

Multi-scale Local Difference Directional Number Pattern for Group-housed Pigs Recognition

Weijia Huang^{1,2}, Weixing Zhu^{1*}, Zhengyan Zhang², Yizheng Guo³

¹ School of Electrical and Information Engineering, Jiangsu University
Zhenjiang, Jiangsu, China

[e-mail: huangweijia@just.edu.cn, zwxbest@126.com]

² School of Electronics and Information, Jiangsu University of Science and Technology
Zhenjiang, Jiangsu, China

[e-mail: zhangzhengyan@just.edu.cn]

³ Nanjing Normal University Taizhou College
Taizhou, Jiangsu, China

[e-mail: guoyizheng0523@163.com]

*Corresponding author: Weixing Zhu

*Received November 23, 2020; revised January 10, 2021; accepted August 13, 2021;
published September 30, 2021*

Abstract

In this paper, a multi-scale local difference directional number (MLDDN) pattern is proposed for pig identification. Firstly, the color images of individual pig are converted into grey images by the most significant bits (MSB) quantization, which makes the grey values have better discrimination. Then, Gabor amplitude and phase responses on different scales are obtained by convoluting the grey images with Gabor masks. Next, by calculating the main difference of local edge directions instead of traditionally edge information, the directional numbers of Gabor amplitude and phase responses are encoded. Finally, the block histograms of the encoded images are concatenated on each scale, and the maximum pooling is adopted on different scales to avoid the high feature dimension. Experimental results on two pigsties show that MLDDN impressively outperforms the other widely used local descriptors.

Keywords: Multi-scale local difference directional number, Gabor filter, Pig identification.

This work was part of a project funded by “The National Natural Science Foundation of China” (Grant No.31872399), “The Doctoral Program of the Ministry of Education of China” (Grant No. 2010322711007), “The Priority Academic Program Development of Jiangsu Higher Education Institutions,” “The Graduate Student Scientific Research Innovation Projects of Jiangsu Ordinary University” (Grant No. CXLX13_664), and PhD Research Project of Jiangsu University of Science and Technology(Grant No. 1032931604).

1. Introduction

Identification of individual animals is important for welfare management in agricultural animals. Although the radio frequency identification (RFID) technology has been widely used, it is still in an invasive way and costly for current practical applications. Jover et al., “[1]” marked on pig’s back and the location of pigs was detected based on image processing technique. Ahrendt et al., “[2]” marked the seed points on pig’s body and traced the location by estimating the spatial location of pigs in each new frame. Kashiha et al., “[3]” identified pigs based on colour and shape markers on pig’s back. These methods overcome the shortcomings of RFID. However, they all depend on printing pattern, number and other marks on the body of pigs. The marks are difficult to preserve for a long time due to the unhygienic environment of pigsty and the growth of pigs, which seriously affects the recognition performance.

To utilize the inherent physiological or behavioural characteristics of animals, more stable biological characteristics were extracted for livestock recognition. Relevant researches of livestock mainly focused on muzzle print image, retinal vascular patterns and iris patterns [4]. Muzzle print reflects the skin lines on the noses of cattle which are different from each other. Kumar et al., “[5]” proposed a method for cattle recognition with group sparse representation, based on their muzzle points and face features. Sian et al., “[6]” extracted features from muzzle print image by the fusion of the improved weber local descriptor and local binary pattern. Lu et al., “[7]” extracted the local and global features of cow iris by 2D complex wavelet transform for cow recognition. Recognition based on the inherent characteristics of biology has the advantages of convenience, low cost and difficulty in counterfeiting. However, these methods are based on the fact that the muzzle images, retinal vessels and iris patterns are different from each other for livestock. Therefore, Livestock usually need to go to a specific position for collecting the images of specific areas of the body. Manual work is usually required owing to the limitations and heavy workloads of these methods.

Kim et al., “[8]” designed an image-processing system to recognize Holstein cows by their body patterns. A charge-coupled device (CCD) was installed near a passageway to capture the side image of each cow. The neural network was applied to Holstein cow recognition. Corkery et al., “[9]” captured face image and identified sheep by independent component analysis. Santosh et al., “[10]” built a facial images database of cattle and used face recognition algorithms based on computer vision for cattle recognition. Zhao and He [11] collected the video when cows walked through a fixed narrow passageway. The side view images of the cows were captured and the convolution neural network (CNN) was used for cow identification. Hansen et al., “[12]” collected the face of pigs and adopted techniques from the human face recognition, such as Fisherfaces, VGG-Face pre-trained and CNN. The internal characteristics of livestock are extracted from side and face image. Compared with muzzle print, retinal vascular and iris image, this method is more convenient for image acquisition. However, livestock still need to go to a specific position or maintain a specific posture. It is difficult to apply these methods to behavior recognition applications. Therefore, although the appearances of pigs share strong visual similarities, it is still important to extract the internal characteristics to distinguish each pig, especially in the application of behaviour recognition. It is of great significance to find the appropriate characteristics of pig body surface in the condition of free movement rather than certain position or posture.

In our previous work, Guo et al., “[13]” segmented individual pigs in a region of interest and a histogram, colour moments, a grey level co-occurrence matrix and shape features were extracted for pig identification by support vector machine (SVM). In real farming

environments, imbalanced local illumination usually affects the feature extraction of colour and contour. Therefore, the insensitivity of texture features to color and illumination changes are discussed. A method based on the combination of Gabor and local binary pattern (LBP) was proposed, and the possibility of individual recognition based on the texture features of pig body surface was explored [14].

In this paper, a novel local feature descriptor, multi-scale local difference directional number (MLDDN) pattern, was proposed for pig identification by extracting the texture features of pig body surface. More discriminative features were obtained. Local adjacent edge directional information of pig images was encoded to improve the robustness of coding. Moreover, two pigsties were taken as samples to verify the effectiveness of the proposed method. The proposed method based on pigsty is universal for identification of group-housed pigs. Pigs were monitored in the pigsty that they do not need to go to a specific position or maintain a specified posture for recognition. Therefore, it's more suitable for the applications of pig abnormal behaviour recognition.

The main contributions of this paper are as follows:

(1) For better use of the colour information of individual pig, the most significant bits (MSB) quantization method is applied for colour image quantization. MSB method could produce values which better discriminate the colours [15]. In this way, pig images with different colours, or patches of different colours on the bodies, are quantized to gray value with larger variation. It is conducive to distinguishing pigs of different colours and produces compact features.

(2) In order to extract more useful information and enhance the discriminative of feature descriptor, the filtering template of single scale used in LDN method is improved. Gabor masks with different scales and directions are used to extract multi-scale structural information from pig images. Moreover, a real pigsty always under complex light conditions and pigs may move to different positions in the pigsty. Therefore, Gabor phase response which is insensitive to illumination conditions [16,17] is also used for feature extraction.

(3) In the process of local image coding, the changes of pixels in neighborhood are usually used to describe the local structure information of the image. The numbers of prominent direction are encoded by calculating the value and symbol of edge response. When there is a small change caused by noise or illumination variance, the prominent direction of local neighborhood will change, which will lead to the change of LDN code. However, noise and illumination variation are inevitable in the real pig farm environment. In order to enhance the robustness of the coding, the differences between the adjacent direction of filtering responses are encoded, which are more stable under the influence of noise and illumination variance.

(4) Although multi-scale local information provides more features for pig identification, the feature dimension is also increased. In order to tackle this problem, maximum pooling is conducted on different scales. By this way, more representative features are encoded while the feature dimension of multi-scale is still the same as a one scale. Therefore, the increase of scale does not affect the feature dimension and more discriminative features are extracted.

2. Materials and Methods

2.1 Image Acquisition and Pre-processing

Experimental videos were captured from pig farm of Zhenjiang Xima Development Company, based at Jiangsu University. There were several pigsties in the farm. Each pigsty covers an area of 4 square meters (2 meters long and 2 meters wide) with 6 to 10 pigs in it. By rebuilding

the pigsty, the camera of FL3-U3-88S2C-C from Point Grey Research Inc. (Riverside Way V6w 1k7 Richmond, BC, Canada) was installed 3 meters above the experimental pigsty, as shown in **Fig. 1** (a). Several videos of group-housed pigs were captured on a sunny afternoon in June 2015 and May 2017 with 1760×1840 pixels, as illustrated in **Fig. 1** (b) and (c). Each video was recorded about 3 minutes and divided into image frames. The images of individual pigs were extract from frames by an adaptive partitioning and multilevel thresholding segmentation method [18] and normalized to the same size based on the centroid and labelled, as shown in **Fig. 1** (d) and (e).

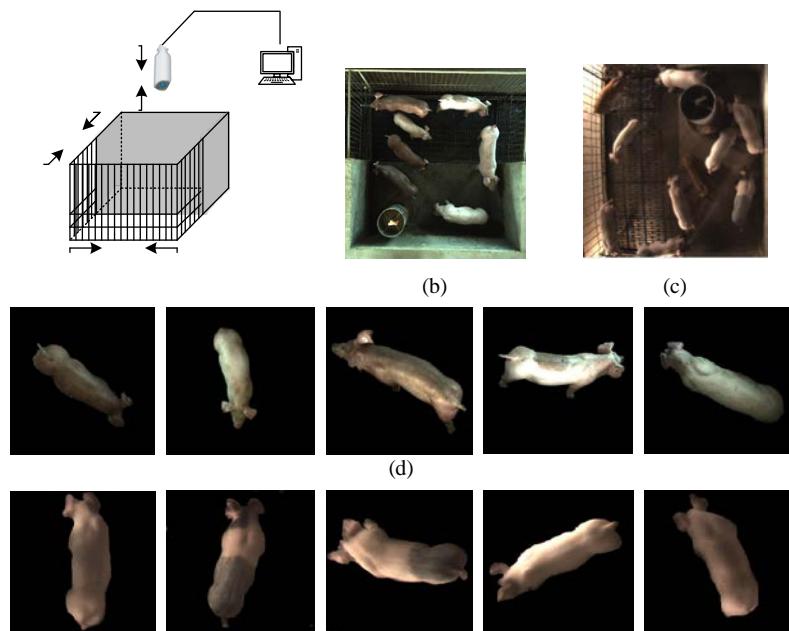


Fig. 1. Video capture system and pig image samples: (a) video capture system of pigsties in the farm, (b) image frames of pigsty No.1, (c) image frames of pigsty No.2, (d) samples of pigsty No.1, (e) samples of pigsty No.2.

Two pigsties were used for the experiments, named as No.1 and No.2. There were 7 pigs in No.1 and 10 pigs in No.2. In the early research, 7 pigs were selected and mixed from other pigsties. The pigs were about 60 days old with an average weight of 24 kilograms, and their colour, body pattern and size were obviously different from each other. 350 individual images of the 7 pigs were taken as samples. Later, in order to further verify the effectiveness of the proposed method, an ordinary pigsty with 10 pigs was used, in which pigs were not specially selected. The 10 pigs were more similar to each other. The pigs were about 45 days old and the average weight was 19 kilograms. 500 individual images of the 10 pigs were taken as samples. All the images were normalized to 100×100 pixels and labelled.

2.2 Relevant Work and Analysis

2.2.1 Local directional number (LDN)

Local pattern has attracted much attention in many applications, such as face analysis, texture classification and scene classification. The edge information with eight directions in

the neighbourhood is encoded in Local Binary Pattern (LBP) [19] and Local Direction Pattern (LDiP) [20]. The transition of intensity change is encoded in Local transitional pattern (LTrP) [21]. Binary Pattern of Phase Congruency (BPPC) applies wavelet transform to the logarithmic Gabor features [22]. However, the dimension is relatively high. The local descriptors are based on the fact that image sub-blocks contain a large number of information. If more structural information is extracted from sub-blocks, more discriminative features can be obtained.

Rivera et al., “[23] proposed a local directional number (LDN) pattern for face analysis, which achieves good performance for face and facial expression classification [24]. Local direction information of texture is encoded in LDN. The prominent direction information is encoded with the aid of the Kirsch compass masks, as shown in **Fig. 2**.

$$\begin{array}{cccc}
 \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} & \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\
 m_0 & m_1 & m_2 & m_3 \\
 \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \\
 m_4 & m_5 & m_6 & m_7
 \end{array}$$

Fig. 2. Kirsch compass masks.

By convoluting images with Kirsch compass masks, edge responses in eight directions of the image can be obtained by:

$$Q_i = I * M_i \quad (1)$$

where I represents the original image, M_i is the Kirsch mask in the i th direction and $*$ denotes convolution operation. Q_i is the edge response of the original image in the i th direction.

The numbers of prominent direction by calculating the value and symbol of edge responses are encoded by a compact way in LDN method. The maximum positive and minimum negative values of the edge response are encoded by:

$$\text{LDN}(x, y) = 8 \times N_1(x, y) + N_2(x, y) \quad (2)$$

where (x, y) is the center pixel in the neighborhood, $N_1(x, y)$ is the directional number of the maximum positive value of edge responses, and $N_2(x, y)$ is the directional number of the minimum negative value of edge responses, which can be calculated by:

$$N_1(x, y) = \arg \max_i \{Q_i \mid 0 \leq i \leq 7\} \quad (3)$$

$$N_2(x, y) = \arg \min_j \{Q_j \mid 0 \leq j \leq 7\} \quad (4)$$

2.2.2 Gabor Wavelet Transform

Gabor wavelet transform is widely used for texture representation. It has different scales and different directions, which make it sensitive to the edge.

Two-dimensional Gabor function [25, 26], $\Psi(x, y)$, is defined as:

$$\Psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\pi\left[\frac{(x-x_0)^2}{\sigma_x^2} + \frac{(y-y_0)^2}{\sigma_y^2}\right]\right\} \exp\{j[u_0x + v_0y]\} \quad (5)$$

where (x_0, y_0) is the center in the space domain and (u_0, v_0) is the optimal spatial frequency in the frequency domain. σ_x and σ_y are standard deviations along X and Y axes. Gabor filtering can be expressed as the convolution operation [27]:

$$F_{u,v}(x, y) = I(x, y) * \Psi_{u,v}(x, y) = A_{u,v}(x, y) e^{j\theta_{u,v}(x, y)} \quad (6)$$

where $F_{u,v}(x, y)$ is the filtered image, $I(x, y)$ represents the original image, and $*$ is convolution operation. $A_{u,v}(x, y)$ and $\theta_{u,v}(x, y)$ represent the Gabor amplitude and phase responses in the u th direction and v th scale, respectively.

2.3 The proposed method

The flow chart of pig identification based on MLDDN is shown in **Fig. 3**. Firstly, RGB image of individual pig is quantized into grey image by MSB quantization method. Secondly, Gabor amplitude and phase responses are obtained by convoluting the grey image with Gabor masks. The main difference of local edge direction is calculated and the directional numbers of Gabor amplitude and phase responses are encoded on each scale. Thirdly, the encoded image is divided into several sub-blocks and the histograms of Gabor amplitude and phase responses are calculated, respectively. Maximum pooling is conducted on different scales to reduce the feature dimension. Finally, the histograms of Gabor amplitude and phase responses are cascaded and the SVM classification is conducted for training and recognition.

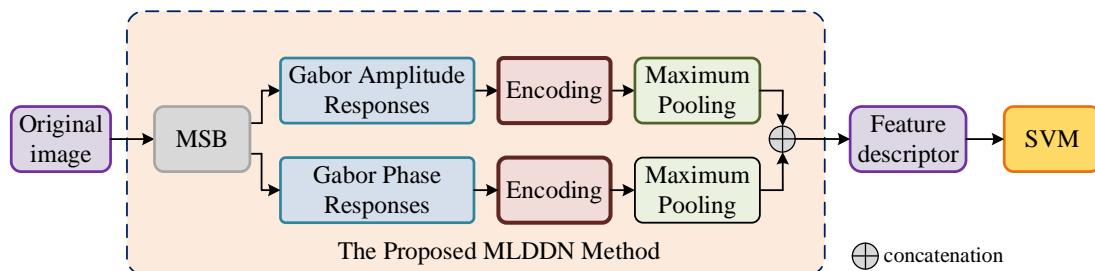


Fig. 3. Flow chart of the proposed method.

2.3.1 Image quantization

To better describe colour information, features can be extracted from each colour channel of the RGB colour image. However, the feature dimension will be three times as large as that of the grey image. The common solution is to reduce the number of colours into one channel. Luminance is the most popular method of quantization on the basis of human brightness perception [15]. It computes a weighted combination of the RGB channels:

$$Q_{Luminance} = 0.587G + 0.299R + 0.114B \quad (7)$$

Recent research has shown that quantization based on significant bits has better effect on feature extraction [15]. The main idea is to combine the pixel values based on the most important bits in RGB channels. By this way, more discriminative pixel values can be obtained. Each pixel in the quantized image can be represented by an 8-bit binary number. The bit importance increases from the 0th bit (lowest) to the 7th bit (highest). The binary template is defined by:

$$GM = \sum_{i=(8-Ng)}^7 G_i \cdot 2^i \quad (8)$$

$$RM = \sum_{i=(8-Nr)}^7 R_i \cdot 2^{i-Ng} \quad (9)$$

$$BM = \sum_{i=(8-Nb)}^7 B_i \cdot 2^{i-(Ng+Nr)} \quad (10)$$

where G_i , R_i and B_i represent the i th bit code of the G, R, B colour channel, N_g , N_r and N_b are the amount of bits used from channels G, R and B, respectively. Thus, the image based on MSB method can be defined by:

$$P_{MSB} = RM + GM + BM \quad (11)$$

The results of quantization based on Luminance and MSB methods are shown in **Fig. 4**. It can be obviously seen that the image intensity change based on MSB method is much more remarkable. MSB quantization method doesn't follow human perception like Luminance method, but records the most significant bits of each colour channel. Therefore, the MSB method can produce better discriminative grey values which are helpful for feature classification [15]. Pigs of different colours or patches of different colours on pig body are quantized to gray values with larger variation, which is helpful to distinguish pigs of different colours and produce compact features.

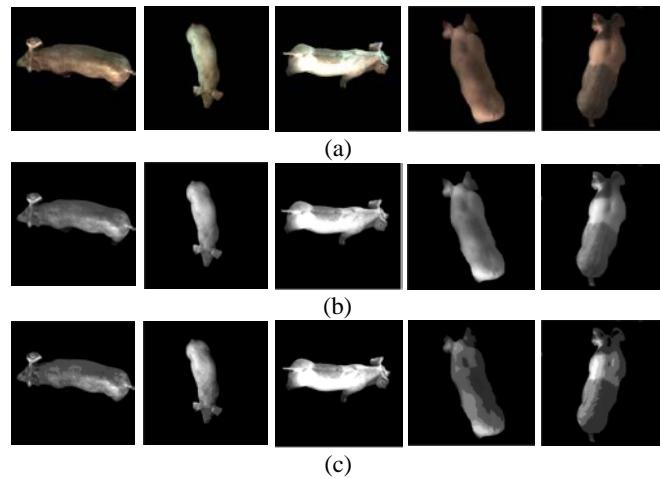


Fig. 4. Results of Luminance and MSB quantization: (a) original RGB images, (b) quantization results of Luminance method, (c) results of MSB method.

2.3.2 Gabor Filtering

According to human observation experience, information of different structures could be observed on different scales. Specially, some details in the image could only be seen on a certain scale. Therefore, if multi-scale information could be described efficiently, more useful features can be extracted. For this purpose, Gabor wavelets with multi-scale are used as compass masks in the proposed method. The grey images are convolved with Gabor wavelets and the responses of Gabor amplitude and phase in eight directions are calculated by Eq (6).

The Gabor responses in one direction with multi scales are shown in **Fig. 5**. **Fig. 5** (a) is a grey image. **Fig. 5** (b) are images of Gabor amplitude response and Gabor phase response. Obviously, different grey structures are shown on different scales in **Fig. 5** (b). More local information can be found from images on the left and more global information can be found on the right. That's because with the increase of Gabor filter window (from left to right), the filtered results gradually change from local features to global features [14]. By this way, more abundant texture features caused by shape, colour and hairs on pig's body surface can be extracted.

It is natural that different characteristics can be observed on different scales by human vision. Compared with feature extraction on single scale, it is easier to get more discriminative information on multi scales, which helps to explore important texture features on pig body surface.

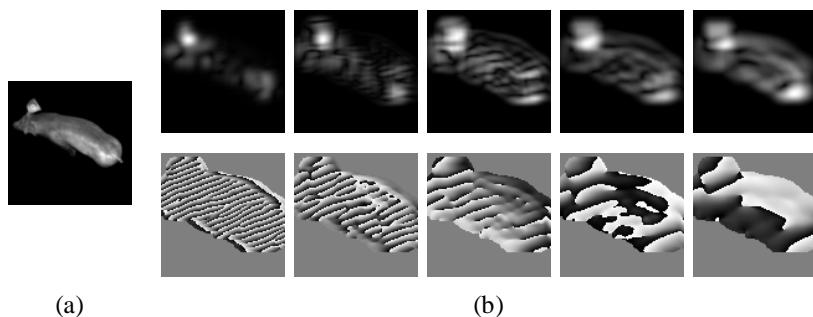


Fig. 5. Images of Gabor amplitude and phase response: (a) grey image based on MSB quantization, (b) images of Gabor amplitude response in the first row and Gabor phase response in the second row with different scales in the second direction.

2.3.3 Encoding

In LDN method, the numbers of prominent direction are encoded by calculating the value and symbol of edge response. However, the prominent direction information of local neighborhood will be changed by the influence of noise and illumination variation. Since the noise and illumination variation are inevitable in a real pig farm environment.

Image edges are more stable to noise and illumination variation, which can greatly reduce the data to be processed while preserve the shape of the object in the image. Inspired by this, the difference of filtered responses between adjacent direction is encoded, as defined by:

$$D_{u,v}(x, y) = [A_{u,v}(x, y) - A_{u+1,v}(x, y)], \quad u = 0, 1, \dots, 6 \quad (12)$$

$$D_{7,v}(x, y) = [A_{7,v}(x, y) - A_{0,v}(x, y)] \quad (13)$$

where $A_{u,v}(x, y)$ represents Gabor amplitude response in the u th direction and the v th scale, $D_{u,v}(x, y)$ denotes the corresponding difference. Gabor wavelets with five scales ($v = 0, 1, \dots, 4$) and eight directions ($u = 0, 1, \dots, 7$) were used in the experiments.

The main difference of Gabor amplitude response is encoded as:

$$\text{MLDDN}_A(x, y) = 8 * Q_{1,v}(x, y) + Q_{2,v}(x, y) \quad (14)$$

where $Q_{1,v}(x, y)$ and $Q_{2,v}(x, y)$ are directional numbers corresponding to the maximum and minimum difference of Gabor amplitude response on the v th scale, which are defined by:

$$Q_{1,v}(x, y) = \arg \max_u \{D_{u,v} \mid 0 \leq u \leq 7\} \quad (15)$$

$$Q_{2,v}(x, y) = \arg \min_u \{D_{u,v} \mid 0 \leq u \leq 7\} \quad (16)$$

Gabor amplitude response is commonly used due to its stability and slower changes. However, pigs may move to different positions in a real pigsty with complex illumination. Therefore, Gabor phase response is also calculated in the proposed method because its insensitivity to illumination conditions [16,17]. The main difference of Gabor phase response on the v th scale is encoded by the same way, which is described as $\text{MLDDN}_P(x, y)$.

Take one scale as an example, the encoding process of Gabor amplitude responses are shown in **Fig. 6**. Firstly, the responses in eight directions were calculated, as shown in **Fig. 6** (a). It can be seen that the response of 0.61 corresponds to directional number 0 and 0.43 corresponds to directional number 1, and so on. Then, difference between adjacent directions is calculated in a counterclockwise, as shown in **Fig. 6** (b). It can be seen that the maximum difference is 0.41 with directional number 5 and the minimum difference is -0.37 with directional number 2. Finally, the directional numbers of the maximum and minimum are encoded, as shown in **Fig. 6** (c).

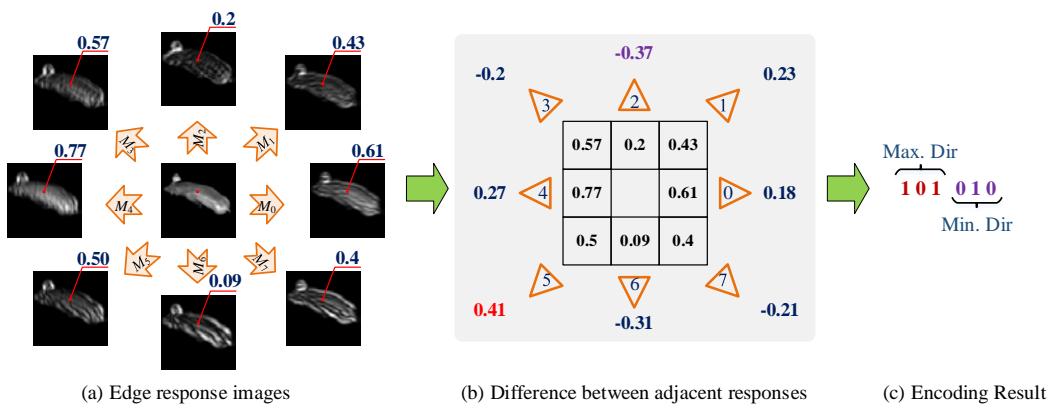


Fig. 6. Example of the encoding process of Gabor amplitude responses on a certain scale.

Fig. 7 shows the comparison between MLDDN and other similar local descriptors, such as LDIP and LDN. The edge responses in eight directions are encoded. Taking a 3×3 neighborhood as an example, and the encoding results of LDIP, LDN and MLDDN are listed below. **Fig. 7** (a) is the original condition and **Fig. 7** (b) and (c) are cases of noise or illumination variance that some pixels changes in the neighborhood.

It is clear that the encoding results of LDIP and LDN have changed in different cases. Conversely, the result of MLDDN remains unchanged in all three cases. One of the reasons is that the MLDDN method encodes the difference between adjacent edge responses. When there is a small change caused by noise or illumination variance, the difference between adjacent directions which describes the main characteristics of local regions remains the same.

0.59	0.40	0.43	0.59	0.60	0.43	0.79	0.80	0.63
0.77		0.61	0.77		0.61	0.77		0.61
0.50	0.09	0.40	0.50	0.09	0.40	0.50	0.09	0.40
(a)			(b)			(c)		
LDP=10011000			LDP=10101000			LDP = 00111000		
LDN=100110			LDN=100110			LDN = 010110		
MLDDN=101110			MLDDN=101110			MLDDN = 101110		

Fig. 7. Comparison between MLDDN and other local descriptors in different cases.
 (a) shows a neighborhood and its encoding. (b) shows the case that the value in one direction changes probably due to noise. (c) shows the case that there are more changes in three directions.

2.3.4 Feature Descriptor

After encoding on all the scales, the feature vector is formed by histogram statistics of the encoded image. However, the histogram only shows the frequency of code value in the image. Many location-related information such as edge, point and corner which could distinguish different pigs will be lost. In order to integrate the location-related information, the coded images of Gabor amplitude and phase responses are divided into N sub-blocks $\{R_1, R_2, \dots, R_N\}$. The histograms of each sub-block are calculated and cascaded as:

$$ha_v^n = \prod_{i=1}^{56} ha_{v,i}^n = \{ha_{v,1}^n, ha_{v,2}^n, \dots, ha_{v,56}^n\} \quad (17)$$

$$hp_v^n = \prod_{i=1}^{56} hp_{v,i}^n = \{hp_{v,1}^n, hp_{v,2}^n, \dots, hp_{v,56}^n\} \quad (18)$$

Here,

$$ha_{v,i} = \sum_{(x,y) \in R_n} \delta(\text{MLDDN_}A_v(x,y), S_i), i=1,2,...,56 \quad (19)$$

$$hp_{v,i} = \sum_{(x,y) \in R_n} \delta(\text{MLDDN_}P_v(x,y), S_i), i=1,2,\dots,56 \quad (20)$$

$$\delta(m, g) = \begin{cases} 1, & m = g \\ 0, & otherwise \end{cases} \quad (21)$$

where ha_v^n and hp_v^n denote the histograms of the v th region on the n th scale of the encoded images of Gabor amplitude and phase responses, respectively. \prod denotes concatenation operation. S is a code value and (x, y) is the pixel in the n th region.

Then, maximum pooling is conducted on different scales for reducing feature dimension. The maximum of the scales in the same bins is selected, the histograms are defined by:

$$hist_{amp} = \prod_{n=1}^N (\max \sum_{v=1}^5 ha_v^n) = \{ \max \sum_{v=1}^5 ha_v^1, \max \sum_{v=1}^5 ha_v^2, \dots, \max \sum_{v=1}^5 ha_v^N \} \quad (22)$$

$$hist_{pha} = \prod_{n=1}^N (\max \sum_{v=1}^5 hp_v^n) = \{ \max \sum_{v=1}^5 hp_v^1, \max \sum_{v=1}^5 hp_v^2, \dots, \max \sum_{v=1}^5 hp_v^N \} \quad (23)$$

where $hist_{amp}$ and $hist_{pha}$ represent the histogram of Gabor amplitude and phase responses. Correspondingly, the histogram dimension of multi-scale is the same as one scale. The increase of scale does not affect the dimension of histogram.

Finally, the histograms of Gabor magnitude and phase responses are cascaded as the feature vector of a pig image by:

$$HIST = \prod (hist_{amp}, hist_{pha}) \quad (24)$$

If an image is divided into 4×4 sub-blocks, the feature dimension of MLDDN is $4 \times 4 \times 56 \times 2 = 1792$, where the number of histogram bins of each sub-block is 56.

Fig. 8 shows the histograms of Gabor amplitude and phase responses of different pigs. A square block based on the centroid of pig image was used to calculate the histograms. As can be seen, the histograms extracted from different pigs are different. Additionally, it can be seen that the difference of phase histograms is relatively more obvious. Therefore, Gabor phase response also provide useful and discriminative information for pig identification.

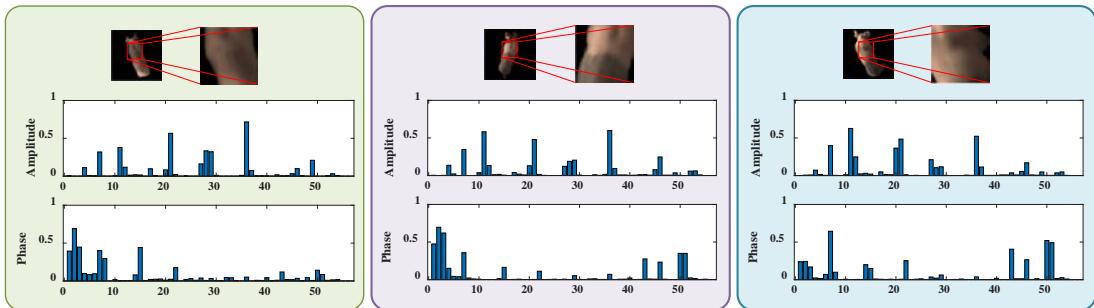


Fig. 8. The encoding results of Gabor amplitude and phase responses of different pigs.

3. Experimental Results and Discussions

3.1 Experiment setup

In the experiment, the individual pig images were randomly divided into five groups for five-fold cross validation. Four groups of them were set as training data and the rest were set as test data. Each group was set as the test data once by turn. The recognition rate was the average of the five records on the test data. Gabor masks with five scales and eight directions were used to extract features. All the images were divided into 4×4 sub-blocks for histogram statistics. SVM with linear, polynomial and RBF kernel functions was conducted for classification. LIBSVM [28] was used on the MATLAB R2018a. The computer processor was

the Intel® coreTM i5-8250U CPU@1.60GHz. The physical memory was 24GB. The GPU was NVIDIA GeForce MX150. The operating system was Microsoft Windows 10. The order of polynomial kernel function was 3 and the RBF kernel function penalty factor C was 100. Principal components analysis (PCA) [29] was also conducted to verify the effectiveness of the proposed method after feature dimension reduction. In the experiment, the data of pigsty No.2 was taken as an example to illustrate the method. Later, the proposed method was applied to pigsty No.1 to solve the problem of identification of group-housed pigs.

3.2 Results of pig identification for pigsty No.2

3.2.1 Results of different quantization methods.

Fig. 9 shows the results of pig identification with Luminance and MSB quantization methods. The blue column represents the result of Luminance quantization. The yellow and grey columns denote the results of MSB quantization method with 256 and 64 colours, respectively. It can be seen that the recognition rates of Luminance quantization were lower than that of MSB quantization by SVM with linear and RBF kernel function. That is because the MSB method produces values which better discriminate the colours [16,17]. Thus, it can better describe the colour information of different pigs. Moreover, the recognition rates of quantized version with 64 colours were higher than that of 256 colours, which further verifies the effectiveness of MSB quantization for classification.

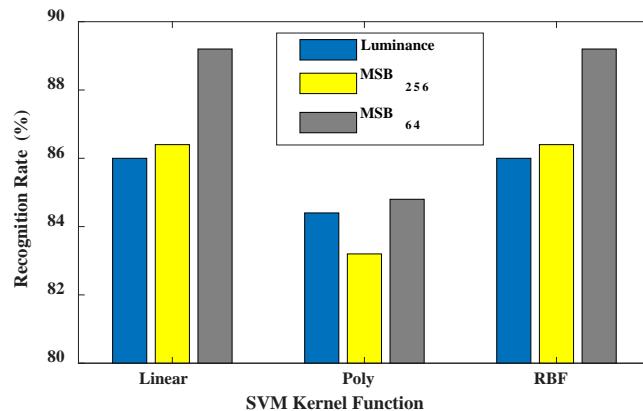


Fig. 9. Recognition rates of MLDDN with the Luminance and MSB quantization methods. Three groups of histograms correspond to the result by SVM classification with linear, polynomial (third order) and RBF ($C = 100$) kernel functions.

3.2.2 Results of different scales.

Table 1 lists the recognition rates and experimental times of MLDDN with different scales by SVM with a linear kernel function. It can be seen that with the increase of scales, the recognition rate was increased. That's because more representative characteristics can be extracted from multi scales. However, with the increase of scale, the computing time was also increased. Lee [30] showed that when Gabor transform is used to represent image nondestructively, eight discrete directions and five scales are needed in each discrete position. Therefore, five scales were used in the experiment.

Table 1. Recognition rates (%) and executed time (s) of MLDDN in different scales.

Number of Scales	2	3	4	5	6	7	8
Recognition rate	85.60	88.80	89.20	89.80	90.20	90.20	91.00
Time	22.34	31.53	41.25	52.67	61.14	74.15	80.11

3.2.3 Results of different local descriptors.

Table 2 shows the recognition rates of pig identification with other local patterns. It can be seen that MLDDN achieved higher recognition rate than other methods. The results of MLDDN were 89.8%, 87.0%, 89.8% by SVM with linear, polynomial and RBF kernel functions, respectively, which were 0.8%, 5% and 1.2% higher than LDN. The most likely reason is: (1) Kirsch compass masks used in LDN are one scale, while Gabor masks used in MLDDN describe the details of pig images on multiple scales. Thus, more important features can be extracted from pig's body surface for identification. (2) Not only the Gabor amplitude response, but also the Gabor phase response were used, which is insensitive to illumination conditions. It is helpful to reduce the impact of complex illumination variation of individual pig images in real pigsty. (3) Considering the impact of noise and illumination variation in the real pig farm environment, the relationship between edge responses of adjacent directions was encoded. So that the robustness of code was increased. (4) MSB quantization method was used to better describe the colour information of different pigs. **Fig. 10** shows the confusion matrix of pig identification for pigsty No.2 based on MLDDN by SVM with linear kernel function.

Table 2. Recognition rates (%) of pig identification with different methods for pigsty No.2.

Method	Linear	Polynomial	RBF(C=100)
LTrP	85.20	83.20	85.00
LDiP	86.00	82.80	86.40
BPPC	88.20	84.80	88.20
LDN	89.00	82.00	88.60
MLDDN	89.80	87.00	89.80

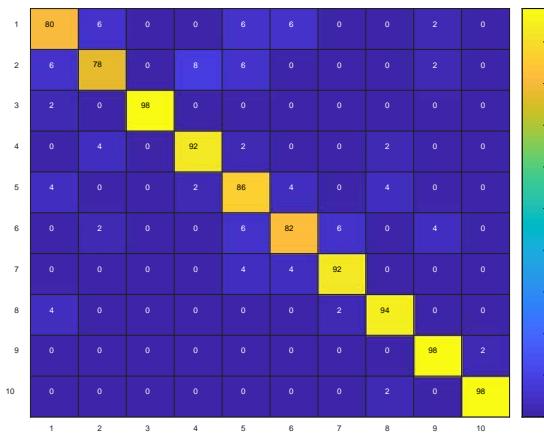
**Fig. 10.** Confusion matrix of pig identification based on MLDDN by SVM with linear kernel function.

Fig. 11 shows the recognition rates of pig identification after dimensionality reduction with PCA. The vertical axis represents the recognition rates with linear kernel SVM and the horizontal axis represents the PCA parameters varying from 0.85 to 0.99. It can be seen that

the MLDDN method also achieves higher recognition rates compared with other methods. MLDDN also keeps higher results while the results of LDN, LDiP and LTrP decrease rapidly with the dimension reduction. The results of BPPC are better than LDN, LDiP and LTrP after the dimension reduction. But the feature vector dimension of BPPC is much higher than other descriptors and MLDDN. Moreover, the results of Gabor amplitude response code $\text{MLDDN}_{\text{amp}}$ and Gabor phase response code $\text{MLDDN}_{\text{pha}}$ were also reported. It can be seen that $\text{MLDDN}_{\text{amp}}$ and $\text{MLDDN}_{\text{pha}}$ also achieves good performance and the fusion of them achieves better results, which further illustrates that Gabor phase information is complementary to Gabor amplitude information.

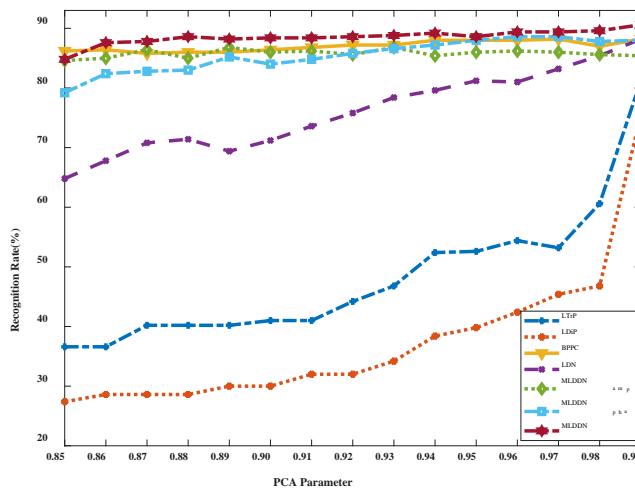


Fig. 11. Recognition rates (%) of different method with different PCA parameters for pigsty No.2.

3.3 Results of pig identification for pigsty No.1

In this paper, experiments of pig identification were also conducted for the pigsty No.1. The 7 pigs in the pigsty were selected and mixed from other pigsties in the early research. Their colour, body pattern and size were relatively different from each other, and the classification number of categories also decreased. Therefore, the experiment results are higher than the pigsty with pigsty No.2.

Table 4 lists the results of identification with different methods. It can be seen that MLDDN method achieves higher recognition rates by SVM with linear, polynomial and RBF kernel functions, which were 95.71%, 93.57% and 95.71%, respectively. **Fig. 12** shows the confusion matrix of pig identification for pigsty No.1 based on MLDDN by SVM with linear kernel function.

Table 4. Recognition rates (%) of pig identification with different methods for pigsty No.1.

Method	Linear	Polynomial	RBF(C=100)
LTrP	87.14	85.71	87.14
LDiP	87.14	85.71	87.14
BPPC	88.57	85.00	89.29
LDN	91.43	86.43	92.14
MLDDN	95.71	93.57	95.71

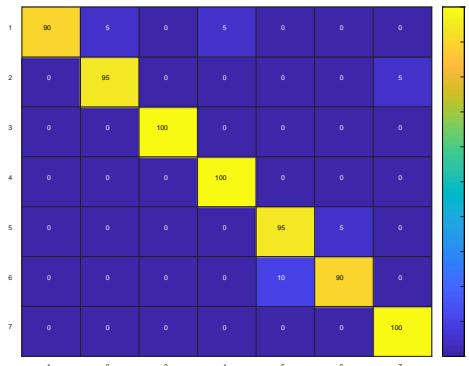


Fig. 12. Confusion matrix of pig identification based on MLDDN by SVM with linear kernel function.

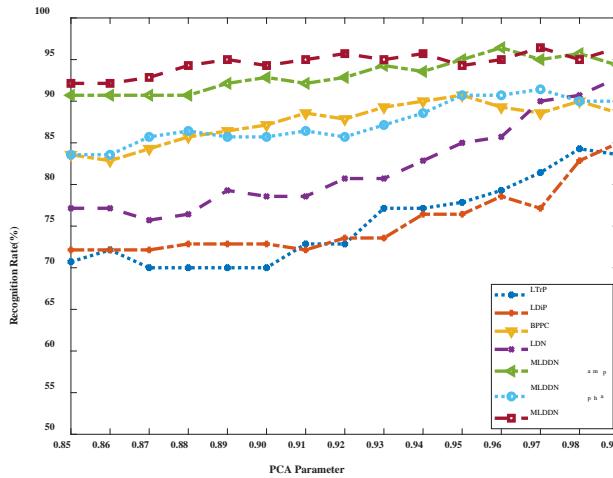


Fig. 13. Recognition rates (%) of different method with different PCA parameters for pigsty No.1.

Fig. 13 shows the recognition rates of pig identification after dimensionality reduction with PCA for pigsty No.1. Similar to the results of pigsty No.2, the recognition rate of MLDDN still keeps higher than other methods when the feature dimension was decreased.

4. Conclusion

This paper proposes a new local feature descriptor for pig identification. The directional information of Gabor amplitude and phase responses on multiple scales were encoded. More discriminative and robust information of adjacent directions was extracted for pig identification. In order to verify the effectiveness of the proposed method, two pigsties were taken as samples. Extensive experiments were conducted and the recognition rates achieved 89.8% and 95.71%. The proposed method for pig recognition was carried out in the real pigsty without any limitation, such as designated location and fixed posture. The proposed method can be used for video analysis of animal individual recognition and behavior recognition. Furthermore, although this method is used for pig identification, the local descriptor can also be applied to identification of other livestock.

References

- [1] J. N. Jover, M. Alcaniz-Raya, V. Gomez, S. Balasch, J. R. Moreno, V. Grau-Colomer and A. Torres, "An automatic colour-based computer vision algorithm for tracking the position of piglets," *Spanish Journal of Agricultural Research*, vol. 7, no. 3, pp. 535-549, March 2009. [Article \(CrossRef Link\)](#)
- [2] P. Ahrendt, T. Gregersen and H. Karstoft, "Development of a real-time computer vision system for tracking loose-housed pigs," *Computers and Electronics in Agriculture*, vol. 76, no. 2, pp. 169-174, May 2011. [Article \(CrossRef Link\)](#)
- [3] M. A. Kashiha, C. Bahr, S. Ott, C. Moons, T. Niewold, F. Ödberg and D. Berckmans, "Automatic identification of marked pigs in a pen using image pattern recognition," *Computers and Electronics in Agriculture*, vol. 93, pp. 111-120, April 2013. [Article \(CrossRef Link\)](#)
- [4] Larregui J I, Cazzato D, Castro S M, "An image processing pipeline to segment iris for unconstrained cow identification system," *Open Computer Science*, vol. 9, no. 1, pp. 145-159, September 2019. [Article \(CrossRef Link\)](#)
- [5] Kumar S, Singh S K, Abidi A I, et al., "Group Sparse Representation Approach for Recognition of Cattle on Muzzle Point Images," *International journal of parallel programming*, vol. 46, no. 5, pp. 812-837, 2018. [Article \(CrossRef Link\)](#)
- [6] Sian C, Jiye W, Ru Z, et al., "Cattle identification using muzzle print images based on feature fusion," in *Proc. of IOP Conference Series Materials Science and Engineering*, vol. 853, no. 1, pp.012051-012059, 2020. [Article \(CrossRef Link\)](#)
- [7] Y. Lu, X. He, Y. Wen and P. S. P. Wang, "A new cow identification system based on iris analysis and recognition," *International Journal of Biometrics*, vol. 6, no. 1, pp. 18-32, March 2014. [Article \(CrossRef Link\)](#)
- [8] H. T. Kim, H. L. Choi, D. W. Lee and Y. C. Yoon, "Recognition of individual Holstein cattle by imaging body patterns," *Asian Australasian Journal of Animal Sciences*, vol. 18, no. 8, pp. 1194-1198, August 2005. [Article \(CrossRef Link\)](#)
- [9] G. P. Corkery, U. A. Gonzales-Barron, F. Butler, K. Mc Donnell and S. Ward, "A preliminary investigation on face recognition as a biometric identifier of sheep," *Transactions of the ASABE*, vol. 50, no. 1, pp. 313-320, January 2007. [Article \(CrossRef Link\)](#)
- [10] Kumar S, Tiwari S, Singh S K, "Face Recognition of Cattle: Can it be Done?," *Proceedings of the National Academy of Sciences India*, vol. 86, no. 2, pp. 137-148, April 2016. [Article \(CrossRef Link\)](#)
- [11] K. X. Zhao and D. J. He, "Recognition of individual dairy cattle based on convolutional neural networks," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 2015, no. 5, pp. 181-187, March 2015. [Article \(CrossRef Link\)](#)
- [12] M. F. Hansen, M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter, M. Farish and B. Grieve, "Towards on-farm pig face recognition using convolutional neural networks," *Computers in Industry*, vol. 98, pp. 145-152, June 2018. [Article \(CrossRef Link\)](#)
- [13] Y. Z. Guo, W. X. Zhu, C. H. Ma and C. Chen, "Top-view recognition of individual group-housed pig based on Isomap and SVM," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 32, no. 3, pp. 182-187, February 2016. [Article \(CrossRef Link\)](#)
- [14] W. Huang, W. Zhu, C. Ma, Y. Guo and C. Chen, "Identification of group-housed pigs based on Gabor and local binary pattern features," *Biosystems Engineering*, vol. 166, pp. 90-100, February 2018. [Article \(CrossRef Link\)](#)
- [15] M. Ponti, T. S. Nazaré and G. S. Thumé, "Image quantization as a dimensionality reduction procedure in color and texture feature extraction," *Neurocomputing*, vol. 173, no. P2, pp. 385-396, January 2016. [Article \(CrossRef Link\)](#)
- [16] Fan C, Wang S, Zhang H, "Efficient Gabor phase based illumination invariant for face recognition," *Advances in Multimedia*, vol.2017, pp.1-11, November 2017. [Article \(CrossRef Link\)](#)
- [17] Zhang Z, Lu G, Yan J, et al., "Compact local Gabor directional number pattern for facial expression recognition," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, no. 3, pp.1236-1248, May 2018. [Article \(CrossRef Link\)](#)

- [18] Y. Z. Guo, W. X. Zhu, P.P. Jiao, C. H. Ma and J. J. Yang, "Multi-object extraction from topview group-housed pig images based on adaptive partitioning and multilevel thresholding segmentation," *Biosystems engineering*, vol. 135, pp. 54-60, May 2015. [Article \(CrossRef Link\)](#)
- [19] T. Ojala, M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51-59, May 1996. [Article \(CrossRef Link\)](#)
- [20] T. Jabid, M. H. Kabir and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI journal*, vol. 32, no. 5, pp. 784-794, October 2010. [Article \(CrossRef Link\)](#)
- [21] T. Jabid and O. S. Chae, "Facial Expression Recognition Based on Local Transitional Pattern," *International Information Institute (Tokyo). Information*, vol. 15, no. 5, pp. 2007-2018, May 2012. [Article \(CrossRef Link\)](#)
- [22] S. Shoaieilangari, W. Y. Yau, J. Li and E. K. Teoh, "Feature extraction through binary pattern of phase congruency for facial expression recognition," in *Proc. of ICARCV*, Guangzhou, China, pp. 166-170, December 5-7, 2012. [Article \(CrossRef Link\)](#)
- [23] A. R. Rivera, J. R. Castillo and O. O. Chae, "Local directional number pattern for face analysis: Face and expression recognition," *IEEE transactions on image processing*, vol. 22, no. 5, pp. 1740-1752, May 2013. [Article \(CrossRef Link\)](#)
- [24] Cigdem T, Kin-Man L, "Histogram-based Local Descriptors for Facial Expression Recognition (FER): A comprehensive Study," *Journal of Visual Communication and Image Representation*, vol. 55, pp. 331-341, August 2018. [Article \(CrossRef Link\)](#)
- [25] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Research*, vol. 20, no. 10, pp. 847-856, August 1980. [Article \(CrossRef Link\)](#)
- [26] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America. A, Optics and image science*, vol. 2, no. 7, pp. 1160-1169, July 1985. [Article \(CrossRef Link\)](#)
- [27] V. Štruc and N. Pavešić, "The complete Gabor-Fisher classifier for robust face recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1-26, April 2010.
- [28] Chen T, Ju S, Ren F, et al., "EEG emotion recognition model based on the LIBSVM classifier," *Measurement*, vol. 164, November 2020. [Article \(CrossRef Link\)](#)
- [29] Salem N, Hussein S, "Data dimensional reduction and principal components analysis," *Procedia Computer Science*, vol. 163, pp. 292-299, 2019. [Article \(CrossRef Link\)](#)
- [30] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959-971, October 1996. [Article \(CrossRef Link\)](#)



Weijia Huang received the BS degree in electronic information engineering and the MS degree in signal and information processing from Jiangsu University of Science and technology, China, in 2003 and 2006, respectively. She is currently pursuing the Ph.D. degree at the School of Electrical and Information Engineering, Jiangsu University, China. Her current research interests include pattern recognition, machine learning and computer vision.



Weixing Zhu received the Ph.D. degree in mechanical design and theory from Jiangsu University, China, in 2001. He is currently a professor at the school of electrical and information engineering, Jiangsu University, China. His current research interests include computer vision, intelligent detection and control.



Zhengyan Zhang received the MS degree in signal and information processing from Jiangsu University of Science and technology, Zhenjiang, Jiangsu, China, in 2007. He is currently pursuing the Ph.D. degree at the College of Telecommunications and Information Engineering in Nanjing University of Posts and Telecommunications. His current research interests include pattern recognition, machine learning and computer vision.



Yizheng Guo received the B.Sc., M.Sc., and Ph.D. degrees from Jiangsu University, China, in 2005, 2008, and 2018, respectively. He is currently a associate professor at the Nanjing Normal University Taizhou College of China, where he is involved in image processing, machine vision and artificial intelligence.