CrossMark

Gesture recognition based on an improved local sparse representation classification algorithm

Yang He[1] · Gongfa Li[1,2] · Yajie Liao[1] · Ying Sun[1,2] · Jianyi Kong[1,2] · Guozhang Jiang[1,2] · Du Jiang[1] · Bo Tao[1] · Shuang Xu[1] · Honghai Liu[3]

Abstract The sparse representation classification method has been widely concerned and studied in pattern recognition because of its good recognition effect and classification performance. Using the minimized l_1 norm to solve the sparse coefficient, all the training samples are selected as the redundant dictionary to calculate, but the computational complexity is higher. Aiming at the problem of high computational complexity of the l_1 norm based solving algorithm, l_2 norm local sparse representation classification algorithm is proposed. This algorithm uses the minimum l_2 norm method to select the local dictionary. Then the minimum l_1 norm is used in the dictionary to solve sparse coefficients for classify them, and the algorithm is used to verify the gesture recognition on the constructed gesture database. The experimental results show that the algorithm can effectively reduce the calculation time while ensuring the recognition rate, and the performance of the algorithm is slightly better than KNN-SRC algorithm.

Keywords Gesture recognition · l_2 norm · Sparse representation · Classification algorithm

✉ Gongfa Li
  ligongfa@wust.edu.cn

[1] The Key Laboratory of Metallurgical Equipment and Control of Education Ministry, Wuhan University of Science and Technology, Wuhan 430081, China

[2] Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

[3] School of Computing, University of Portsmouth, Portsmouth PO1 3HE, UK

1 Introduction

Vision-based gesture recognition is a hot topic in human-computer interaction, which is in line with human's natural communication habits, and can be used for long-distance and non-contact interaction. Vision-based gesture recognition involves technologies such as gesture detection segmentation, feature extraction and classification recognition. These techniques have developed in recent years, but there are still imperfections, including recognition rate and real-time [1,2]. At present, sparse representation theory has been gradually introduced into pattern recognition, which mainly uses the test sample to identify the linear sparse representation of the dictionary composed of training samples, and has good robustness and recognition effect [3].

After the proposed method of SRC, the theory of sparse representation has been paid more and more attention by researchers, and it has great development potential and broad application prospects in the fields of pattern recognition [4–9]. In the construction of redundant dictionaries, the researchers put forward various methods, such as Aharon et al. [10] based on the K-means algorithm proposed K-Singular Value Decomposition (K-SVD) algorithm, using the tracking method for sparse coding, and the use of SVD singular value decomposition for the dictionary update and the corresponding sparse coefficient update, which greatly accelerating the training algorithm convergence rate. In order to improve the calculation speed of SRC algorithm, Li et al. [11] proposed a local SRC algorithm, KNN-SRC, based on sparse representation algorithm, and KNN algorithm was used to select K training samples are similar with test samples to form a local dictionary. However, the computational complexity of the above algorithm was large and the calculation time was long. Qi et al. [12] proposed a classification method based on sparse neighbor representation, which is suitable

for non-linear distribution of sample set and low-dimensional data classification problem. It can improve the efficiency while ensuring the recognition rate, but the choice of local dictionary needs further study. In the sparse solution algorithm, there are greedy algorithm and convex optimization algorithm, respectively, to solve the minimized l_0 norm and minimize l_1 norm problem. Yang et al. [13,14] analyzed five existing fast minimized l_1 norm algorithms, namely Gradient Projection, Homotopy, Iterative Shrinkage-thresholding, Proximal Gradient, and Alternating Direction. In the face recognition experiment, Homotopy has a better effect in recognition rate and calculation time. Zhai et al. [15] also uses Homotopy algorithm to solve the sparse coefficient, combined with color information, which effectively improves the robustness of face recognition in the case of occlusion.

The convex optimization algorithm has high precision, but high computational complexity. In order to solve the problem of computational complexity faced by l_1 norm solving algorithm, we introduce the idea of local sparse representation, according to the steps of KNN-SRC algorithm, we use the minimized l_2 norm to select the local dictionary to calculate, the l_2 norm local sparse representation classification algorithm was proposed. In order to verify the validity of the algorithm, the influence of the parameter K on the recognition rate was verified on two gesture libraries, and the optimal K value was selected. Finally, the recognition rate and the average run time of the algorithm on the two gestures were compared with the change of the dimension. The result shows that the algorithm can guarantee that Recognition rate while improving efficiency.

2 Sparse representation classification algorithm

According to the relevant references, the test samples can be represented by the sparse linear combination of the dictionary elements, and then the sparse reconstruction algorithm is used to solve the sparse coefficients. Sparse representation classification algorithm mainly involves two aspects, namely the redundant dictionary structure and the sparse coefficient solution [16,17]. The following mainly introduces the framework model of sparse representation classification algorithm.

(1) The test sample is a linear representation of the training sample

y can be represented linearly by A_j:

$$y = A_j v_j = a_{j,1} v_{j,1} + a_{j,2} v_{j,2} + \cdots + a_{j,n_j} v_{j,n_j} \tag{1}$$

where n is the total number of sample;

$A_j = [a_{j,1}, a_{j,2} \ldots, a_{j,n_j}] \in \Re^{m \times n_j}$;

$v_j = [v_{j,1}, v_{j,2} \ldots, v_{j,n_j}] \in \Re^{n_j \times 1}$ is the linear representation coefficient.

The type of test sample has not been determined, so the entire C class of gesture samples constitute a redundant dictionary matrix:

$$A = [A_1, \ldots, A_j, \ldots, A_C]$$
$$= [a_{1,n_1}, \ldots, a_{j,n_j}, \ldots, a_{C,n_C}] \in \Re^{m \times n}$$

y can be denoted by A as:

$$y = Ax \tag{2}$$

where $x = [0, \cdots 0, v_{j,1}, v_{j,2}, \cdots, v_{j,n_j}, 0, \cdots, 0]$ is the sparse coefficient of test sample, and non-zero elements are corresponding to the j-th class. If $S = \|x\|_0$ and $S \ll n$, then that x is sparse, S is the degree of sparseness.

(2) A sparse calculation method with minimized norm

Since the dimension of the dictionary A is less than the number of samples, that is $m < n$, Eq. (2) belongs to the undetermined equation and the solution is not unique [18]. Equation (2) can be solved by minimizing the l_0 norm:

$$\hat{x} = \arg \min_x \|x\|_0 \quad s.t. \ y = Ax \tag{3}$$

According to the theory of compression perception, when x is sufficiently sparse, Eq. (3) can be equivalent to solving the problem of minimizing l_1 norm as:

$$\hat{x} = \arg \min_x \|x\|_1 \quad s.t. \ y = Ax \tag{4}$$

Due to environmental factors, the actual sample collection will be affected by noise, light, etc., the test sample cannot be better represented by linear combination of training sample. And the recognition rate is easily affected. To improve robustness, a noise constraint can be added:

$$\hat{x} = \arg \min_x \|x\|_0 \quad s.t. \ \|y - Ax\|_2 \le \varepsilon \tag{5}$$
$$\hat{x} = \arg \min_x \|x\|_1 \quad s.t. \ \|y - Ax\|_2 \le \varepsilon. \tag{6}$$

(3) The minimum residual classification

The sparse coefficients of the linear combination are obtained, then the test samples can be reconstructed by using sparse coefficients and redundant dictionaries, and then the reconstructed samples and test samples for each class are compared [19,20], judge the type of test sample based on the minimum residual.

$$\gamma_j(y) = \|y - A\delta_j(\hat{x})\|_2 \quad j = 1, 2, \ldots, C \tag{7}$$
$$I(y) = \arg \min_j r_j(y) \tag{8}$$

where $\gamma_j(y)$ is the residual for the j-th class of the sample, $\delta_j(\hat{x})$ is the coefficient of the corresponding position of the
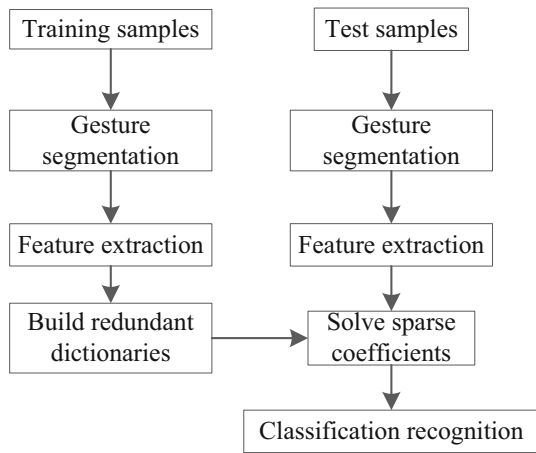
Training samples

Test samples

Gesture segmentation

Gesture segmentation

Feature extraction

Feature extraction

Build redundant dictionaries

Solve sparse coefficients

Classification recognition

Fig. 1 Gesture recognition framework

j-th class sample, and the other position is 0; $I(y)$ is the category of the test sample.

The SRC based gesture recognition framework is shown in Fig. 1.

When the gesture is recognized, the test sample may not belong to the gesture in the gesture sample library or not the gesture image. The classifier needs to judge the invalid gesture, and increasing the validity judgment can improve the robustness of the algorithm [21]. When the gesture recognition is done by sparse representation of the classification algorithm, the sparse coefficient is obtained on all the training samples [22]. In theory, the effective test gestures belong to the gestures in the sample library and can be expressed linearly by the similar gestures. The sparse coefficients are nonzero only in the same sample, the invalid test gestures may be distributed in a variety of samples [23]. The sparse coefficient and residual graphs of the ineffective test samples on the grab palette are shown in Fig. 2. It can be seen from the figure that the larger nonzero terms in the sparse coefficients are more distributed on the multi-class training samples. The

residuals on the class are not very different, even if the result of the classification based on the minimum residual will be erroneous and meaningless.

Therefore, a Sparse Concentration Index (SCI) can be defined according to the distribution of sparse coefficients,

$$\text{SCI} = \frac{C \cdot \max_j \|\hat{x}_j\|_1 / \|\hat{x}\|_1 - 1}{C - 1} \in [0, 1] \qquad (9)$$

where C is number of training samples; $\hat{x}$ is sparse coefficient.

The SCI field is [0, 1], the effective test sample corresponds SCI value is 1 and the valid test sample corresponds SCI value is 0 in theory. In fact, SCI value is not strictly 0 or 1, you can set a threshold $\tau \in (0, 1)$, when SCI $\geq \tau$, the test sample is valid and the next step of classification, otherwise it is judged to be invalid gesture. So before the SRC step 5 needs to determine the SCI value, which is better at judging invalid gestures than using the minimum residuals alone. The result of the classification based on the minimum residual will be erroneous and meaningless, the result was determined by SCI classification is easy to be understood and accepted.

3 Local sparse representation classification

3.1 Thought of local sparse representation of thought

For a typical gesture recognition scene, a training sample data set containing a class C sample is given to identify a test image $y \in \Re^m$. Suppose that the class $j \in [1, \ldots, C]$ contains n_j gesture samples, which is a column vector matrix $A_j = [a_j, \ldots, a_{n_j}]$, which contains the matrix $A = [A_1, \ldots, A_C] \in \Re^{m \times n}$ of the gesture samples of all classes, where m is the dimension of the sample feature, $n = n_1 + \cdots + n_j$ is the number of all training samples,
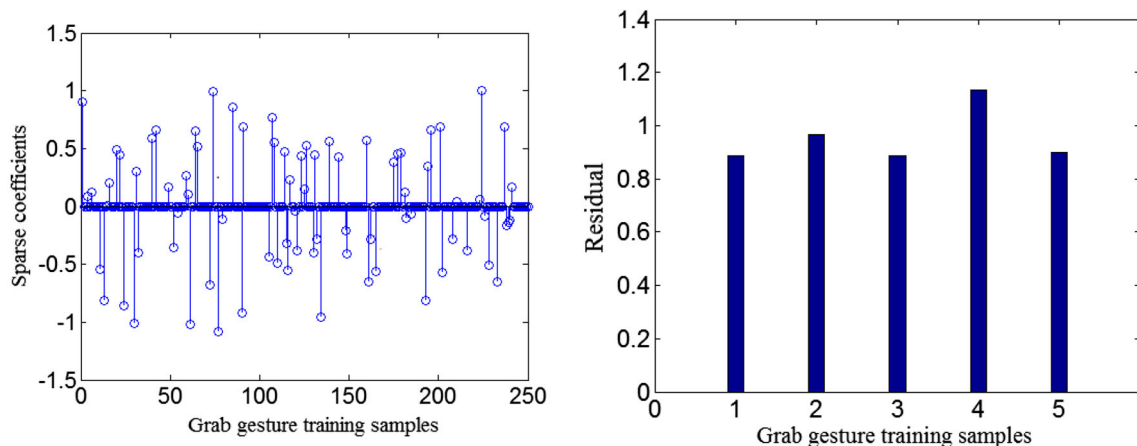

Fig. 2 Sparse coefficient and residual graph of invalid test samples

assuming that the test image y can be represented by the training sample set, then the problem can be expressed as $y = Ax + z$, where x is the coefficient vector representing the relationship between y and some of the columns of A.

According to the principle of compression perception, when solving the sparse coefficients, the minimization of l_1 norm problem can be transformed into the minimization of l_2 norm convex optimization problem. In this paper, Homotopy method is used to solve the minimized l_1 norm problem. Homotopy algorithm is an unconstrained optimization problem, which can be iteratively solved in the direction of the nearest convergence center in the process of finding the optimal solution [24,25]. However, compared with the greedy algorithm, the computational complexity is high, so it is necessary to propose a method to approximate the minimization of l_1 norm problem, which can without reducing the recognition rate under the premise of effectively reducing the overall running time. Regularization technique is one of the effective methods to deal with constraint optimization. Therefore, the Lagrangian objective function $v(x)$ of the minimization of l_1 norm is defined as a series of vector operations. The optimization problem becomes:

$$v(x) = \left\| y - \sum_{i=1}^{n} a_i x_i \right\|_2^2 + \lambda \sum_{i=1}^{n} |x_i| \tag{10}$$

where $a_i \in \Re^m$ is the i-th column of A, x_i is the i-th element of the coefficient vector, and λ is the sparse control parameter. It is assumed that the maximum number of K elements in the coefficient vector is nonzero, and any of the remaining elements $x_j = 0$, then $\left\| a_j x_j \right\|_2 = 0$, a_j has no effect on $v(x)$. That is, due to the existence of zero coefficient, resulting in the corresponding training samples have no effect on the results of the solution, and other training samples related to test samples, which is the reason why the test sample can be linearly represented by similar samples. The objective function can be redefined as:

$$v(\alpha) = \left\| y - \sum_{i=1}^{K} \omega_i \alpha_i \right\|_2^2 + \lambda \sum_{i=1}^{K} |\alpha_i| \tag{11}$$

The corresponding sample of the k elements constitutes the new dictionary Ω, then ω_i is the i-th column of the matrix Ω. Since the error estimate does not depend on the zero element in x, the new objective function $v(\alpha)$ is theoretically the same as $v(x)$, and $K \ll n$, it is calculated faster. With the new dictionary Ω and the coefficient vector α, the minimization of the l_1 norm can be redefined:

$$\hat{\alpha} = \arg\min_{\alpha} \| y - \Omega\alpha \|_2^2 + \lambda \|\alpha\|_1. \tag{12}$$

3.2 KNN-SRC classification algorithm

The KNN classifier is widely used, and the conventional KNN classifier simply computes the K neighbors of the test sample and classifies the test samples into the category containing the most neighbors. The conventional KNN classifier, regardless of the proximity of the sample distance from the near sample treats each neighbor in the classification of the test sample. Although the classification criterion of KNN is not strict, it can detect the training samples which are not similar or far away from the test samples. Therefore, a local SRC method KNN-SRC is proposed [26], it combines the KNN algorithm with the sparse representation algorithm. The method is divided into two steps: the first step to calculate the distance between the training sample and the test sample, select K test samples of the neighbors from the training sample:

$$d_i(y) = \| y - a_i \|^2 \tag{13}$$

The second step does not consider the influence of other training samples on the final classification decision, the selected K neighboring test samples as sub-dictionary $\Omega = [\omega_1, \ldots, \omega_K]$, and then use (13) to solve the sparse coefficient, and according to the minimum residual classification.

The steps of the KNN-SRC algorithm are summarized as follows:

Step 1. Input: training sample $A \in \Re^{m \times n}$, test sample $y \in \Re^{m \times 1}$;
    2. For each column of A and y are normalized by l_2 norm;
    3. Calculate the distance between the training sample and the test sample:

$$d_i(y) = \| y - a_i \|^2$$

    4. The distances are sorted in descending order, and the K samples of the test samples are searched from the training samples to form the subset $\Omega \in \Re^{m \times k}$;
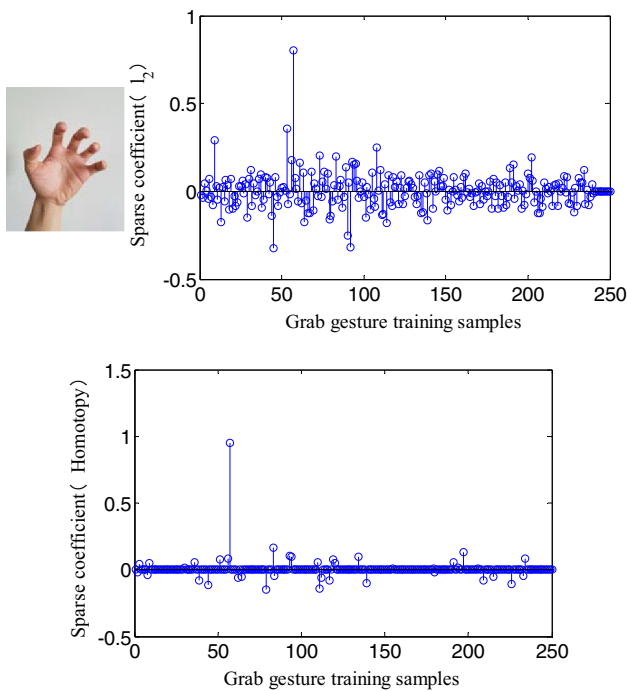   5. Solving minimize l_1 norm on subset Ω:

$$\hat{\alpha} = \arg\min_{\alpha} \| y - \Omega\alpha \|_2^2 + \lambda \|\alpha\|_1$$

    6. Calculate the residual for each class $j \in [1, C]$:

$$r_j(y) = \left\| y - \Omega\delta_j(\hat{\alpha}) \right\|_2$$

    7. Calculate SCI:

$$SCI = \frac{C \cdot \max_j \|\alpha_j\|_1 / \|\hat{\alpha}\|_1 - 1}{C - 1}$$

Fig. 3 Sparse coefficient of test samples

8. Output: If $SCI \geq \tau$, Category $I(y) = \arg \min_j r_j(y)$, otherwise it is an invalid gestures.

3.3 l_2 Norm local sparse representation classification algorithm

Besides the l_0 norm and the l_1 norm, the minimum l_2 norm method can be used to find the estimate of x in the sparse coefficient solution of SRC.

$$\hat{x}_{l_2} = \arg \min_x \|y - Ax\|_2^2 \tag{14}$$

where x is the coefficient vector, y is the test sample, and A is the training sample.

The above equation can be solved by the least squares, it is calculated by pseudo inverse matrix:

$$\hat{x}_{l_2} = (A^T A)^{-1} A^T y \tag{15}$$

Since the pseudo inverse matrix $(A^T A)^{-1} A^T$ is the same for all y, it is only necessary to solve once and save it, and then solve the corresponding coefficients for each y, so the method is simpler and faster to calculate. But the coefficient vector $\hat{x}_{l2}$ is relatively dense, as shown in Fig. 3, it can be seen that the sparsity is not stronger than the l_1 norm. However, the larger values of the sparse coefficients solved by the l_2 norm are basically the same as those calculated by the l_1 norm.

Therefore, we can combine the K corresponding coefficients in $|\hat{x}_{l_2}|$ to form a subset $\Omega = [\omega_1, \ldots, \omega_K]$, which not

only selects the dictionary atoms most likely to be related to the test samples, but also reduces the number of atoms in the new dictionary, and reduces the amount of subsequent calculations, finally the dictionary uses a sparse representation algorithm that minimizes the l_1 norm to classify the test samples.

It is not easy to find the exact useful sample by minimizing the l_2 norm method, but it has a high degree of sample approximation [27]. As mentioned earlier, the calculation of the minimized l_2 norm is fast. Although the coefficients l_2 method of solution is dense, some of the peaks are similar to the nonzero coefficients in the l_1 method, that is, the training image corresponding to the matching test image category.

l_2 Norm local sparse representation classification (l_1-LSRC) uses the least squares method to approximate the l_1 norm to minimize, and obtains the computational speed of the least squares and the robustness of the SRC [28,29]. In the recognition phase, the linear regression approximates SRC. Firstly, the pseudo inverse $(A^T A)^{-1} A^T$ is used to quickly calculate the coefficient vector $\hat{x}_{l_2}$, the elements in the vector $|\hat{x}_{l_2}|$ is arranged in descending order, selecting the K samples that corresponding to the maximum coefficient in $|\hat{x}_{l_2}|$, and establishing the approximate dictionary matrix Ω. And use this smaller dictionary Ω is used to as input of the minimize l_1 norm solution, a new sparse vector α is calculated. Finally, the sparse and concentrated index SCI is calculated for the test sample image, and the minimum residue is used to determine the most likely category.

The l_2 norm local sparse representation classification method steps are summarized as follows:

Step 1. Input: training samples $A \in \Re^{m \times n}$, test samples $y \in \Re^{m \times 1}$;

2. Each column of A and y is normalized by l_1 norm;

3. Use the inverse operator $(A^T A)^{-1} A^T$ to solve the minimization of l_1 norm:

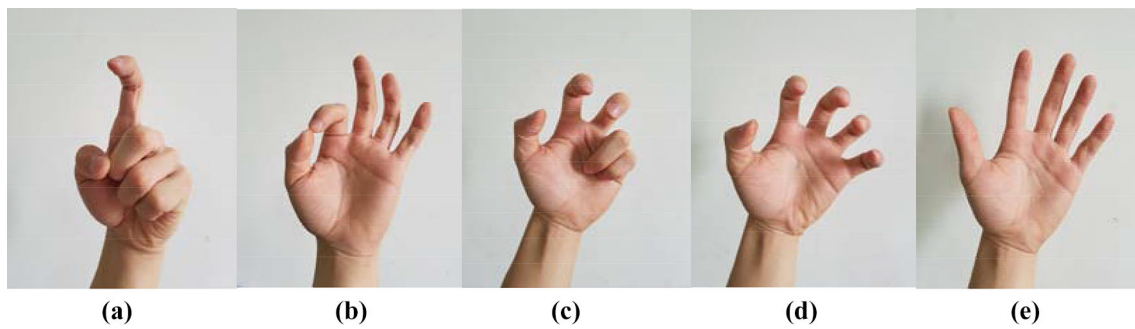$$\hat{x}_{l_2} = (A^T A)^{-1} A^T y$$

4. Select the K samples that corresponding to the maximum coefficient from A to form a subset $\Omega \in \Re^{m \times k}$;

5. Solve the minimum l_1 norm by an approximate subset dictionary Ω:

$$\hat{\alpha} = \arg \min_\alpha \|y - \Omega\alpha\|_2^2 + \lambda \|\alpha\|_1$$

6. Calculate residuals for each class $j \in [1, C]$:

$$r_j(y) = \|y - \Omega\hat{\alpha}_j\|_2$$

7. Output: if $SCI \geq \tau$, category $I(y) = \arg \min_j r_j(y)$, otherwise the gesture is invalid.

Fig. 4 Five types of grabbing hand samples. **a** Five fingers grabbing, **b** three fingers grabbing, **c** two fingers pinching, **d** single finger hook, **e** openning

In this algorithm, the coefficients are quickly calculated by the least squares method, and the training samples corresponding to K larger coefficients are selected to reduce the number of training samples to minimize the l_1 norm and remove the jamming samples. Although only the third step is added, the K samples are extracted by the least squares method greatly reduce the computational complexity, especially when K is small, the recognition rate can be effectively accelerated.

4 Simulation of experiment results of gesture recognition

4.1 Establishment of gesture sample library

In order to verify the recognition effect of the gesture recognition algorithm, the gesture image is collected to establish the gesture sample library, and the influence of each factor on gesture recognition is analyzed. In the experiment, the specified gesture samples are selected, the ellipse model is established in YCbCr color space for segmentation gestures, and the Hu invariant moments and HOG characteristics are extracted [30,31]. (1) Grabbing gesture sample library

The camera is used to collect five kinds of typical grabbing hand samples for five people, namely, single finger hooking, two fingers pinching, three fingers grabbing, five fingers grabbing, and opening, which as shown in Fig. 4. Each of the five types of gestures is collected 20 pictures, that is, 100 pictures of each type of gesture, a total of 500 pictures can be used for experiments [32]. In the process of collecting gestures, also consider gesture rotation, scale, light and background changes.

(2) ASL sample library

The ASL sample library contains 26-letter gestures (the letters j and z gestures are dynamic, so this is not considered here) as shown in Fig. 5, which were collected for five individuals by Kinect in the same light and scene. Each operator and letter has a color picture and a depth picture, and each of

which collects 20 color pictures of 24 letters, a total of 2400 color pictures can be used for the experiment [33], the part of color pictures as shown in Table 1.

4.2 Effect of parameter K on algorithm recognition rate

The l_1-LSRC and KNN-SRC algorithms use the local sparse representation to select the K training samples most relevant to the test samples for classification recognition, where $K < n$, which excludes unrelated training samples, reduces the computational complexity while maintaining the robustness of the SRC algorithm. In this paper, we experimentally validate the influence of different K values on the recognition rate to select the best K. In the experiment, 50 training samples are selected for each class, 50 as test samples, the principal component analysis (PCA) is used to reduce dimension, the features are reduced to 100, Homotopy algorithm was used to solve the problem of minimizing l_1 norm. The experimental results are shown in Figs. 6 and 7.

The change of recognition rate with K on the grasping gesture library was shown in Fig. 4. It can be seen from the figure that the recognition rate of the SRC algorithm does not change with K, and it is stable at 91.4%, which is the reference of the other two algorithms. As the value of K increases, the more relevant sample increases, and the recognition rate of l_1-LSRC and KNN-SRC algorithm gradually increases. And the recognition rate of l_1-LSRC algorithm is almost unchanged after $K > 125$, it is about 90%, which closes to the SRC algorithm recognition rate; when $K > 150$, KNN-SRC algorithm recognition rate is almost unchanged at about 88%, which closes to the SRC algorithm recognition rate. Therefore, the best K value of the l_1-LSRC algorithm is 125 and the KNN-SRC algorithm is 150 in grabbing gesture sample library. When the value of K is low, the recognition rate of the l_1-LSRC algorithm is higher than the KNN-SRC algorithm, the reason is that the sample selected by the minimization l_1 norm is more relevant than the sample selected by KNN.
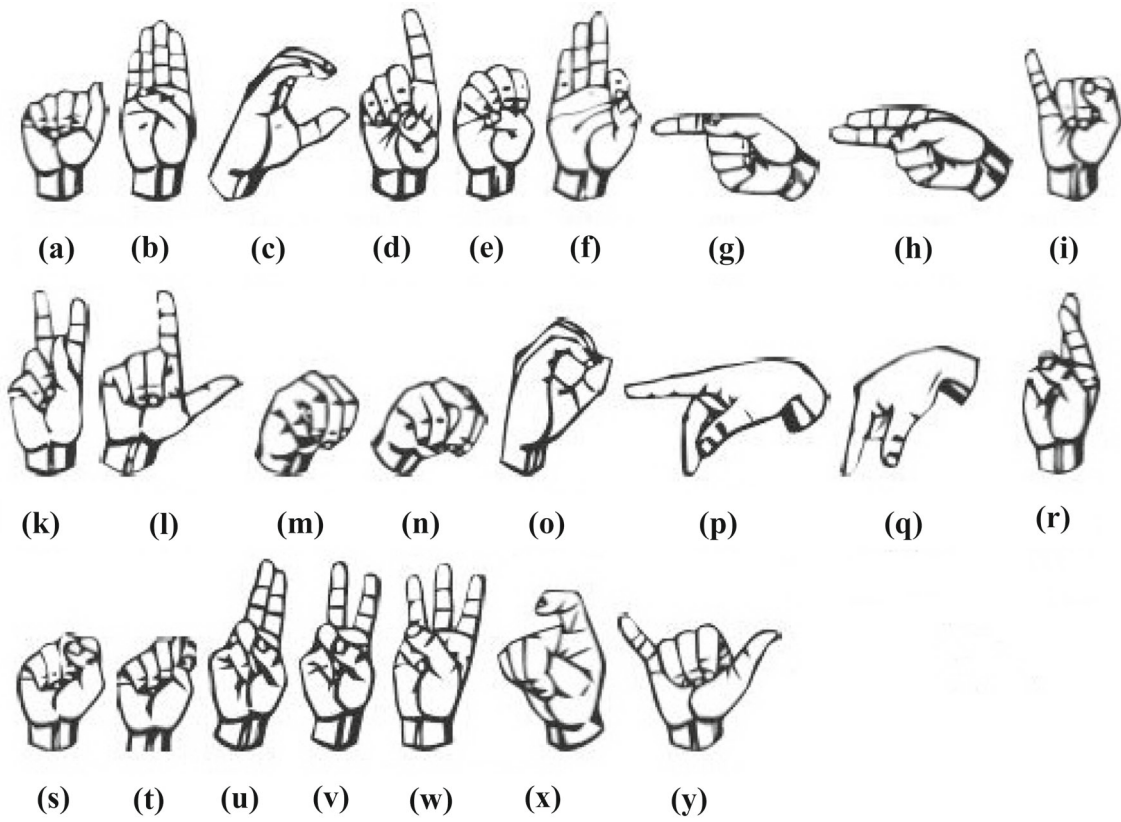
Fig. 5 ASL finger spelling alphabet
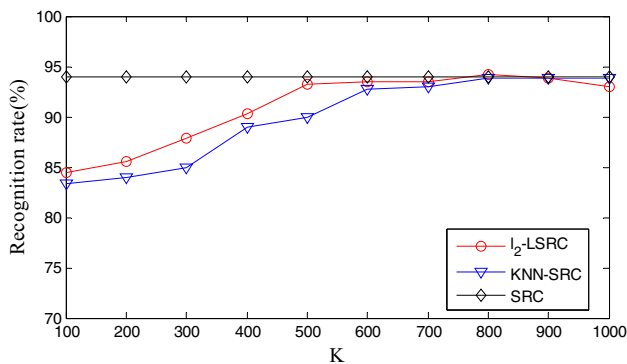
Table 1 Part of the ASL gesture sample

The change of recognition rate with K on the ASL sample library was shown in Fig. 7. It can be seen from the figure that the recognition rate of SRC algorithm is not affected by K, it is stable at 94.2%, and the recognition rate of l_1-LSRC and KNN-SRC algorithm increases with the increase of K. When K changes between [500, 1000], the recognition rate of l_1-SRC algorithm is relatively stable, and it is about 93%.

When K changes between [600, 1000], the recognition rate of KNN-SRC algorithm is relatively stable and about 92%, which both are close to SRC algorithm. Therefore, the best K value of the l_1-LSRC algorithm on the ASL sample library is 500, and the best K value of the KNN-SRC algorithm is 600. It can be seen from the figure that when K is low, the

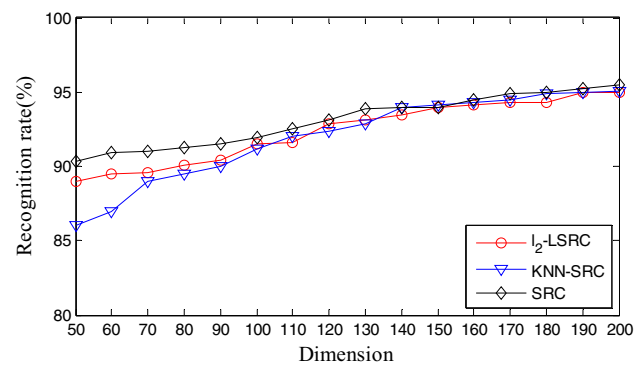**Fig. 6** Change of recognition rate with K on the grasping gesture library
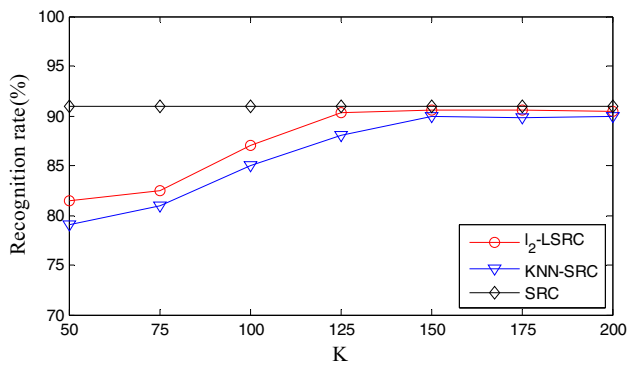
Fig. 8 Change of recognition rate with dimension on the grasping gesture library

Fig. 7 Change of recognition rate with K on the ASL sample library

Fig. 9 Change of average running time with dimension on the grasping gesture library

recognition rate of l_1-LSRC algorithm is higher than KNN-SRC algorithm.
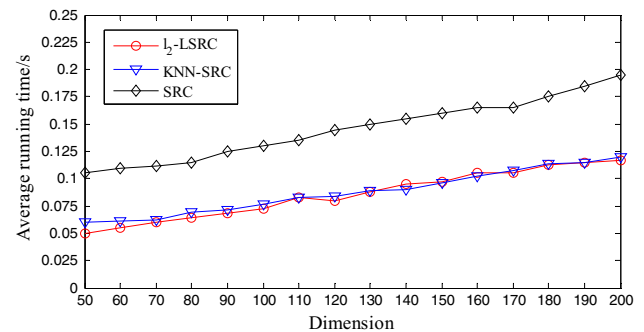
4.3 Comparison of the performance of algorithms

The SRC, KNN-SRC and l_1-SRN algorithm are compared in terms of recognition rate and average run time. At the time of experiment, 50 training samples and 50 test samples were randomly selected from each of the grabbing library, and the PCA dimensionality was extracted after extracting the features. In addition, setting the parameters of each algorithm, $K = 125$ in the l_1-LSRC algorithm, $K = 150$ in the KNN-SRC algorithm, the experimental results were shown in Figs. 6 and 7.

It can be seen from Fig. 8 that the recognition rate of each algorithm increases with the increase of the dimension. When the dimension is low, the change of recognition rate is great and the difference is large. When the dimension is high, the recognition rate tends to be flat and the difference is small. The SRC algorithm has higher recognition rate than the l_1-LSRC and KNN-SRC algorithms under the same dimension. When the dimension is low, the recognition rate of l_1-LSRC algorithm is higher than KNN-SRC algorithm, which is smaller both than SRC algorithm. When the dimen-

sion is greater than 100, the recognition rate of the l_1-LSRC algorithm and KNN-SRC algorithm is close to the recognition rate of SRC algorithm, which is about 91%. It can be seen from Fig. 9 that the average running time of each algorithm increases with the increase of the dimension, the SRC algorithm has the highest average running time, the lowest is 0.105 s, the highest is 0.195 s, followed by l_1-LSRC and KNN-SRC algorithm, the average running time is between [0.051, 0.119 s], and the average running time of l_1-LSRC algorithm is lower than KNN-SRC algorithm.

In order to further compare the algorithm performance, 50 training samples and 50 test samples were selected randomly from each class in the ASL sample library, and after the feature is extracted. The principal component analysis (PCA) is used to reduce dimension. The goal of PCA is to express high-dimensional data in low-dimensional subspace. It mainly projects the original sample data into a low-dimensional space by linear transformation. And $K = 500$ in the l_1-SRC algorithm, $K = 600$ in the KNN-SRC algorithm, the experimental results were shown in Figs. 8 and 9.

It can be seen from Figs. 8 and 9 that the recognition rate and the average calculation time increase with the increase of the dimension. Due to the large number of samples was selected on the ASL sample library, the recognition rate and the average run time were relatively high compared to the
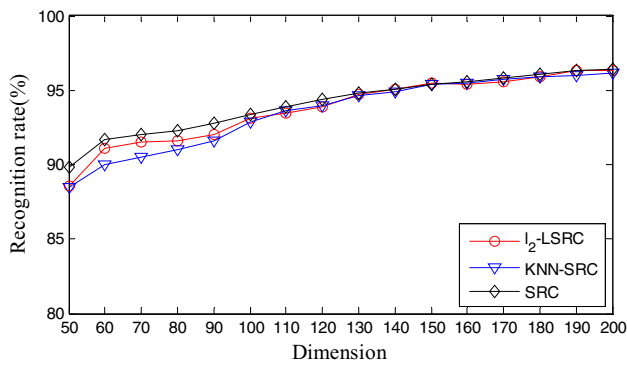
Fig. 10 Change of recognition rate with dimension on ASL sample library
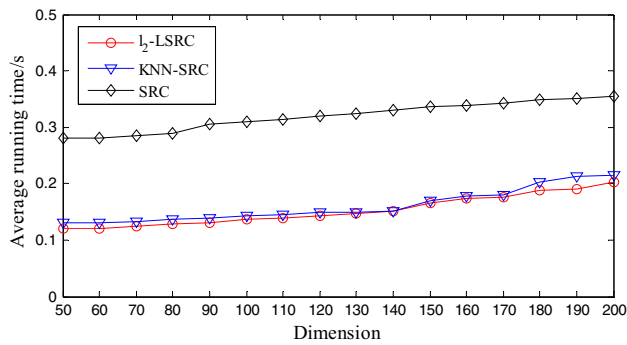
Fig. 11 Change of average running time with dimension on ASL sample library

crawl hand library. It can be seen from Fig. 10 that the SRC algorithm has the highest recognition rate. When the dimension is low, the recognition rate of l_1-LSRC algorithm is higher than KNN-SRC algorithm, which is smaller both than SRC algorithm. When the dimension is greater than 100, the recognition rate of the l_1-LSRC algorithm and KNN-SRC algorithm is close to the recognition rate of SRC algorithm, which is about 94%. It can be seen from Fig. 11 that the lowest of the average running time of the SRC algorithm is 0.282s, the highest is 0.355s, and the average running time of the l_1-LSRC and KNN-SRC algorithm is between [0.121s, 0.214s]. The l_1-LSRC algorithm is better than the KNN-SRC algorithm in the computation time.

The specific results data was shown in Tables 2 and 3, and the comparison of the performance of algorithms under different gesture sample library as shown in Tables 2 and 3. According to the data in the table, considering the recognition rate and the average running time, the l_1-LSRC algorithm is better than the KNN-SRC algorithm. Therefore, the l_1-LSRC and KNN-SRC algorithm have chosen the K most relevant training samples rather than all training samples, making the recognition rate and SRC little difference, but $K < n$, the average run time is much lower than SRC. So they improve the efficiency while ensuring algorithm accuracy. In addition, the l_1-LSRC and KNN-SRC algorithm have different way in choosing the K, the KNN-SRC algorithm is executed for each test sample are required to calculate the distance, while the l_1-LSRC algorithm only needs to compute a pseudo inverse

Table 2 Comparison of the performance of algorithms on the grasping gesture library

Dimension	60	80	100	120	140	160	180	200
Recognition rate								
SRC	90.9	91.3	91.9	93.1	94.0	94.5	95.0	95.5
KNN-SRC	87.0	89.5	91.2	92.4	94.0	94.3	94.9	95.1
l_1-LSRC	89.5	90.1	91.5	92.9	93.5	94.1	94.3	95.0
Average running time								
SRC	0.1095	0.1150	0.1300	0.1450	0.1550	0.1650	0.1750	0.1950
KNN-SRC	0.0610	0.0695	0.0770	0.0840	0.0900	0.1020	0.1140	0.1195
l_1-LSRC	0.0550	0.0645	0.0725	0.0800	0.0950	0.1050	0.1125	0.1170

Table 3 Comparison of the performance of algorithms on ASL sample library

Dimension	60	80	100	120	140	160	180	200
Recognition rate								
SRC	91.7	92.3	93.4	94.4	95.1	95.6	96.1	96.4
KNN-SRC	90.0	91.0	92.9	94.0	94.9	95.5	95.9	96.2
l_1-LSRC	91.1	91.6	93.1	93.9	95.1	95.4	95.9	96.3
Average running time								
SRC	0.280	0.289	0.309	0.320	0.330	0.339	0.349	0.355
KNN-SRC	0.130	0.137	0.143	0.149	0.152	0.179	0.202	0.214
l_1-LSRC	0.120	0.120	0.137	0.144	0.152	0.173	0.189	0.202

matrix, and it is more relevant to calculate the sample selected by l_1 norm. So the l_1-LSRC algorithm is better than the KNN-SRC algorithm in recognition rate and computation time.

5 Conclusions

Aiming at the problem of high computational complexity based on l_1 norm solving algorithm, this paper proposed l_1 norm local sparse representation classification algorithm based on the introduction of local sparse representation method. The local dictionary was chosen by minimizing l_1 norm and was carried out on the dictionary to solving sparse coefficients. In order to verify the validity of the algorithm, the influence of the parameter K on the recognition rate was verified in two gesture libraries. The optimal K value was selected and the recognition rate and the average running time of each algorithm were compared. The experimental results showed that using the idea of local sparse representation, the l_1-LSRC algorithm can effectively reduce the computation time while ensuring the recognition rate. And the performance of l_1-LSRC algorithm is slightly better than that of KNN-SRC algorithm.

References

1. Guan, R., Xu, X.G., Luo, Y.Y., et al.: A computer vision-based gesture detection and recognition technique. Comput. Appl. Softw. **30**(1), 155–159 (2013)
2. Yi, J.G., Chneg, J.H., Ku, X.H.: Review of gestures recognition based on vision. Comput. Sci. **43**(z1), 103–108 (2016)
3. Wright, J., Yang, A.Y., Ganesh, A., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210–227 (2009)
4. Li, X.Z., Wu, J., Cui, Z.M., et al.: Sparse representation method of vehicle recognition in complex traffic scenes. J. Image Graph. **17**(3), 90–95 (2012)
5. Cui, M., Prasad, S.: Class-dependent sparse representation classifier for robust hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. **53**(5), 2683–2695 (2015)
6. Ning, Y.N., Li, D.Z., Han, X., et al.: Gesture recognition method based on sparse representation. Comput. Eng. Des. **37**(9), 2548–2552 (2016)
7. Guo, Y.M., Zhao, G.Y., Pietikainen, M.: Dynamic facial expression recognition with atlas construction and sparse representation. IEEE Trans. Image Process. **25**(5), 1977–1992 (2016)
8. Yang, W.J., Kong, L.F., Wang, M.Y.: Hand gesture recognition using saliency and histogram intersection kernel based sparse representation. Multimed. Tools Appl. **75**(10), 6021–6034 (2016)
9. Meng, F.R., Tang, Z.Y., Wang, Z.X.: An improved redundant dictionary based on sparse representation for face recognition. Multimed. Tools Appl. **76**(1), 895–912 (2017)
10. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process. **54**(11), 4311–4322 (2006)
11. Li, C.G., Guo, J., Zhang, H.G.: Local sparse representation based classification. In: 2010 20th International Conference on Pattern Recognition (ICPR), pp. 649–652 (2010)
12. Wang, Q., Hui, K.H.: Sparse neighbor representation for classification. Comput. Eng. Des. **34**(4), 1425–1431 (2013)
13. Yang, A.Y., Ganesh, A., Sastry, S.S., et al.: Fast l_1-minimization algorithms and an application in robust face recognition: a review. In: IEEE International Conference on Image Processing, pp. 1849–1852 (2010)
14. Yang, A.Y., Zhou, Z., Ganesh, A., et al.: Fast l_1-minimization algorithms for robust face recognition. IEEE Trans. Image Process. **22**(8), 3234–3246 (2013)
15. Zhai, Y.K., Gan, J.Y., Li, J.W.: Study of occluded robust face recognition approach based on homotopy algorithm and color information fusion. J. Signal Process. **27**(11), 1762–1768 (2011)
16. Miao, W., Li, G.F., Jiang, G.Z., et al.: Optimal grasp planning of multi-fingered robotic hands: a review. Appl. Comput. Math. **14**(3), 238–247 (2015)
17. Li, G.F., Liu, J., Jiang, G.Z., et al.: Numerical simulation of temperature field and thermal stress field in the new type of ladle with the nanometer adiabatic material. Adv. Mech. Eng. **7**(4), 1–13 (2015)
18. Li, G.F., Gu, Y.S., Kong, J.Y., et al.: Intelligent control of air compressor production process. Appl. Math. Inf. Sci. **7**(3), 1051–1058 (2013)
19. Liao, Y.J., Li, G.F., Sun, Y., et al.: Simultaneous calibration: a joint optimization approach for multiple kinect and external cameras. Sensors **17**(7), 1491 (2017)
20. Donoho, D.: For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. Commun. Pure Appl. Math. **59**(6), 797–829 (2006)
21. Miao, W., Li, G.F., Sun, Y., et al.: Gesture recognition based on sparse representation. Int. J. Wirel. Mob. Comput. **11**, 348–356 (2016)
22. Li, G.F., Miao, W., Jiang, G.Z., et al.: Intelligent control model and its simulation of flue temperature in coke oven. Discret. Contin. Dyn. Syst. S (DCDS-S) **8**(6), 1223–1237 (2015)
23. Ashfaq, T., Khurshid, K.: Classification of hand gestures using Gabor filter with Bayesian and Naïve Bayes classifier. Int. J. Adv. Comput. Sci. Appl. **7**, 276–279 (2016)
24. Khaledian, M., Menhaj, M.B.: Real-time vision-based hand gesture recognition using SIFT features. Telkomnika Indones. J. Electr. Eng. **15**, 162–170 (2015)
25. Yu, B., Qin, Y.M.: Generating test case for algebraic specification based on Tabu search and genetic algorithm. Cluster Comput. **20**(1), 277–289 (2017)
26. Chen, D., Li, G.F., Sun, Y., et al.: An Interactive image segmentation method in hand gesture recognition. Sensors **17**(2), 253 (2017)
27. Ghotkar, A., Vidap, P., Ghotkar, S.: Vision based hand gesture recognition using Fourier descriptor for Indian sign language. Signal Image Process. **7**, 29–38 (2016)
28. Escalante, H.J., Guyon, I., Athitsos, V.: Principal motion components for one-shot gesture recognition. Pattern Anal. Appl. **20**, 167–182 (2017)
29. Zhuo, T.: Face recognition from a single image per person using deep architecture neural networks. Cluster Comput. **19**(1), 73–77 (2016)
30. Cui, R.X., Yan, W.S.: Mutual synchronization of multiple robot manipulators with unknown dynamics. J. Intell. Robot. Syst. **68**(2), 105–119 (2012)
31. Ju, Z.J., Liu, H.H.: A unified fuzzy framework for human-hand motion recognition. IEEE Trans. Fuzzy Syst. **19**(5), 901–913 (2011)

32. Chen, W., Xiong, C.H., Yue, S.G.: Mechanical implementation of kinematic synergy for continual grasping generation of anthropomorphic hand. IEEE/ASME Trans. Mechatron. **20**(3), 1249–1263 (2015)
33. Pugeault, N., Bowden, R.: Spelling it out: real-time ASL fingerspelling recognition. In: 1st IEEE Workshop on Consumer Depth Cameras for Computer Vision, pp. 1–6 (2011)

Yang He was born in Hubei province, P.R. China, in 1991. He received B.S. degree in mechanical design manufacture and automation from City College of Wuhan University of Science and Technology, Wuhan, China. He is currently occupied in his M.S. degree in mechanical engineering at Wuhan University of Science and Technology. His current research interests include image processing and intelligent controls.

Gongfa Li received the Ph.D. degree in Wuhan University of Science and Technology, Wuhan, China. He is currently a professor in Wuhan University of Science and Technology. His major research interests are computer aided engineering, mechanical CAD/CAE, modelling and optimal control of complex industrial process.

Yajie Liao was born in Hubei province, P.R. China, in 1991. He received B.S. degree in mechanical engineering and automation from WenHua College, Wuhan, China. He is currently occupied in his M.S. degree in mechanical engineering at Wuhan University of Science and Technology. His current research interests include image processing and intelligent controls.

Ying Sun is currently an associate professor in Wuhan University of Science and Technology. Her major research focuses on teaching research in Mechanical Engineering.

Jianyi Kong received the Ph.D. degree in Helmut Schmidt University, Germany. He is currently a professor in Wuhan University of Science and Technology. His research interests are intelligent machine and controlled mechanism, mechanical and dynamic design and fault diagnosis of electrical system, mechanical CAD/CAE, intelligent design and control.

Guozhang Jiang received the Ph.D. degree in Wuhan University of Science and Technology, China. He is currently a professor in Wuhan University of Science and Technology. His research interests are computer aided engineering, mechanical CAD/CAE and industrial engineering and management system.
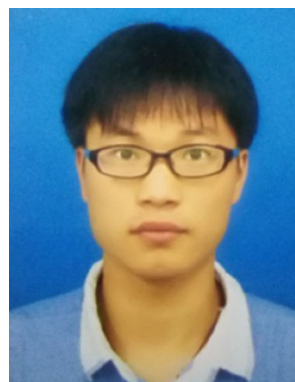
Du Jiang was born in Hubei province, P.R. China, in 1993. He received B.S. degree in mechanical engineering and automation from Wuhan University of Science and Technology, Wu Han, China, in 2017. He is currently occupied in his M.S. degree in mechanical design and theory at Wuhan University of Science and Technology. His current research interests include image processing and intelligent controls.

Bo Tao received the Ph.D. degree in University of Science and Technology of China, China. He is currently a lecturer in Wuhan University of Science and Technology. His research interests are artificial intelligence and sensing technology.

Honghai Liu received the Ph.D. degree in intelligent robotics from Kings College, University of London, London, U.K. He is currently a Professor of intelligent systems in Portsmouth University, Portsmouth, U.K. His research interests are approximate computation, pattern recognition, multi-sensor based information fusion and analytics, human machine systems, advanced control, intelligent robotics and their practical applications.

Shuang Xu received the Ph.D. degree in Huazhong University of Science and Technology, China. She is currently a lecturer in Wuhan University of Science and Technology. Her research interests are artificial intelligence and image processing.