × 4 (cells) × 9 (bins) =2304 dimensions. It is clear that the original HOG feature vector has high dimensionality, and there is a lot of redundant information on account of fragmentary segmentation (**Fig. 5(a)**, where only the initial cell in the figure is illustrated due to grid size limitations), which affects both the recognition accuracy and efficiency. In this paper, the sizes of cell and block were adjusted to reduce feature dimensionality, namely one block comprised 64×64 pixels and one cell containted 16×16 pixels, and consequently each block was composed of 4×4 cells (as shown in **Fig. 5(b)**). Thus, the feature vector based on the new division has 4 (blocks) × 16 (cells) × 9 (bins) =576 dimensions. This shows that the feature dimension is reduced by 75%, and the recognition accuracy and efficiency are both significantly improved.

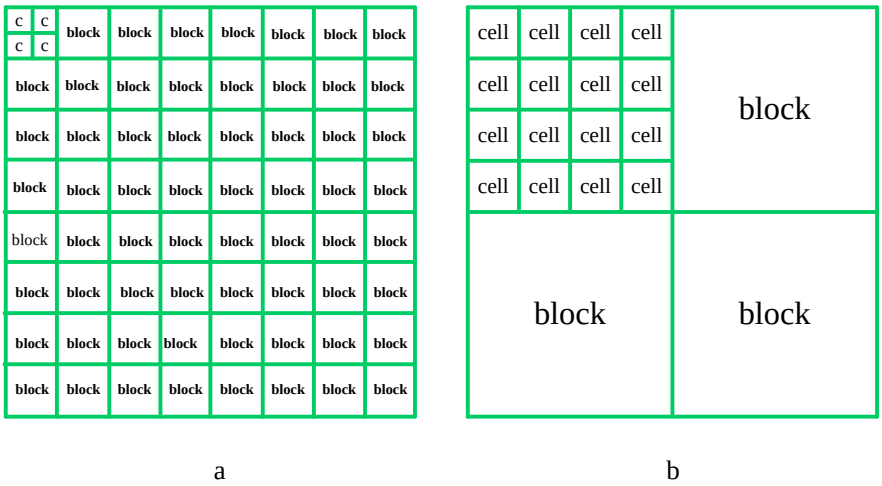


Fig. 5. Graphical representation on image division into blocks and cells.
(a) original division, (b) proposed division

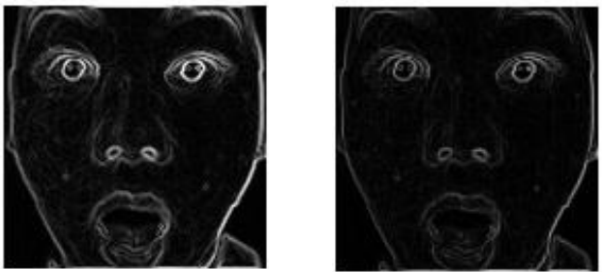
As a further variation, the Canny operator is used to replace the 1-D central symmetric template in this paper; the corresponding gradient calculation equations are shown in (8)-(11). We used these gradient operators on database images, with example results shown in **Fig. 6**. It can be seen that the facial noise of gradient magnitude images produced by the Canny operator (as shown in **Fig. 6(b)**) is less than that produced by the 1-D central symmetric template (**Fig. 6(a)**), while the edges are also more distinct. Therefore, feature extraction based on the variant is more discriminative of inter-personal differences but robust to intra-personal variations.

$$G_x(x,y)=\frac{1}{2}[f(x+1,y)-f(x,y)+f(x+1,y+1)-f(x,y+1)], \tag{8}$$

$$G_y(x,y)=\frac{1}{2}[f(x,y)-f(x,y+1)+f(x+1,y)-f(x+1,y+1)], \tag{9}$$

$$G(x,y)=\sqrt{[G_x(x,y)]^2+[G_y(x,y)]^2}, \tag{10}$$

$$\theta(x,y)=\arctan\frac{G_y(x,y)}{G_x(x,y)} \tag{11}$$



a b

Fig. 6. Gradient Magnitude Images. (a) Gradient Magnitude Images of 1-D central symmetric template (b) Gradient Magnitude Images of Canny operator

4 Experiments

In this section, we evaluate the role of image filtering and the validity of our feature extraction approach in FER extensively, with particular emphasis on robustness to noise. Experiments were conducted on two publicly available databases: the JAFFE database and the CK database.

The JAFFE database contains 213 images of Japanese female facial expressions classified using 6 classes, namely anger, disgust, fear, happiness, sadness and surprise. Each class has 10 subjects. The size of the images is 256×256 pixels. In this paper, we chose 3 images per subject for each expression, which resulted into 180 expression images. The CK database consists of 100 university students aged between 18-30 years old; 65% were female, 15% were African-American, and 3% were Asian or Latino. Subjects were instructed to perform a series of facial expression displays, starting from neutral or nearly neutral one to one of six target prototypic expressions. Each image is 640×480 pixels or 640×690 pixels. In this paper, we chose the five most expressive images taken from each sequence, which resulted into 1665 expression images.

The images contained in these two databases were all monotonous and are freely available for anyone to use. However, before the experiments, some preprocessing was necessary to crop the original images of databases based on the positions of the eyes, and the images were resized to 128×128 pixels. The template is shown in **Fig. 7,** and example images of each prototypic expression are shown in **Fig. 8**.

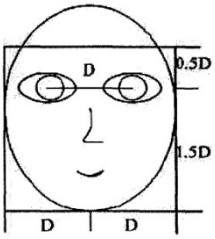


Fig. 7. Cropping template of facial expression image

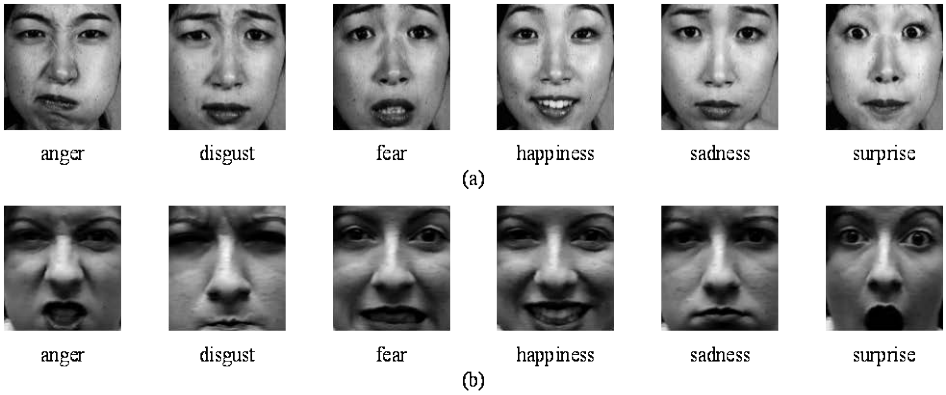


Fig. 8. Sample images of each prototypic expression from (a) JAFFE database (b) CK database

In this paper, we carried out the experiments using two strategies, namely the "leave-one-sample-out" (hereafter referred to as L-O-Sam-O) and "leave-one-subject-out" (hereafter referred to as L-O-Sub-O) [8]. The L-O-Sam-O is a familiar sample selection strategy. During each run, one image is selected as the test sample, while the remaining N-1 images are selected as the training samples (N is the number of total samples, equal to 180 for the JAFFE database sample set and to 1665 for the CK database set). N-cross validation was carried out for each sample so that the recognition result is more accurate. The L-O-Sub-O is an unfamiliar sample selection strategy. During each run, the test set contains all expression images of one subject, but the training set does not contain any images that correspond to the same test subject. We carried out M-cross validation (M is expressed as the total number of subjects, which equalled 10 in JAFFE database set, and 8 in the CK set). The result is more indicative of the ability of the system to generalize its performance and recognize a new person's expression. KNN was adopted to classify samples from these two databases.

Three experiments were conducted. First, the performance of the Canny-HOG descriptor was compared with five classic descriptors without any pre-filtering operations. Secondly, we discuss the performance of the combination of the improved P-M filter and Canny-HOG descriptor on FER compared with other combinations. Finally, the robustness of the proposed scheme is presented for recognizing the noisy database from two aspects.

4.1 Evaluation of Canny-HOG descriptor

In this section, the default parameters of original the HOG, LBP, Gabor, LDP and LDNP descriptors were used according to the corresponding references. **Table 1** lists the default parameters for each descriptor. The feature size and the feature extraction time for each descriptor are listed in **Table 2**. From the results, it is seen that the Canny-HOG descriptor's feature size is far smaller than other descriptors' and its feature extraction time is only 5.75s, which is desirable in engineering applications with lower computational costs. In addition, although the feature size of the LDNP is 4096, which is less than LBP, it needs to choose between positive and negative directions and encode the numbers, which results in higher computational costs.

Table 1. Default parameter settings for each descriptor

Descriptor	Parameters	Value
LBP[9]	Block size	16*16 pixels
	Bit numbers	8
Gabor[7]	Direction numbers	8
	Scale numbers	5
	Sample interval	4
LDP[13]	Block size	16*16 pixels
	Bit numbers	8
LDNP[28]	Block size	16*16 pixels
	Bit numbers	6
Original HOG[14]	Cell size	8*8 pixels
	Block size	2*2 cells
Canny-HOG	Cell size	16*16 pixels
	Block size	4*4 cells

Table 2. The feature size and feature extraction time of different descriptors

Descriptor	Feature size	Feature extraction time(s)
LBP[9]	16384	28.71
Gabor[7]	40960	3852.31
LDP[13]	16384	28.73
LDNP[28]	4096	23.41
Original HOG[14]	2304	8.14
Canny-HOG	576	5.75

The performance of different descriptors was evaluated using the JAFFE and CK database sets. The recognition rates for the two sets are presented in **Tables 3** and **4**. From the results, three main conclusions can be drawn:

(1) In **Table 3**, Canny-HOG shows the best performance on the JAFFE database set compared with other descriptors. First, Gabor is more robust to illumination variation, noise and other interference than other descriptors [8], but in this experiment Canny-HOG slightly outperformed Gabor by 0.55% and reduced the classification time by a factor of nearly 100 on the L-O-Sam-O strategy. Second, on the L-O-Sub-O strategy, Canny-HOG was significantly better than the other descriptors, achieving a maximum recognition improvement of over 13% compared to the worst performer. Thus, it can be seen that Canny-HOG captures facial expression information more accurately and is helpful for eliminating the different expressions' influence of the same person on face recognition field.

(2) Compared with **Table 3**, the performance of all the descriptors using the L-O-Sap-O strategy was better on the CK database images, where the recognition rates achieved by all descriptors were over 95%. This is because the five most expressive images taken from each CK sequence were similar. In **Table 4**, although the recognition rate of LDP is slightly above Canny-HOG on the L-O-Sap-O strategy, Canny-HOG significantly outperforms LDP by 14.59% on the L-O-Sub-O strategy, and the mean recognition rate of Canny-HOG is 84.9% which is higher than the other descriptors. In general, on the L-O-Sub-O strategy Canny-HOG is significantly better than other descriptors.

(3) To further illustrate the effectiveness of the Canny operator in Canny-HOG, we used the center symmetric template to replace the Canny operator without changing any other parameters. This variant is named the CS-HOG descriptor. The recognition rates of CS-HOG on JAFFE database were 96.11% and 60.00% on the L-O-Sap-O and L-O-Sub-O strategies respectively. Comparing with **Table 3**, the recognition rate of Canny-HOG on the L-O-Sap-O strategy was not decreased, but the recognition rate on the

L-O-Sub-O strategy dropped by 5%. Thus, the Canny operator is helpful in eliminating the influence of different human faces on FER.

Table 3. Recognition rates and classification times for different descriptors on JAFFE database set

Descriptor	L-O-Sap-O		L-O-Sub-O	
	Recognition rate (%)	Classification time	Recognition rate (%)	Classification time
LBP[9]	86.67	109.54	52.78	76.28
Gabor[7]	95.56	315.72	57.22	209.36
LDP[13]	89.44	24.23	59.44	16.43
LDNP[28]	88.33	25.13	58.33	21.84
Original HOG[14]	90.56	16.72	59.44	12.32
Canny-HOG	**96.11**	**3.14**	**65.00**	**2.83**

Table 4. Recognition rates and classification times for different descriptors on CK database set

Descriptor	L-O-Sap-O		L-O-Sub-O	
	Recognition rate (%)	Classification time	Recognition rate (%)	Classification time
LBP[9]	96.82	115448.58	63.89	73.09
Gabor[7]	96.78	276621.34	65.36	193.65
LDP[13]	**97.60**	25685.41	58.33	11.42
LDNP[28]	97.30	31292.41	62.50	11.89
Original HOG[14]	96.04	18087.29	72.22	7.26
Canny-HOG	96.88	**4251.02**	**72.92**	**1.81**

4.2 Evaluation of the performance of the combination of the improved P-M filter and Canny-HOG descriptor

We have also compared the filtering result of the improved P-M filter with the original P-M filter, as discussed in Section 2.3, where it was observed that the improved P-M filter can effectively distinguish the facial speckle noise and the texture details of muscle wrinkle deformations, thus providing a foundation for the subsequent accurate expression feature extraction. Meanwhile in Section 4.1 the performance of the Canny-HOG descriptor was shown to be superior to five other state-of the-art descriptors on the JAFFE and CK database sets. In this section, we compare the combination of the improved P-M filter and the Canny-HOG descriptor with other combinations on the JAFFE and CK database sets. The experimental results are shown in **Table 5**-**Table 8**. The following conclusions can be drawn:

(1) The original images were first pre-processed using the original P-M filter, and then the the facial expression features where extracted using spatial descriptors. The experimental results are shown in **Table 5** and **Table 7**. In the last column, we can observe that the performance of the original P-M filter and the Canny-HOG descriptor was slightly slightly compared with the Canny-HOG descriptor. The reason is the original P-M filter could not distinguish the facial noise and the texture details of muscle wrinkle deformations. It blurred the images and the Canny-HOG descriptor could not obtain as many desirable orientation features either on the JAFFE nor on the CK database sets.

(2) The original P-M filter was replaced with the improved P-M filter during the pre-processing step, and the experimental results of the JAFFE and CK database sets are shown in **Table 6** and **Table 8**. It is evident that the combination of the improved P-M filter and Canny-HOG descriptor performs better than the Canny-HOG descriptor with the original P-M filter. The recognition rates on the JAFFE database set were improved by 1.11% and 2.22% on the L-O-Sap-O and L-O-Sub-O strategies respectively. Furthermore, the recognition rates on the CK database set were improved by 0.48% and 2.08% on the L-O-Sap-O and L-O-Sub-O strategies respectively. The results indicate that the improved P-M filter can further remove interference information and provides a foundation for the subsequent accurate expression feature extraction.

(3) In **Table 8**, although the recognition rate of the improved P-M filter + LDP descriptor is slightly above that of the improved P-M filter + Canny-HOG descriptor on the L-O-Sap-O strategy, the improved P-M filter + Canny-HOG descriptor significantly outperforms the improved P-M filter + LDP descriptor by 13.89% on the L-O-Sub-O strategy, while the mean recognition rate of the improved P-M filter + Canny-HOG descriptor is 86.18% which is higher than other combinations. In general, on the L-O-Sub-O strategy the improved P-M filter + Canny-HOG descriptor is significantly better than other combinations.

Table 5. Recognition rates for different descriptors combined with the original P-M filter on the JAFFE database set

Recognition Rate(%)	Original P-M +LBP	Original P-M+LDP	Original P-M+LDNP	Original P-M+ Original HOG	Original P-M + Canny HOG
L-O-Sap-O(%)	93.33	93.89	90.56	92.78	**95.56**
L-O-Sub-O(%)	51.11	56.67	52.78	61.67	**65.56**

Table 6. Recognition rates for different descriptors combined with the improved P-M filter on the JAFFE database set

Recognition Rate(%)	Improved P-M +LBP	Improved P-M+LDP	Improved P-M+LDNP	Improved P-M+ Original HOG	Improved P-M + Canny HOG
L-O-Sap-O(%)	95	96.11	92.78	93.33	**97.22**
L-O-Sub-O(%)	55	55.56	53.33	64.44	**67.22**

Table 7. Recognition rates for different descriptors combined with the original P-M filter on the CK database set

Recognition Rate(%)	Original P-M +LBP	Original P-M+LDP	Original P-M+LDNP	Original P-M+ Original HOG	Original P-M + Canny HOG
L-O-Sap-O(%)	97.11	97.90	**98.02**	96.16	96.64
L-O-Sub-O(%)	61.11	61.81	63.89	**75**	72.22

Table 8. Recognition rates for different descriptors combined with the improved P-M filter on the CK database

Recognition Rate(%)	Improved P-M +LBP	Improved P-M+LDP	Improved P-M+LDNP	Improved P-M+ Original HOG	Improved P-M + Canny HOG
L-O-Sap-O(%)	96.16	**97.96**	97.78	96.46	97.36
L-O-Sub-O(%)	63.89	61.11	61.11	70.83	**75**

4.3 Noise robustness analysis of the improved P-M filter + Canny-HOG descriptor

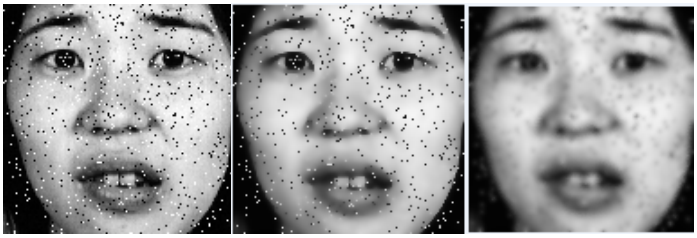
In the previous sections we discussed the validity of the proposed method, which has been shown to be suitable for FER without noise interference. In this section we primarily discuss the noise robustness of the proposed method from two aspects: firstly, noise robustness analysis of two filters, namely the original P-M filter and the improved P-M filter; secondly, noise robustness analysis of the combination of the improved P-M filter and the Canny-HOG descriptor. The experimental results will give further confirmation of the validity of the proposed method. In this section, experiments were performed using only the JAFFE images due to the high computational cost of involved for CK images.

4.3.1 Noise robustness analysis of the improved P-M filter

First, we analyze the two filters' performance on Gaussian white noise and salt & pepper noise environments. To better illustrate the advantages of the improved P-M filter, the same parameters for the two filters were chosen, namely, the number of iterations was 5 and the gradient threshold value was 15. In addition, the Gaussian white noise variance was 0.001, while the salt & pepper noise variance was 0.05, with a mean of 0 in both cases. The filtering results are shown in **Fig. 9** and **Fig. 10**, respectively. From the figures, it can be observed that the improved P-M filter is better than the original P-M filter both in Gaussian white noise and in salt & pepper noise environments. The significant reason is that the improved P-M filter makes use of the relative gray value difference to adjust the diffusion coefficient, which improves the filter's noise removal properties, as shown in **Fig. 9(c)** and **Fig. 10(c).** Thus, the improved P-M filter can obtain the desirable effect of with only a small number of iterations and speed up its convergence compared with the original P-M filter.

(a) (b) (c)

Fig. 9. Filtering results of the two AD filters. (a) Original image with Gaussian noise, (b) Original P-M filter, (c) Improved P-M filter result.



(a) (b) (c)

Fig. 10. Filtering results of the two AD filters. (a) original image with salt & pepper noise, (b) Original P-M filter, (c) Improved P-M filter.

4.3.2 Noise robustness analysis of the improved P-M filter + Canny-HOG descriptor

Finally, we analyze further the noise robustness of the combination of the improved P-M filter and the Canny-HOG descriptor. We added zero-mean Gaussian white noise with a variance of 0.01 and zero-mean salt & pepper noise with a variance of 0.05 to the images of the JAFFE database. From **Table 3** and **Table 4** it is evident that the Canny-HOG descriptor has the best performance compared with other spatial descriptors, therefore we only evaluated three cases: the Canny-HOG descriptor, the original P-M filter + Canny-HOG descriptor and the improved P-M filter + Canny-HOG descriptor. The experimental results are shown in **Table 9**. Our conclusions are twofold:

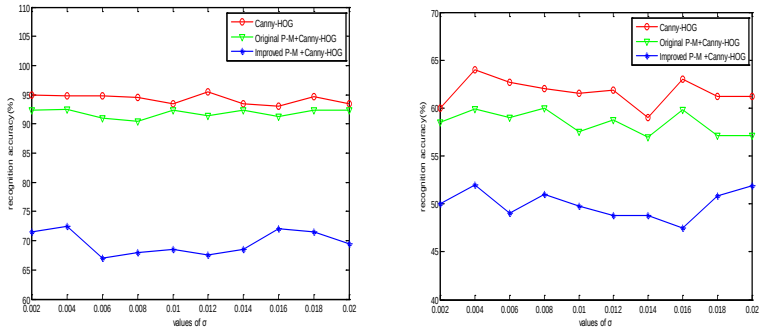
(1) Because the image quality deteriorated due to the presence of the Gaussian white noise and salt & pepper noise, the recognition rate of the Canny-HOG descriptor was sharply reduced. Especially on the L-O-Sam-O strategy, the recognition rate decreased nearly 25% compared with the results of **Table 3**. This indicates that the Canny-HOG descriptor by itself is adversely affected by noise interference.

(2) If the original P-M filter or the improved P-M filter are used process the noisy images before using the Canny-HOG descriptor to extract the expression features, the recognition rates are all considerably improved. The performance improvement obtained using the improved P-M filter + Canny-HOG was more pronounced. Thus, when the image quality is poor, preprocessing becomes particularly important, and the combination proposed in this paper is more conductive to FER in the presence of noise interference.

In order to further confirm the superiority of the proposed method, we added Gaussian white noise and salt & pepper noise with different variance levels to the JAFFE images, and used the KNN classifier to identify different expressions. The recognition rate curves of the three algorithms using the two kinds of classification strategies are shown in **Fig. 11** and **Fig. 12**. It can be observed that, overall, the curve of the proposed method is above other two curves and particularly higher than the curve of the Canny-HOG descriptor. Therefore, the combination of the improved P-M filter and Canny-HOG descriptor is an effective method for FER in different noise environments.

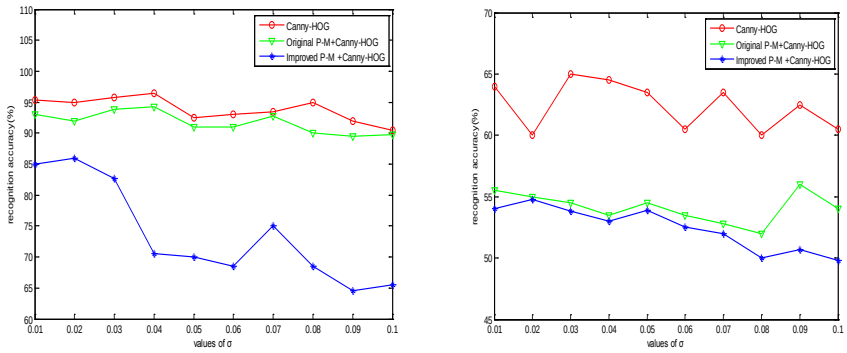
Table 9. Recognition rates for Canny-HOG and combinations with the P-M filter and the improved P-M filter in different noise environments.

		Canny-HOG	Original P-M + Canny-HOG	Improved P-M+ Canny-HOG
Gaussian	L-O-Sam-O(%)	71.11	91.67	**94.44**
	L-O-Sub-O(%)	51.67	61.67	**65.00**
Salt & Pepper	L-O-Sam-O(%)	73.89	93.33	**96.11**
	L-O-Sub-O(%)	56.11	61.11	**66.67**



(a) (b)

Fig. 11. Recognition rates for the above three algorithms on Gauss white noise with different variances. (a) Recognition rate of L-O-Sam-O classification strategy, (b) Recognition rate of L-O-Sub-O classification strategy



(a) (b)

Fig. 12. Recognition rates for the above three algorithms on Salt & Pepper noise with different variances. (a) Recognition rate of L-O-Sam-O classification strategy, (b) Recognition rate of L-O-Sub-O classification strategy

5 Conclusions

Due to the degradation of face image quality, FER is a challenging task. Solving this problem demands work in at least two areas: development of an effective filter distinguishing the expression texture details from the facial noise and an accurate feature extraction descriptor. To achieve this goal, the work presented in this paper proposed the following contributions:

(1) We presented an improved P-M filter which combined the relative gray value difference with the gradient magnitude and modified the diffusion coefficient to distinguish the background noise and the muscle deformation texture correctly.

(2) We presented the Canny-HOG descriptor instead of Canny operator to calculate gradient magnitude and gradient orientation, and the block and cell were adjusted accordingly. Experimentation on JAFFE and CK databases showed that the Canny-HOG descriptor is more discriminative of different expression and has lower computational costs.

(3) We combined the proposed filter and feature descriptor and evaluated their performance against other more traditional combinations on images of the JAFFE and CK databases, in both noisy and noiseless conditions. Detailed experiments show that the combination of the improved P-M filter and the Canny-HOG descriptor offers major advantages for FER.

Acknowlegements

This work is financially supported by NSFC Projects No. 61471162, Program of International science and technology cooperation (0S2015ZR1083), NSF of Jiangsu Province (BK20141389, BK20160781), NSF of Hubei Province (2014CFB589), Technology Research Program of Hubei Provincial Department of Education (D20141406), Key Laboratory of meteorological detection and information processing in Jiangsu Province (KDXS1503), and NIT fund for Young Scholar (CKJB201602).

References

[1] K.T. Song and S.C. Chien, "Facial Expression Recognition Based on Mixture of Basic Expressions and Intensities,' in *Proc. of IEEE International Conf. on Systems,Man,and Cybernetics*, pp. 3123-3128,Oct. 2012.Article (CrossRef Link)

[2] L.H.ZHAO, C.K.YANG, F.PAN et al., "Face Recognition Based on Gabor with 2DPCA and PCA,' in *Proc. of 24th Chines Control and Decision Conference*, pp.2632-2635, 2012. Article (CrossRef Link)

[3] C.F.Beckmann, S.M. Smith., "Probabilistic independent component analysis for functional magnetic resonance imaging," *IEEE Trans on Medical. Imaging Magazine*, vol. 23, no. 2, pp. 137-152, Feb. 2004. Article (CrossRef Link)

[4] S.Nikitidis, A.Tefas et.al., "Maximum Margin Projection Subspace Learning for Visual Data Analysis," *IEEE Trans on Image ProcessingMagazine*,vol.23, no.10,pp.4413-4425, Oct. 2014. Article (CrossRef Link)

[5] Q.X. Gao, J.J. Liu et al., "Enhanced fisher discriminant criterion for image recognition," *Pattern Recognition Magazine*, vol.45, no.10, pp. 3717-3724, Oct. 2012. Article (CrossRef Link)

[6] G.Lu,Z.Lin et al., "Face recognition using discriminant locality preserving projections based on maximum margin criterion,' '*PatternRecognitionMagazine*, vol.43, no.10, pp.3572-3579, Oct. 2010. Article (CrossRef Link)

[7] J.X.Ruan, "Study on Key Technology for Multi-Pose Face Detection and Facial Expression Recognition," *South China University of Technology-Thesis*, pp.64-83, Jun. 2010.

[8] M.Kyperountas, A.Tefas, I. Pitas., "Salient feature and reliable classifier selection for facial expression classification," *PatternRecognitionMagazine*, vol. 43, no. 3, pp.972-986, Mar. 2010. Article (CrossRef Link)

[9] T. Ojala, M. Pietikainen, D. Harwood., "A comparative study of texture measures with classification based on feature distributions," *PatternRecognitionMagazine*, vol. 29, no. 1, pp. 51-59, Jan. 1996. Article (CrossRef Link)

[10] D. Huang,C.F.Shan,M. Ardabilian,et al., "Local Binary Patterns and Its Application to Facial Image Analysis:A Survey," IEEE *Transactions on Systems Man and Cybernetics-Part C:Applications and ReviewMagazine*, vol. 41, no. 6, pp. 765-781, Nov. 2011. Article (CrossRef Link)

[11] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Transactions on Pattern Analysis and Machine IntelligenceMagazine*, vol. 29, no. 6, pp. 915-928, Jun. 2007. Article (CrossRef Link)

[12] J.Caoand C. Tong, "Facial expression recognition based on LBP-EHMM," in *Proc. of Congress on Image and Signal Processing*, vol. 2, pp. 371-375, May. 2008. Article (CrossRef Link)

[13] T. Jabid, M.H.Kabir, O. Chae, "Robust Facial Expression Recognition Based on Local Directional Pattern," *ETRIJournalMagazine*, vol. 32, no. 5, pp. 784-794, Oct. 2010. Article (CrossRef Link)

[14] N. Dalal,B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 886-893, June. 2005. Article (CrossRef Link)

[15] F.Y. Song, X.Y.Tan,X.Liu,S.C. Chen, "Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients," *Pattern Recognition Magazine*, vol. 47, no. 9, pp. 2825-2838, Sep. 2014. Article (CrossRef Link)

[16] R.Minetto, N.Thome, M. Cord et al., "T-HOG: An effective gradient-based descriptor for single line text regions," *Pattern Recognition Magazine*, vol. 46, no. 3, pp. 1078-1090, Mar. 2013. Article (CrossRef Link)

[17] G.Gilanie, M.Attique, S.Naweed, "Object extraction from T2 weighted brain MR image using histogram based gradient calculation," *Pattern Recognition Letters Magazine*, vol. 34, no. 12, pp. 1356-1363, Sep. 2013. Article (CrossRef Link)

[18] L.P.LI., "A Research on Image De-noising Methods," Yangtze University-Thesis, Jun. 2012.

[19] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence Magazine*, vol. 12, no. 7, pp. 629-639,Jul. 1990. Article (CrossRef Link)

[20] Z.YAN, H.M. GU,C.G.CAI, "Enhancement Processing on Seismic Image based on anisotropic Diffusion," *Chinese Journal of Geophysics Magazine*, vol. 48, no. 3, pp. 390-394,Oct. 2013. Article (CrossRef Link)

[21] J.L. CHEN, X.J. LIU, “On Digital Image Watermarking Based on Anisotropic Diffusion," *Computer Applications and Software Magazine*, vol. 30, no. 7, pp. 151-153, Jun. 2013. Article (CrossRef Link)

[22] L.J. FU, Y. YAO, Z.L.FU, "Filtering method for medical images based on median filtering and anisotropic diffusion," *Journal of Computer ApplicationsMagazine*, vol. 34, no. 1, pp. 145-148, May. 2014. Article (CrossRef Link)

[23] O.Déniz, G.Bueno, J.Salido, etal, “Face recognition using Histograms of Oriented Gradients,” *Pattern Recognition Letters Magazine*, vol. 32, no. 12, pp. 1598-1603, Sep. 2011. Article (CrossRef Link)

[24] B.C.Zhang, Y.S.Gao, S.Q.Zhao, et al., “Local Derivative Pattern Versus Local Binary Pattern : Face Recognition with High-Order Local Pattern Descriptor,” *IEEE Transactions on Image Processing Magazine*, vol. 19, no. 2, pp. 533-544, Feb. 2010. Article (CrossRef Link)

[25] Y.B. Zheng, X.S. Huang, S.J.Feng, "An Image Matching Algorithm Based on Combination of SIFT and the Rotation Invariant LBP," *Journal of Computer 2 Aided Design & Computer Graphics Magazine*, vol. 22, no. 2, pp. 286-292, Sep. 2010. Article (CrossRef Link)

[26] X. Tan and B. Triggs, "Enhanced local texture feature sets for facerecognition under difficult lighting conditions," in *Proc. of Anal. Model Faces Gestures*, pp.168–182, Oct. 2007. Article (CrossRef Link)

[27] Z. Xiang,H.L. Tan,Z.M.Ma, "Performance Comparison of Improved HOG,Gabor and LBP," *Journal of Computer-Aided Design & Computer Graphics Magazine*, vol. 24, no. 6, pp. 787-792, Feb. 2012. Article (CrossRef Link)

[28] A.R. Rivera, J.R. Castillo, O.Chae, "Local Directional Number Pattern for Face Analysis : Face and Expression Recognition," *IEEE Transactions on Image processing Magazine*, vol. 22, no. 5, pp. 1740-1752, May. 2013. Article (CrossRef Link)

[29] F.Ahmed, "Gradient directional pattern: a robust feature descriptor for facial expression recognition," *Electronics Letters Magazine*, vol. 48, no. 19, pp. 1203-1204, Sep. 2012. Article (CrossRef Link)

[30] C.X. Ding, J.Y. Choi, D.C. Tao, et al., "Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence Magazine*, vol. 38, no. 3, pp. 518-531, Mar. 2016. Article (CrossRef Link)

[31] C.X. Ding and D.C. Tao., 'Robust Face Recognition via Multimodal Deep Face Representation," *IEEE Transactions on Multimedia Magazine*, vol. 17, no. 2, pp.2049-2058, Nov. 2015. Article (CrossRef Link)

[32] Wang Q, Ma L, Gao Q, et al., "Adaptive maximum margin analysis for image recognition," *Pattern Recognition*, vol. 61, pp.339-347, 2017. Article (CrossRef Link)

[33] Zhang L, Wang L, Lin W., "Semi-supervised biased maximum margin analysis for interactive image retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2294-2308, 2012. Article (CrossRef Link)

[34] Gao Q, Liu J, Cui K, et al., "Stable locality sensitive discriminant analysis for image recognition," *Neural Networks*, vol. 54, pp. 49-56, 2014. Article (CrossRef Link)

[35] Wong W K, Zhao H T., "Supervised optimal locality preserving projection," *Pattern Recognition*, vol. 45, no. 1, pp. 186-197, 2012. Article (CrossRef Link)

[36] Chao W L, Ding J J, Liu J Z., "Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection," *Signal Processing*, vol. 117, pp. 1-10, Mar. 2015. Article (CrossRef Link)

[37] Wolf L, Hassner T, Taigman Y., "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 33, no. 10, pp. 1978-1990, 2011. Article (CrossRef Link)

[38] Wolf L, Hassner T, Taigman Y., "Descriptor based methods in the wild," in *Proc. of Workshop on faces in 'real-life' images: Detection, alignment, and recognition.* 2008.

[39] Vu N S, Caplier A., "Enhanced patterns of oriented edge magnitudes for face recognition and image matching," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1352-1365, 2012. Article (CrossRef Link)

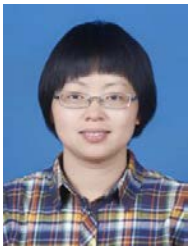
[40] Lei Z, Pietikäinen M, Li S Z., "Learning discriminant face descriptor,"*IEEE Transactions on pattern analysis and machine intelligence*, vol. 36, no. 2, pp. 289-302, 2014. Article (CrossRef Link)

[41] Shan C, Gong S, McOwan P W., "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009. Article (CrossRef Link)

[42] ZHAO De, HE Chuan-Jiang, CHEN Qiang, "Anisotropic Diffusion Model Combined with Local Entropy," *PR & AI*, vol. 25, no. 4, pp. 642-647, Aug. 2012. Article (CrossRef Link)

[43] Schroder J B., "Smoothed aggregation solvers for anisotropic diffusion," *Numerical Linear Algebra With Applications*, vol. 19, no. 2, pp. 292-312, 2012. Article (CrossRef Link)

[44] Wang H, Wang Y, Ren W, et al., "Image denoising using anisotropic second and fourth order diffusions based on gradient vector convolution," *Computer Science and Information Systems*, vol. 9, no. 4, pp. 1493-1511, Mar. 2012. Article (CrossRef Link)



Ying Tong received her B.Sc. from PLA University of Science and Technology in 2001 and received her M.Sc. from Nanjing University of Posts and Telecommunications in 2006. Now she is an associate professor in Nanjing Institute of Technology. Her research interests include Signal processing, Pattern recognition and Sparse coding.

Yuehong Shen received his PhD degree in 1999 in Nanjing University of Science and Technology. Now he is a professor and doctoral supervisor of PLA University of Science and Technology. His research interests include wireless MIMO communication, digital signal processing, statistical signal analysis, wireless statistic division multiplexing.

Bin Gao received his B.Sc. and M.Sc. both from PLA University of Science and Technology in 2007 and 2010 respectively. Now he is an education/school administrator in PLA University of Science and Technology. His research interests include compressive sensing, optimization methods and machine learning.

Fenggang Sun received his B.Sc. and M.Sc. both from Shandong University in 2006 and 2009 respectively. Now he is a lecturer in Shandong Agricultural University. His research interests include compressive sensing, array signal processing and cognitive radio networks.

Rui Chen received the B.E degree, the M.E. degree from Southeast University, Nanjing, China, in 1991 and 1996, respectively. She received PhD degree from Nanjing University of Post and Telecommunications in 2013. Her major research interests include distributed video coding and wireless multimedia communication.

Yefeng Xu received his B.Sc. from PLA University of Science and Technology in 1990 and received his M.Sc. from University of Electronic Science and Technology of China in 2013. Now he is a lecturer in PLA University of Science and Technology. His research interests include 5G communication,and sparse signal representation.