IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 2, 2018, accepted December 11, 2018, date of publication December 18, 2018, date of current version March 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2887223

Dynamic Gesture Recognition in the Internet of Things

GONGFA LI[1,2], (Member, IEEE), HAO WU[1], GUOZHANG JIANG[3,4], SHUANG XU[1,4], AND HONGHAI LIU[5,6], (Senior Member, IEEE)

[1]Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China
[2]Research Center of Biologic Manipulator and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan 430081, China
[3]Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
[4]3D Printing and Intelligent Manufacturing Engineering Institute, Wuhan University of Science and Technology, Wuhan 430081, China
[5]School of Computing, University of Portsmouth, Portsmouth PO1 3HE, U.K.
[6]Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Gongfa Li (ligongfa@wust.edu.cn)

This work was supported in part by the Grants of the National Natural Science Foundation of China under Grant 51575407, Grant 51575338, Grant 51575412, and Grant 61733011, and in part by the Grants of the National Defense Pre-Research Foundation of the Wuhan University of Science and Technology under Grant GF201705.

ABSTRACT Gesture recognition based on computer vision has gradually become a hot research direction in the field of human–computer interaction. The field of human–computer interaction is an important direction in the Internet of Things (IoTs) technology. Human–computer interaction through gestures is the direction of continuous research on IoTs technology. In recent years, the Kinect sensor-based gesture recognition method has been widely used in gesture recognition, because it can separate gestures from complex backgrounds and is less affected by illumination and can accurately track and locate gesture motions. At present, the Kinect sensor needs to be further improved on the recognition of complex gesture movements, especially the problem that the recognition rate of dynamic gestures is not high, which hinders the development of human–computer interaction under the IoTs technology. In this paper, based on the above problems, the Kinect-based gesture recognition is analyzed in detail, and a dynamic gesture recognition method based on HMM and D-S evidence theory is proposed. Based on the original HMM, the tangent angle and gesture change at different moments of the palm trajectory are used as the characteristics of the complex motion gesture, and the dimension of the trajectory tangent is reduced by the number of quantization codes. Then, the parameter model training of HMM is completed. Finally, combined with D-S evidence theory, combinatorial logic is judged, dynamic gesture recognition is carried out, and a better recognition effect is obtained, which lays a good foundation for human–computer interaction under the IoTs technology.

INDEX TERMS Gesture recognition, hidden Markov model (HMM), D-S evidence theory, Internet of Things (IoT).

I. INTRODUCTION

Techniques related to visual gesture recognition include gesture detection segmentation, tracking positioning, feature extraction, classification recognition, etc. These technologies have been greatly developed in recent years, but there are still some imperfections, including: recognition rate, robustness, real-time, stability, practicality, etc. [1]–[3]. At present, due to the uncertainty and diversity of gestures, the research of vision-based gesture recognition has certain challenges Such as inaccurate gesture changes, occlusion, speed of movement, etc. As well as debugging experiments in the gesture recognition system in the real environment, problems such as complex background environment, uncertain lighting conditions, variety of image space information, and computer processing capabilities are often encountered [4], [5]. Currently, the research on gesture recognition technology has become an important research direction in digital image processing, artificial intelligence, computer vision, human interaction and other related fields [6]–[8]. With the continuous development of Internet of Things technology, human-computer interac-

The associate editor coordinating the review of this manuscript and approving it for publication was Chunsheng Zhu.

tion has become more and more convenient, which promote the development of non-contact interaction. The non-contact interaction is to recognize the dynamic gesture of the person, and the machine responds to the recognized dynamic gesture.

In the recognition of static gestures, great achievements have been made, and the classification and recognition of static gestures have achieved good results. However, the recognition of dynamic gestures takes into account the temporal and spatial differences of gestures, so it has been searching for better dynamic gesture recognition methods. This paper improves the recognition rate of dynamic gestures, and then lays a foundation for improving the human-computer interaction experience under the Internet of Things technology. This article first introduces the feature extraction of gestures, and then extends to the extraction of gesture motion trajectories, and next extracts the features of gesture motion trajectories from the extracted motion gestures. The hidden Markov model is a method for dynamic gesture classification. The paper then introduces the concept of hidden Markov model and gesture classification and training with hidden Markov model, and analyzes the HMM model. Finally, a gesture recognition algorithm based on HMM and D-S evidence theory is proposed. Through experiments and results analysis, the HMM algorithm and the method of combining HMM with Naïve Bayesian model are compared.

II. RELATED WORK

Currently, after decades of research on gesture recognition, some breakthrough research progress has been made, and many methods for gesture recognition technology have also been formed. The data glove-based gesture recognition method was extensively studied at the beginning. This method first measures various types of information of the hand through the glove, and then uses the measured information for the modeling and recognition of the gesture. In 1983, Grimes pioneered the use of data gloves for gesture recognition, sparking a wave of gesture recognition research [9]–[11]. After collecting the gesture information, Yang *et al.* [12] proposed a new hand feature extraction method. Xiong *et al.* [13] combined the Back-Propagation Neural Network (BPNN) and (Particle Swarm Optimization, PSO) algorithms on the basis of traditional algorithms, which greatly reduced the training time and improved the accuracy of gesture recognition. The data glove-based gesture recognition method has high recognition rate and high speed [14]–[16], but this method requires artificial wear data gloves and placement of position trackers. In addition, the purchase price of data gloves is high and does not apply to natural human-computer interaction.

The computer-based recognition process uses camera equipment to capture image information or record video information, and uses related image analysis methods to obtain the required gesture map information [17]. The outstanding advantage of this method is that the input device is cheaper, and the camera becomes very popular among various consumer electronic products, besides, the most important thing is that it does not require any additional requirements, which makes the interaction between the computer and the person more natural [18]–[20]. Therefore, the gesture recognition technology based on computer vision has been extensively studied, and the recognition rate, real-time and robustness have been greatly improved.

At present, the most commonly used methods of dynamic gesture classification are Dynamic Time Warping (DTW) and Hidden Markov Model (HMM). HMM have two states: implicit state and observed state, which represent basic state values and observed values [21], [22]. HMM has a constant property on the time scale range, and can be automatically segmented for time series. Zhang and Deng [23] propose dynamic gesture track recognition strategy based on HMM-CART model structure, which has higher accuracy of dynamic gesture recognition. But the data needs to be converted to the same format. Wang *et al.* [24] used Baum-Welch algorithm to train the HMM (Hidden-Markov-Model), which can handle time-based information of different lengths and has high recognition efficiency for dynamic gestures. However, with the development of interactive applications and the improvement of recognition accuracy requirements, a single HMM has been unable to meet the needs of the application, therefore, Wang *et al.* [25] Proposed a HMM-FNN-based model structure that combines the HMM with a fuzzy neural network to establish a more complex gesture by setting up fuzzy rules. Indeed, it has improved the accuracy, but still has the shortcomings such as the dependence of fuzzy rules on prior experience and insufficient convergence rate due to more hidden layer nodes. Barros *et al.* [26] propose a method to recognition dynamic gesture using convexity approach. But the Convexity Approach relies on the contour of the hand, and thus is strongly dependent on a proper segmentation method. Palacios-Alonso *et al.* [27] combined HMM and naive Bayesian model and trained the Hidden Markov Model-Naive Bayesian Classifier (HMM-NBC) model to classify and recognize ten types of gestures. DTW's recognition idea is to find the best match between the two modes, and can automatically seek the optimal planning from the starting point to the end point [28]–[30]. The feature extraction and classification recognition of gestures have been progressing rapidly in recent years, but there are still some shortcomings, mainly in the aspects of low recognition rate, low robustness and poor practicality. At present, due to the repeated changes in gestures, it is inaccurate, will be affected by the occlusion, and even its speed of movement will affect the gesture. When we actually perform gesture recognition simulation, the environment background, different illumination intensity and dynamic information are required to be high, so there is still room for improvement in gesture recognition research [31]–[33].

This paper proposes a dynamic gesture recognition method based on HMM and D-S evidence theory. Since dynamic complex gestures involve temporal and spatial transformations, the motion of the hand can be identified by motion information and gesture changes [34], [35]. Therefore, the tangential angle of the hand-center trajectory and the change of the gesture can be used as the characteristics of the motion gesture. The dimension of the feature vector of the trajectory angle of the motion trajectory is reduced by quantizing the number of codes, and the parameter model of the HMM is trained by the forward and backward algorithm, the Viterbi algorithm and the Baum-Welch algorithm. The D-S evidence theory is used to fuse the parameter models, and then the logic is determined to realize the classification identification.

III. FEATURE EXTRACTION OF GESTURES

The process of feature parameter extraction is a process of abstracting the motion of a gesture in time and space into a number that can be used for calculation. Different gestures correspond to different feature vectors. As the gesture changes, the corresponding feature vector also changes. Therefore, the gesture change and gesture motion are used as feature vectors [36]–[38].

The process of gesture movement can be simplified to the process of changing the position of the palm of the hand over time. The gesture movement can be represented by a trajectory in the parameter space, and the trajectory includes features such as position, velocity and direction [39]–[41].

The distance between the center of the current frame gesture and the starting point is denoted by $S(x, y)$:

$$S(x, y) = \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2} \tag{1}$$

The Euclidean distance through the center of the gesture between two adjacent frames is denoted by $D(x, y)$:

$$D(x, y) = \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2} \tag{2}$$

The angle between the center of the gesture between two adjacent frames is represented by $\theta(x, y)$:

$$\theta(x, y) = \arctan\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right) \tag{3}$$

A. EXTRACTION OF GESTURE TRAJECTORIES

The extraction of the gesture trajectory generally obtains the gesture trajectory by tracking the gesture. In this paper, Kinect's bone tracking technology is used to locate and track the joints of the opponent. For each frame of depth data obtained, the three-dimensional coordinates of the palm are directly obtained to obtain the motion information of the human hand in space [42], [43].

1) DETERMINATION OF THE STARTING POINT AND ENDING POINT OF THE GESTURE TRACK

There will be a pause at the beginning and end of the gesture. Using this feature, the starting and ending points of the gesture track can be judged according to the speed $V(x, y)$ of the gesture, and the gestures at the starting and ending points can be recorded. When the speed of the palm of a certain depth data is lower than the threshold $V_{\min}$, the position of the palm of the frame may be used as the starting and ending point. After the continuous multi-frame, the speed of the palm is lower than $V_{\min}$ again, and the position of the palm of the frame is used as the termination point. The trajectory between the start and end points of the palm and the end point can be extracted [44], [45].

2) GESTURE TRACK SAMPLING

If feature extraction is performed directly on the trajectory data, the amount of data is very large and there is redundant data. Therefore, the motion trajectory needs to be sampled, and the sampled data not only ensures the movement trend of the gesture, but also greatly reduces the amount of data and facilitates feature extraction. At present, generally has two methods to sampling, one is to perform sampling at intervals, and the other is to calculate the moving distance of the palm, and when the moving distance reaches the set threshold, one sampling is performed [46], [47]. Considering that the same gesture is different in speed, the second method is selected. For the motion trajectory, the sampling is performed once every other distance, and the line connecting the points after the sampling indicates the trajectory of the gesture. Fig. 1 shows the raw data and the motion trajectory formed after sampling.

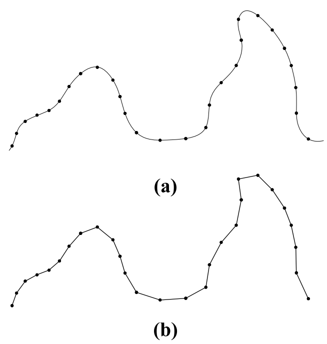


FIGURE 1. Original track and Post-sampled track. (a) Original track. (b) Post-sampling track.

B. OTHER RECOMMENDATIONS FEATURE EXTRACTION OF GESTURE TRAJECTORY

Although the point coordinate position on the gesture trajectory curve can be used as a salient feature of the gesture trajectory recognition, since the coordinate point position of the trajectory changes even with the same gesture, it is not suitable as the feature value. As for the speed of the gesture track, even if they are the same gesture, their speed difference is large, so it is not suitable as a feature value. Different gesture motion trajectories have different tangential angles, and the same gesture motion trajectory has the same or similar tangential angle change. Therefore, this chapter selects the tangential angle change of the hand motion trajectory as the eigenvalue. Since the trajectories drawn by the hand center are composed of sequences of tangential angle changes at different times, the changes are continuous, and the range is $[0, 2\pi]$, if calculated one by one, not only the calculation amount is large, the delay is serious, and the processing effect is not necessarily good. In order to simplify the calculation and increase the calculation speed, it is necessary to quantize the tangent angle [48], [49].

Assuming that the coordinates of the palm of the hand at time t are (x_t, y_t), and the coordinates of the palm of the hand at time $t+1$ are (x_{t+1}, y_{t+1}), the tangential angle $\theta_{t\to t+1}$ of the trajectory of the palm can be expressed as:

$$\theta_{t\to t+1} = \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) \tag{4}$$

According to the obtained angle, uniform quantization coding in the 12 direction is performed, as shown in Fig. 2. This not only simplifies the dimension of the feature vector, but also improves the speed of gesture recognition.

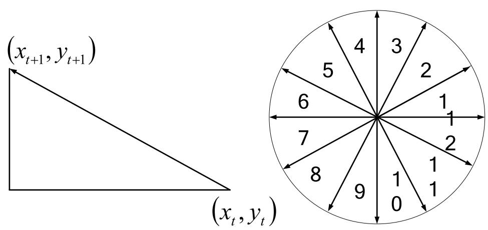


FIGURE 2. Direction chain code extraction feature vector.

IV. MODEL TRAINING

As shown in Fig. 3, the hidden Markov model consists of two parts: the first part is a Markov chain, which is described by the state transition matrix A and the state distribution function π. The second part is a random process, and the matrix B represents the probability distribution description of the state. Hidden Markov models can be represented by quintuple [50]–[52]:

$$\lambda = (N, M, \pi, A, B) \tag{5}$$

where: N—all possible states in the hidden Markov model are $\{s_1, s_2, \ldots, s_t, s_N\}$, The hidden state at time t is q_t, then the sequence of a hidden state is $Q = \{q_1, q_2, \ldots q_T\}$;

M—the number of possible observations in each state is $\{v_1, v_2, \ldots v_M\}$, the observed value at time t is o_t, then an observation sequence is $O = \{o_1, o_2, \ldots o_T\}$;

$\pi -$ the distribution of the initial state, that is $\{\pi_1, \pi_2, \ldots, \pi_N\}$, is used to describe the probability distribution of the state of the observation sequence O at the time T_i, $\pi_i = P(q_i = S_i)$, $1 \le i \le N$, and $\sum_{i=1}^{N} \pi_i = 1$;

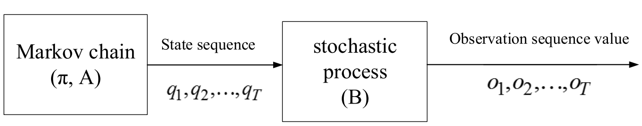


FIGURE 3. Hidden Markov model.

A—The time-independent state transition probability matrix is:$A = \{a_{ij}\}$, $a_{ij} = P\left(q_{t+1} = s_j \middle| q_t = s_i\right)$, $1 \le i, j \le N$, and $\sum_{j=1}^{N} a_{ij} = 1$;

B—state output probability matrix is:

$B = b_j(k) = P\left(o_t = v_k \middle| q_t = s_j\right)$, $1 \le j \le N, 1 \le k \le M$, and $\sum_{k=1}^{M} b_j(k) = 1$.

A. DYNAMIC GESTURE CLASSIFICATION TRAINING

Suppose we know the parameter $\lambda = (\pi, A, B)$ of a hidden Markov model and an observation sequence, and find the probability that this HMM produces this observation sequence [49], [50]. In dynamic gesture recognition, the dynamic gesture is a sequence of observations, that is, a sequence of test samples. The HMM is calculated by matching, and the HMM is the recognized gesture. This chapter uses the forward-backward algorithm to solve it. The steps are as follows:

Forward recursion:

1) Input: hidden Markov model parameter $\lambda = (\pi, A, B)$, observation sequence $(O_1, O_2, \ldots O_t)$;

2) Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1), 1 \le i \le N \tag{6}$$

3) Recursive:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i)\, a_{ij}\right] b_j(O_{t+1}), \quad 1 \le t \le T-1, \quad 1 \le j \le N \tag{7}$$

where: α_{ij}—the probability that the implicit state i in the HMM is transferred to the implicit state j;

B—the probability of observing O_{t+1} under implicit state j.

4) Output: T time is terminated, available

$$P(O/\lambda) = \sum_{j=1}^{N} \alpha_T(j), \quad 1 \le j \le N \tag{8}$$

Backward recursion:

1) Input: hidden Markov model parameter $\lambda = (\pi, A, B)$, observation sequence is $(O_{t+1}, O_{t+2}, \ldots O_T)$;

2) Initialization:

$$\beta_T(i) = 1, \quad 1 \le j \le N \tag{9}$$

3) Advance:

$$\beta_t(j) = \left[\sum_{j=1}^{N} \alpha_{ij} b_j(O_{t+1})\right] \beta_{t+1}(i), \quad 1 \le t \le T-1, \quad 1 \le j \le N \tag{10}$$

4) Output:

$$P(O/\lambda) = \sum_{i=1}^{N} \beta_t(j) \tag{11}$$

B. OPTIMAL DYNAMIC GESTURE CLASSIFICATION TRAINING

An observation sequence is generated for a dynamic gesture, calculate the probability that the sequence of states of all gestures produces the sequence of observations, and the state sequence with the highest probability is found therefrom [55], [56]. It is assumed that the parameter $\lambda = (\pi, \mathrm{A}, \mathrm{B})$ of a hidden Markov model and an observation sequence generated by it are found, and the optimal hidden state sequence most likely to produce this observation sequence is obtained. This chapter uses the Viterbi algorithm to solve the problem. The steps are as follows:

1) Define auxiliary variables:

$$\delta_t(i) = \max_{q1,q2,\ldots qt-1} P(q_1, q_2, \ldots q_{t-1}, q_t = S_i, O_1, O_2, \ldots O_t, \lambda) \tag{12}$$

where: $\delta_t(i)$—in the case of $q_t = s_i$, the maximum probability of sequence $(O_1, O_2, \ldots, O_t)$ is generated at time t.

2) Input: HMM parameter $\lambda = (\pi, A, B)$ and observation sequence is $(O_1, O_2, \ldots, O_t)$;

3) Initialization:

$$\delta_1(i) = \pi_i b_i(O_i), \quad 1 \le i \le N \tag{13}$$

$$\varphi_i = 0, \quad 1 \le i \le N \tag{14}$$

4) Recursive:

$$\delta_t(j) = \max_{1\le i\le N}\left[\delta_{t-1}(i)\, a_{ij}\right] b_j O_t, \quad 2 \le t \le T, \quad 1 \le j \le N \tag{15}$$

$$\varphi_t(j) = \max_{1\le i\le N}\left[\delta_{t-1}(i)\, a_{ij}\right], \quad 2 \le t \le T, \quad 1 \le j \le N \tag{16}$$

5) Termination:

$$P^* = \max_{1\le i\le N}\left[\delta_T(i)\right] \tag{17}$$

$$Q_T^* = \max_{1\le i\le N}\left[\delta_T(i)\right] \tag{18}$$

6) Output status sequence:

$$q_t^* = \varphi_{t+1}\left(q_{t+1}^*\right), \quad 2 \le t \le T \tag{19}$$

C. DYNAMIC GESTURE PARAMETER MODEL TRAINING

Assuming that an observation sequence is known, the parameter $\lambda = (\pi, \mathrm{A}, \mathrm{B})$ of the HMM is required to make the probability $P\left(O/\lambda\right)$ maximum [57], [58]. In the dynamic gesture recognition, the gesture image needs to be collected. After processing, the sample sequence is obtained, and the sequence is stored corresponding to the gesture name, and then trained to determine the parameter $\lambda = (\pi, \mathrm{A}, \mathrm{B})$ model. When the parameters are determined, the model can be used for subsequent identification work. This chapter uses the Baum-Welch algorithm to solve the problem.

Definition $\varepsilon_t(i,j)$ is the probability that the sample sequence O and the model λ are in the S_i state at time t and the $t+1$ time is the S_j state.

$$\varepsilon_t(i,j) = \frac{\alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{i+1}(j)}{P\left(O/\lambda\right)} \tag{20}$$

It is obtained by the algorithm of the front and the back:

$$\varepsilon_t(i,j) = \frac{\alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{i+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{i+1}(j)} \tag{21}$$

At time t, the probability that the Markov chain is in the S_i state is:

$$\gamma(i) = \sum_{j=1}^{N} \varepsilon_t(i,j) = \frac{[\alpha_t(i)\, \beta_t(i)]}{P\left(O/\lambda\right)} \tag{22}$$

where: $\sum_{1\le i\le T-1} \varepsilon_t(i)$—the expected number of S_i state transitions out;

$\sum_{1\le i\le T-1} \varepsilon_t(i,j)$ – the expected number of times the S_i state transitions to the S_j state.

The estimated value of the new model is:

$$\overrightarrow{a_{ij}} = \frac{\sum_{j=1}^{N} \varepsilon_T(i,j)}{\sum_{t=1}^{T} \gamma(i)} \tag{23}$$

$$\overrightarrow{b_{jk}} = \frac{\sum_{t=1,O_t=k}^{T} \varepsilon_t(i,j)}{\sum_{t=1}^{T} \varepsilon_t(j)} \tag{24}$$

$$\vec{\pi}_i = \gamma(i) \tag{25}$$

Through the above formula, the model parameter $\lambda = (\pi, A, B)$ is re-evaluated, and the new model parameter $\vec{\lambda} = \left(\vec{\pi}, \vec{A}, \vec{B}\right)$ can be obtained by calculation, and then the process is repeated. To obtain a converged model parameter, a threshold δ is defined such that when $\log P\left(O/\lambda\right) - \log P\left(O/\lambda_0\right) < \delta$, the resulting model $\vec{\lambda} = \left(\vec{\pi}, \vec{A}, \vec{B}\right)$ is a trained HMM.

We choose left-right banded model Fig. 4 as the HMM topology, because the left-right banded model is good for modeling-order-constrained time-series whose properties sequentially change over time 26. Since the model has no backward path, the state index either increases or stays unchanged as time increases. After finishing the training process by computing the HMM parameters for each type of gesture, a given gesture is recognized corresponding to the maximal likelihood of ten HMM models by using Viterbi algorithm.

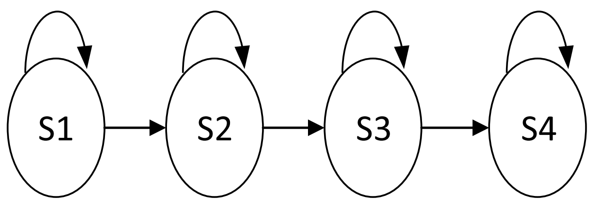


FIGURE 4. HMM topology.

The training process of the HMM is the process of continuously updating the parameter λ. The training result determines the recognition effect of the model on the gesture. As shown in Fig. 5, it is the training flowchart of the dynamic gesture HMM. For a gesture to be trained, a part of the data is extracted from the gesture training sample database, and the gesture change feature observation sequence and the palm

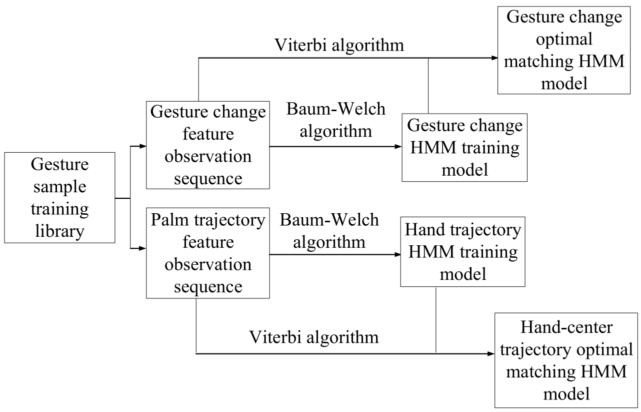


FIGURE 5. Training flow chart of dynamic gesture HMM.

motion trajectory observation sequence are obtained respectively. Then the HMM model of gesture change and the HMM model of palm motion trajectory are respectively trained by using the Baum-Welch algorithm. When the threshold δ condition is satisfied, the training of the gesture change HMM model and the palm motion trajectory HMM model is completed, and the above process is repeated to complete all gesture model training. The remaining part of the gesture data is extracted from the gesture training sample, and then input into the HMM model of gesture change and the palm motion trajectory model which have been trained. And the Viterbi algorithm is used to find the HMM model that best matches the input sample for subsequent classification.

V. GESTURE RECOGNITION ALGORITHM BASED ON HMM AND D-S EVIDENCE THEORY

In this paper, HMM and D-S evidence theory are combined to construct a fusion feature model for dynamic gesture recognition. The main idea is to use the D-S evidence theory to feature the optimal matching HMM model obtained by Baum-Welch algorithm and Viterbi algorithm after the gesture change feature observation sequence and the hand-heart trajectory feature observation sequence are trained. Then the logic is judged according to the D-S evidence theory combination rule, and finally the fusion model is classified, the flow of dynamic gesture recognition is shown in Fig. 6.

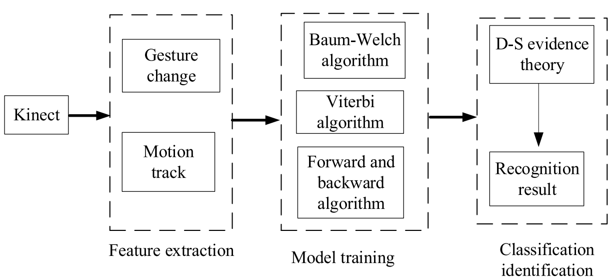


FIGURE 6. Flow chart of Dynamic gesture recognition.

In recent years, the research on D-S evidence theory has been rapidly developed. As the improvement and optimization of Bayesian theory, D-S evidence theory has a better description language when describing things, that is, it can more accurately study things. The D-S evidence theory was first proposed in the late 20th century by A.P. Dempster, and his students further improved the method to make it possible to perform data fusion analysis under limited data and uncertain conditions. D-S evidence theory can be applied to more weak data processing, which can qualitatively analyze the uncertainty of things and meet the needs of actual production and life. Through the information fusion of the evidence on the recognition framework, qualitative analysis is carried out to complete the judgment of things.

According to the traditional D-S evidence theory, the optimal matching HMM model of the gesture change and the optimal matching HMM model of the palm trajectory change are studied. The basic contents are as follows:

A. IDENTIFICATION FRAMEWORK

The D-S evidence theory discusses a "Frame of Discernment" θ, so that possible independent identification results or assumptions about the proposition are defined within this framework. The set of all possible subsets contained in θ is called the power set of θ and is represented by the symbol $\Omega(\theta)$. For example, in complex dynamic gesture recognition, assuming that the category of the sample to be identified may be x, y, z, then in this case, the "identification frame" and "power set" are defined as follows:

$$\theta = \{x, y, z\} \tag{26}$$

$$\Omega(\theta) = \{\phi, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, \}z\}\} \tag{27}$$

The ϕ empty set indicates that the recognition result is not any result of x, y, z, and may be other categories,$\{x, y\}$ indicates that the result may be x or y, and other subsets are similar. It can be seen that when θ contains N elements, $\Omega(\theta)$ has $2^N - 1$ elements.

B. BASIC CREDIBILITY DISTRIBUTION FUNCTION

For the defined "recognition framework" θ, the power set function m is defined as a mapping of the power set $\Omega(\theta)$ to [0,1], that is, the following two conditions are met:

$$\mathrm{m} : \Omega(\theta) \rightarrow [0, 1] \tag{28}$$

$$\mathrm{m}(\phi) \neq 0, \quad \sum_{A \subseteq \Omega(\theta)} m(A) = 1 \tag{29}$$

m is the basic probability assignment function of the proposition. When $A = \theta$, $\mathrm{m}(A)$ indicates that m does not know how to allocate. When A is a subset of θ and $\mathrm{m}(A) \neq 0$, A is called a Focal Function of m.

C. TRUST FUNCTION AND LIKELIHOOD FUNCTION

Let θ be the "recognition framework" and $m : \Omega(\theta) \rightarrow |0, 1|$ be the basic credibility assignment function on framework θ.

The trust function Bel is defined as:

$$Bel : \Omega(\theta) \to [0, 1], \quad Bel(A) = \sum_{B \subseteq A} m(B) \tag{30}$$

The likelihood function Pls is:

$$Pls : \Omega(\theta) \to [0, 1], Pls(A) = 1 - Bel(\bar{A}) = \sum_{B \cap A \neq \phi} m(B) \tag{31}$$

The trust function $Bel(A)$ represents a measure in which the proposition A determines to be established, and the likelihood function $Pls(A)$ represents an uncertainty measure that the proposition A seems to be possible.

D. TRUST UNCERTAINTY INTERVAL

Knowing the trust function $Bel(A)$ of a proposition A and the likelihood function is $Pls(A)$, then the uncertainty trust interval of A is $[Bel(A), Pls(A)]$. As shown in FIG. 5, the two ends are the intervals of trust A and distrust A, and the middle region represents the uncertainty probability of proposition A. The lower bound is the confidence degree of propositions, which represents the minimum probability of propositions based on the direct evidence of sensors; the upper bound is the likelihood degree of propositions, which indicates the trust degree of propositions and the potential probability of propositions. Therefore, this boundary can indicate how much of a certain evidence is determined to support a certain proposition, what proportion is the lack of understanding of the evidence, and what proportion is determined to refute a certain proposition. The uncertainty interval is shown in Fig. 7.

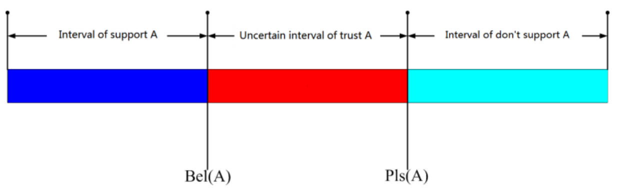


FIGURE 7. Trust uncertainty interval.

E. D-S EVIDENCE COMBINATION RULE

Suppose Bel_1, Bel_2 is the trust function from two different types of feature data under the same "identification framework" θ, and m_1, m_2 is its corresponding basic credibility assignment function. Let the focal elements of Bel_1, Bel_2 be $A_1, A_2, \ldots A_K$ and $B_1, B_2, \ldots B_K$ respectively, then according to the D-S orthogonal principle, the synthesized basic credibility distribution function $m : \Omega(\theta) \to [0, 1]$ is:

$$m(c) = \begin{cases} 0, & C \neq \theta \\ \dfrac{\sum_{A_i \cap B_j = C} m_1(A_i)\, m_2(B_j)}{1 - K}, & C \neq \phi \end{cases} \tag{32}$$

$$K = \sum_{A i \cap B j = \phi} m_1(A_i)\, m_2(B_j) \tag{33}$$

According to the D-S evidence theory combination rule, the characteristic data of the two types of HMM optimal matching models are merged, and then the feature fusion results are logically analyzed to obtain the classification results. The acquired gesture change characteristics and the palm motion trajectory characteristics are trained by the model, and the HMM features of the gesture change and the HMM feature of the palm-track trajectory are obtained respectively. Then the evidence of the two types of optimal matching HMM features is obtained by DS evidence theory interval. The feature fusion is performed by the D-S combination rule, and the logic analysis is performed according to the result, and finally the gesture classification result is obtained. The specific identification flow chart is shown in Fig. 8.

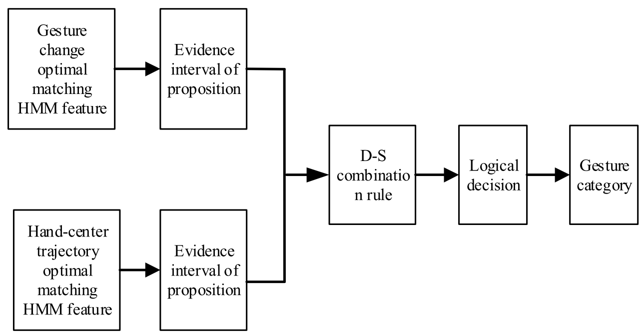


FIGURE 8. Identification flow chart based on D-S evidence theory.

VI. EXPERIMENT AND RESULT ANALYSIS OF DYNAMIC GESTURE RECOGNITION

A. DEFINITION OF COMBINED DYNAMIC GESTURES

This paper divides a complete dynamic gesture into gesture changes and changes in gesture trajectories. This paper uses

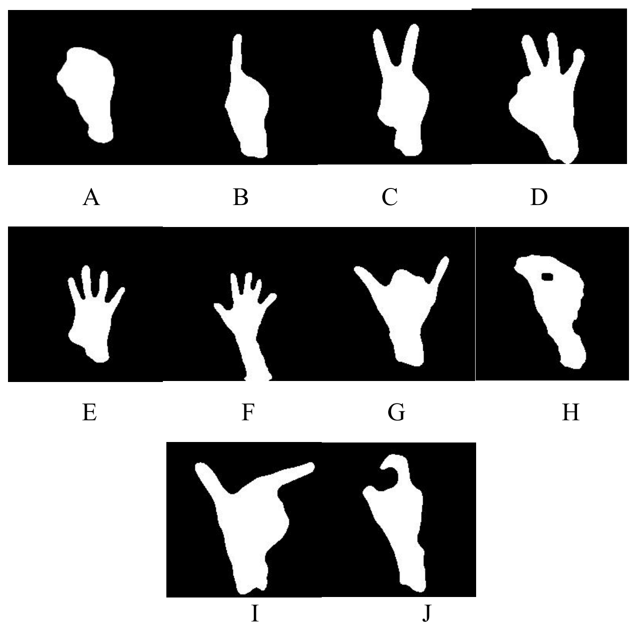


FIGURE 9. Ten static gesture binary graph.

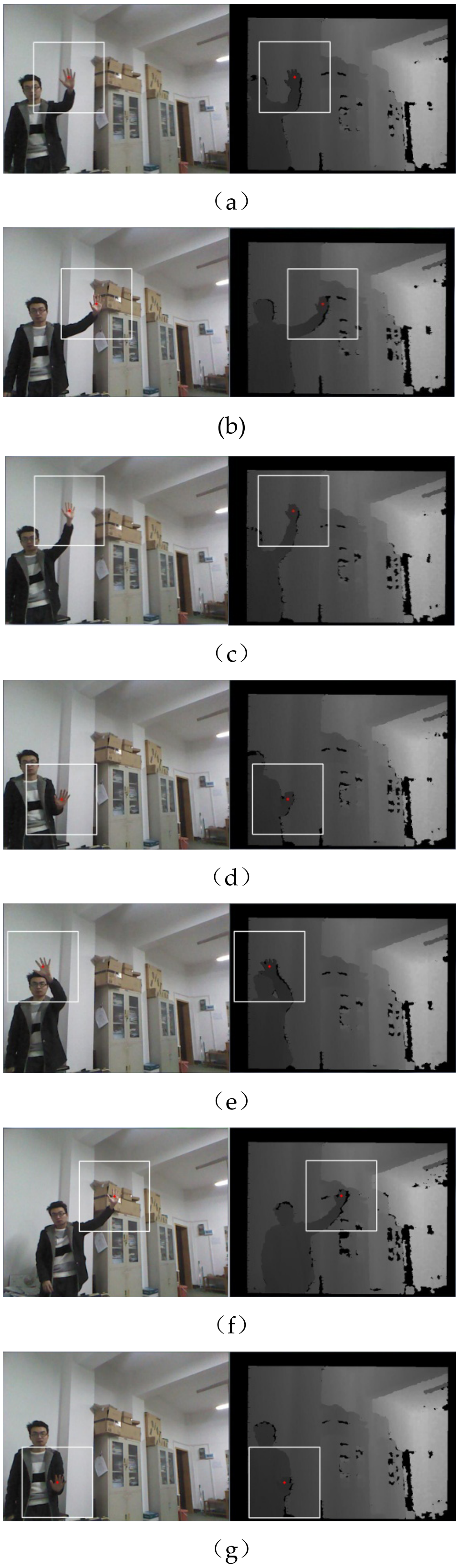


FIGURE 10. Predefined ten types of gesture tracks. (a) Left. (b) Right. (c) Up. (d) Down. (e) Top left. (f) Upper right. (g) Bottom left. (h) Bottom right. (i) Clockwise circle. (j) Counterclockwise circle.

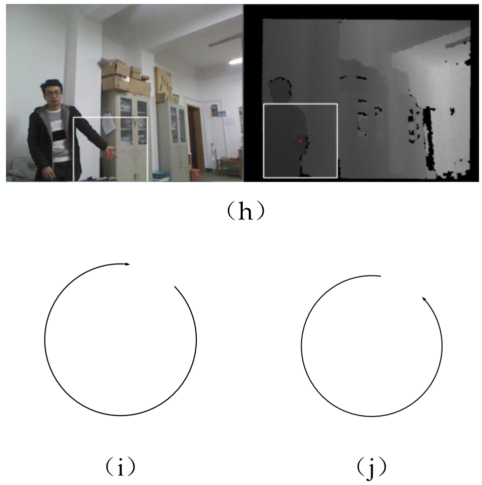


FIGURE 10. *(Continued..)* Predefined ten types of gesture tracks. (a) Left. (b) Right. (c) Up. (d) Down. (e) Top left. (f) Upper right. (g) Bottom left. (h) Bottom right. (i) Clockwise circle. (j) Counterclockwise circle.

TABLE 1. Ten types of custom combination gestures.

Combined gesture type	Gesture change	Motion track
Gesture 1	A-B	Left
Gesture 2	B-C	Right
Gesture 3	C-D	Up
Gesture 4	D-E	Down
Gesture 5	E-F	Clockwise circle
Gesture 6	F-G	Counterclockwise circle
Gesture 7	G-H	To the top left
Gesture 8	H-I	To the lower left
Gesture 9	I-J	To the top right
Gesture 10	J-A	To the bottom right

Kinect to collect ten types of gesture samples from different people, which means the numbers 0-9. For the convenience of subsequent experiments, the ten digits 0-9 are represented by A-J. For each type of gesture, 100 samples are selected, 50 samples are used as test samples, and 50 samples are used as training samples. For A-J ten-type gestures, 1000 samples are required. For gesture changes, the ten types of static gestures collected have been represented by the numbers A-J, as shown in Fig. 9. For gesture motion trajectories, this section defines ten types of motion trajectories such as up, down, left, right, clockwise, inverse time, top left, bottom left, right up, and right down, as shown in Fig. 10. The defined gestures and gesture motion trajectories are combined to form a combined gesture. As shown in Table 1, this paper customizes a combination of 10 different static gestures and motion trajectories.

B. ESTABLISHMENT OF DYNAMIC GESTURE SAMPLE LIBRARY

The Kinect used in the experiment has a maximum frame rate of 30FPS and a resolution of 640∗480. It collects gesture trajectories of 10 people, and each gesture requires 100 samples. Among them, 50 gesture trajectories are used as training samples, and 50 gesture trajectories are left as test samples. Since ten gesture trajectories are defined, 1000 samples are needed. This article uses Kinect Studio's Kinect Studio tool to record gesture video. The running interface is shown in Fig. 11. When recording gesture tracks, the Kinect Studio tool can save multiple data sources such as color, depth and bones to a file at the same time. During playback, the original data of various recorded data sources can be obtained directly from the saved file. For recorded video, data analysis can be performed on different frames of the video selectively, and the video clip of each sample is intercepted by setting a breakpoint. Therefore, in order to collect video clips more conveniently, a video is recorded for each gesture, and the gesture is repeated 100 times in this video, so that 100 data samples can be obtained. For 10 gestures, only 10 videos can be recorded to capture the raw data of 1000 gesture samples.

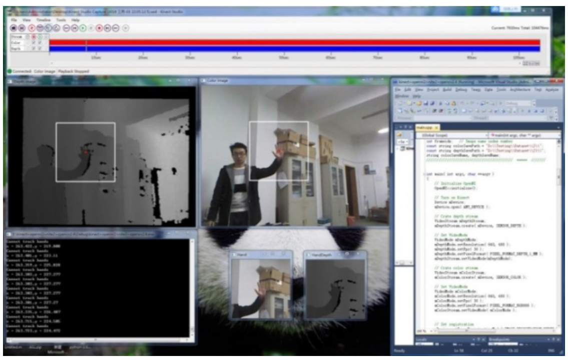


FIGURE 11. Work interface of kinect studio.

C. EXPERIMENTAL RESULTS AND ANALYSIS OF DYNAMIC GESTURE RECOGNITION

1) RECOGNITION EFFECT OF THE METHOD IN THIS PAPER

In the dynamic recognition of gestures, because the recognition rate is not high, the gesture recognition method based on HMM model and D-S evidence theory is proposed. On the basis of the original HMM, the tangential angle and gesture change of the palm trajectory at different moments are extracted as features of the complex motion gesture. In order to reduce the amount of calculation and the running time, the tangential angle of the trajectory is reduced by the quantization code to reduce the dimension. The HMM model is trained and trained by the HMM model, and the optimal matching HMM model and the trajectory optimal matching HMM model are obtained. Then combined with the D-S evidence theory, the two optimal feature models are merged, and then the logic is determined, and then the dynamic gesture is identified. Each type of gesture collects 100 sample data, 50 samples are used for training samples, and the remaining 50 are used for test samples. Each type of gesture is tested. The test results are shown in Table 2.

TABLE 2. Gesture recognition results.

Gesture category	Testing frequency	Correct number of times	Recognition rate
Gesture 1	50	47	94%
Gesture 2	50	48	96%
Gesture 3	50	48	96%
Gesture 4	50	47	94%
Gesture 5	50	46	92%
Gesture 6	50	46	92%
Gesture 7	50	44	88%
Gesture 8	50	43	86%
Gesture 9	50	45	90%
Gesture 10	50	44	88%

It can be seen from Table 2 that the average correct recognition rate of dynamic gestures based on HMM and D-S evidence theory is 91.6%, which is better for complex combined gesture classification and recognition. The recognition rate of gesture 1 to gesture 6 is above 90%, and the recognition rate of gesture 7 to gesture 10 is low, between 85% and 90%. The highest recognition rate is gesture 2 and gesture 3, the recognition rate reaches 96%, and the gesture with the lowest recognition rate is gesture 8, and the recognition rate is 86%. Due to the convenience of the right hand, the upward and rightward gestures have little change in the amplitude of each motion trajectory, and the amplitude of the gesture change is not large, so the recognition rates of gesture 1 and gesture 4 are slightly lower than gesture 2 and gesture 3. Since the tangential angle of the gesture trajectory changes greatly during the trajectory movement of the gesture 5 to the gesture 10, it is easy to cause the gesture plane to generate an excessive angle with the Kinect plane. In the process of gesture change, the feature extraction is affected by the angle of the gesture plane and the Kinect plane, and the gesture change characteristics will change, which affects the recognition rate. In the gesture 8 with the lowest recognition rate, the motion trajectory is to the lower left. Due to the convenience of the right hand, when the gesture trajectory moves to the lower left, the fluctuation of the palm trajectory is large, and the gesture trajectory is misidentified as gesture 1 or gesture 4 . The most influential on the recognition rate is the fluctuation of the palm trajectory. Of course, the gesture plane and the Kinect plane angle also have an influence.

2) COMPARISON OF METHODS IN THIS PAPER WITH OTHER SIMILAR METHODS

In [27], the naive Bayesian classifier was introduced on the basis of the traditional HMM to train the HMM-NBC model to perform dynamic recognition of combined gestures. In [26], a method for recognizing dynamic gestures using the convexity approach. This method uses hand contours as input and is used together with other feature extraction techniques to reference HMM to train all gesture models and perform partial recognition. In [25], based on the traditional

TABLE 3. Comparison of the method in this paper with other similar methods.

Recognition methods	Maximum recognition rate（%）	Minimum recognition rate（%）	Average recognition rate（%）
Literature [24]	96	82	88.4
Literature [23]	90	80	88.3
Method of this paper	96	86	91.6

HMM, the fuzzy neural network was introduced to train the HMM-FNN model, and the dynamic gesture classification and recognition were carried out in the collision type, motion trajectory and axial direction motion respectively. The methods of the above literature were tested and compared in the environment of this experiment. As shown in Table 3, the comparison results of the maximum recognition rate, the minimum recognition rate and the average recognition rate are given. It can be seen that the dynamic gesture recognition method based on the fusion of HMM and D-S evidence theory in this experiment greatly improves the problem of low dynamic gesture recognition rate. Although the recognition rate of some complicated gesture combinations is low, the method of this experiment is optimal for the recognition of these three methods.

VII. CONCLUSION

Firstly, the thesis studies the dynamic gesture motion, and uses the gesture change and motion trajectory as all the features of a complex gesture to extract the tangential angle of the gesture change feature and the palm motion trajectory. The tangent angle of the motion path of the palm is used as the feature vector, and the encoding and quantizing processing in the 12 direction is performed, and the calculation speed is improved. Then the Forward-backward algorithm is used to solve the classification identification problem. The Viterbi algorithm solves the category of the optimal trajectory sequence and the Baum-Welch algorithm to train the parameter model of the HMM. Then, based on the traditional HMM, the gesture change feature and the hand motion trajectory feature model are trained respectively, and the HMM model and the trajectory optimal matching HMM model are obtained. Then use D-S evidence theory to carry out model feature fusion to realize dynamic gesture recognition. By training the defined ten types of gestures, and then using the experimental method to test the test samples, the results of gesture classification are obtained. The results show that the experimental method has achieved good results for dynamic gesture classification and recognition.

The dynamic gesture recognition based on HMM and D-S evidence theory proposed in this paper achieves better recognition effect and has important significance for human-computer interaction under the Internet of Things technology. With the continuous development of Internet of Things technology, non-contact human-computer interaction has become a hot spot. In the fields of smart cars, smart homes, robots, etc., the camera is equipped to capture the movement of the human hand, recognize the gesture, and complete the interaction between the machine and the person. The non-contact human-computer interaction of gesture recognition is a very meaningful study for the use of language barriers or certain occasions. This paper aims to improve the recognition rate of dynamic gestures, lay a foundation for improving human-computer interaction of gesture recognition, and also aims to promote the development of Internet of Things technology.

REFERENCES

[1] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artif. Intell. Rev.*, vol. 43, no. 1, pp. 1–54, Jan. 2015.

[2] G. F. Li, L. Zheng, Y. Sun, and J. Kong, "Towards the sEMG hand: Internet of things sensors and haptic feedback application," *Multimedia Tools Appl.*, to be published. doi: 10.1007/s11042-018-6293-x.

[3] E. Ohn-Bar and M. M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2368–2377, Dec. 2014.

[4] S. P. Priyal and P. K. Bora, "A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments," *Pattern Recognit.*, vol. 46, no. 8, pp. 2202–2219, Aug. 2013.

[5] D. Jiang *et al.*, "Gesture recognition based on binocular vision," *Cluster Comput.*, to be published. doi: 10.1007/s10586-018-1844-5.

[6] Z. J. Ju *et al.*, "An integrative framework of human hand gesture segmentation for human-robot interaction," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1326–1336, Sep. 2015.

[7] Y. Zhou, Y. Fang, K. Li, D. Zhang, H. Li, and K. Li, "sEMG bias-driven functional electrical stimulation system for upper-limb stroke rehabilitation," *IEEE Sensors J.*, vol. 18, no. 16, pp. 6812–6821, Aug. 2018.

[8] K. R. Nirmal and N. Mishra, "3D graphical user interface on personal computer using P5 data glove," *Int. J. Comput. Sci.* , vol. 8, no. 5, pp. 155–160, Sep. 2011.

[9] G. F. Li *et al.*, "Hand gesture recognition based on convolution neural network," *Cluster Comput.*, to be published. doi: 10.1007/s10586-017-1435-x.

[10] Y. F. Fang, D. Zhou, and H. Liu, "Interface prostheses with classifier-feedback-based user training," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2575–2583, Nov. 2016.

[11] Y. He *et al.*, "Gesture recognition based on an improved local sparse representation classification algorithm," *Cluster Comput.*, to be published. doi: 10.1007/s10586-017-1237-1.

[12] X. C. Yang, X. Sun, D. Zhou, Y. Li, and H. Liu, "Towards wearable A-mode ultrasound sensing for real-time finger motion recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 6, pp. 1199–1208, Jun. 2018.

[13] L. Xiong, G. Jiang, Y. Guo, and H. Liu, "A three-dimensional fiber Bragg grating force sensor for robot," *IEEE Sensors J.*, vol. 18, no. 9, pp. 3632–3639, May 2018.

[14] B. Li *et al.*, "Gesture recognition based on modified adaptive orthogonal matching pursuit algorithm," *Cluster Comput.*, to be published. doi: 10.1007/s10586-017-1231-7.

[15] X. Q. Cai, T. H. Guo, X. Wu, and H. Y. Sun, "Gesture recognition method based on wireless data glove with sensors," *Sensor Lett.*, vol. 13, no. 2, pp. 134–137, Feb. 2015.

[16] J. Gałka, M. Mąsior, M. Zaborski, and K. Barczewska, "Inertial motion sensing glove for sign language gesture acquisition and recognition," *IEEE Sensors J.*, vol. 16, no. 16, pp. 6310–6316, Aug. 2016.

[17] O. Luzanin and M. Plancak, "Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network," *Assembly Automat.*, vol. 34, no. 1, pp. 94–105, 2014.

[18] D. S. Chen *et al.*, "An interactive image segmentation method in hand gesture recognition," *Sensors*, vol. 17, no. 2, p. 253, Jan. 2017. doi: 10.3390/s17020253.

[19] Z. Lu, X. Chen, Q. Li, X. Zhang, and P. Zhou, "A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 2, pp. 293–299, Apr. 2014.

[20] Y. Sun *et al.*, "Gesture recognition based on Kinect and sEMG signal fusion," *Mobile Netw. Appl.*, vol. 23, no. 4, pp. 797–805, Aug. 2018.

[21] K. Liu, C. Chen, R. Jafari, and N. Kehtarnavaz, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors J.*, vol. 14, no. 6, pp. 1898–1903, Jun. 2014.

[22] Y. J. Liao *et al.*, "Simultaneous calibration: a joint optimization approach for multiple Kinect and external cameras," *Sensors*, vol. 17, no. 7, p. 1491, 2017. doi: 10.3390/s17071491.

[23] Q. X. Zhang and F. Deng, "Dynamic gesture recognition based on leapmotion and HMM-CART model," *J. Phys. Conf. Ser.*, vol. 910, no. 1, Oct. 2017, Art. no. 012037. doi: 10.1088/1742-6596/910/1/012037.

[24] X. Wang, M. Xia, H. Cai, Y. Gao, and C. Cattani, "Hidden-Markov-models-based dynamic hand gesture recognition," *Math. Problems Eng.*, vol. 2012, pp. 137–149, Jan. 2012. doi: 10.1155/2012/986134.

[25] X. Y. Wang *et al.*, "Recognition of complex dynamic gesture based on HMM-FNN model," *J. Softw.*, vol. 19, no. 9, pp. 2302–2312, Sep. 2008.

[26] P. Barros *et al.*, "A dynamic gesture recognition and prediction system using the convexity approach," *Comput. Vis. Image Understand.*, vol. 155, pp. 139–149, Feb. 2017.

[27] M. A. Palacios-Alonso, C. A. Brizuela, and L. E. Sucar, "Evolutionary learning of dynamic naive Bayesian classifiers," *J. Automated Reasoning*, vol. 45, no. 1, pp. 21–37, Jun. 2010.

[28] Y. F. Fang, H. Liu, G. Li, and X. Zhu, "A multichannel surface EMG system for hand motion recognition," *Int. J. Humanoid Robot.*, vol. 12, no. 2, 2015, Art. no. 1550011. doi: 10.1142/S0219843615500115.

[29] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," *Pattern Recognit.*, vol. 43, no. 3, pp. 706–719, 2010.

[30] W. C. Guo, X. Sheng, H. Liu, and X. Zhu, "Toward an enhanced human-machine interface for upper-limb prosthesis control with combined EMG and NIRS signals," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 4, pp. 564–575, Aug. 2017.

[31] W. Miao *et al.*, "Optimal grasp planning of multi-fingered robotic hands: A review," *Appl. Comput. Math.*, vol. 14, no. 3, pp. 238–247, Sep. 2016.

[32] G. F. Li and K. Veer, "A novel feature extraction method for machine learning based on surface electromyography from healthy brain," *Neural Comput. Appl.*, vol. 18, no. 3, p. 869, Mar. 2018. doi: 10.1007/s00521-018-3887-1.

[33] Y. Kim and B. Toomajian, "Hand gesture recognition using micro-Doppler signatures with convolutional neural network," *IEEE Access*, vol. 4, pp. 7125–7130, 2016.

[34] F. Du *et al.*, "Adaptive fuzzy sliding mode control algorithm simulation for 2-DOF articulated robot," *Int. J. Wireless. Mobile Comput.*, vol. 13, no. 4, pp. 306–313, 2017.

[35] W. L. Ding *et al.*, "D-S evidential theory on sEMG signal recognition," *Int. J. Comput. Sci. Math.*, vol. 8, no. 2, pp. 138–145, Aug. 2017.

[36] Y. Sang, L. Shi, and Y. Liu, "Micro hand gesture recognition system using ultrasonic active sensing," *IEEE Access.*, vol. 6, pp. 49339–49347, 2018.

[37] B. W. Luo *et al.*, "Decomposition algorithm for depth image of human health posture based on brain Health," *Neural Comput. Appl.*, to be published. doi: 10.1007/s00521-018-3883-5.

[38] M. Grzegorczyk, "A non-homogeneous dynamic Bayesian network with a hidden Markov model dependency structure among the temporal data points," *Mach. Learn.*, vol. 102, no. 2, pp. 155–207, Feb. 2016.

[39] Y. J. Huang, X. Yang, Y. Li, D. Zhou, K. He, and H. Liu, "Ultrasound-based sensing models for finger motion classification," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 5, pp. 1395–1405, Oct. 2017.

[40] Q. J. Lv, Y. Qiao, N. Ansari, J. Liu, and J. Yang, "Big data driven hidden Markov model based individual mobility prediction at points of Interest," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5204–5216, Jun. 2017.

[41] Y. Yao and Y. Fu, "Contour model-based hand-gesture recognition using the Kinect sensor," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 11, pp. 1935–1944, Nov. 2014.

[42] S. Bilal, R. Akmeliawati, A. A. Shafie, and M. J. E. Salami, "Hidden Markov model for human to computer interaction: A study on human hand gesture recognition," *Artif. Intell. Rev.*, vol. 40, no. 4, pp. 495–516, Dec. 2013.

[43] W. Miao, C. Li, Y. Sun, G. Jiang, J. Kong, and H. Liu, "Gesture recognition based on sparse representation," *Int. J. Wireless. Mobile Comput.*, vol. 11, no. 4, pp. 348–356, Apr. 2016.

[44] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, "Gesture recognition based on skeletonization algorithm and CNN with ASL database," *Multimedia Tools Appl.*, to be published. doi: 10.1007/s11042-018-6748-0.

[45] M. Hu, X. Bu, X. Sun, Z. Yu, and Y. Zheng, "Rape plant disease recognition method of multi-feature fusion based on D-S evidence theory," *Mathematical. Comput. Appl.*, vol. 22, no. 1, p. 18, 2017. doi: 10.3390/mca22010018.

[46] W. T. Cheng, Y. Sun, G. Li, G. Jiang, and H. Liu, "Jointly network: A network based on CNN and RBM for gesture recognition," *Neural Comput. Appl.*, to be published. doi: 10.1007/s00521-018-3775-8.

[47] D. S. Chen, G. Li, G. Jiang, and Y. Fang, "Intelligent computational control of multi-fingered dexterous robotic hand," *J. Comput. Theor. Nanosci.*, vol. 12, no. 12, pp. 6126–6132, Dec. 2015.

[48] S. Lee, J. Lee, M. H. Hayes, and J. Paik, "Adaptive background generation for automatic detection of initial object region in multiple color-filter aperture camera-based surveillance system," *IEEE Trans. Consum. Electron.*, vol. 58, no. 1, pp. 104–110, Feb. 2012.

[49] D. Ramírez-Giraldo, S. Molina-Giraldo, A. M. Álvarez-Meza, G. Daza-Santacolo, and G. Castellanos-Domínguez, "Kernel based hand gesture recognition using kinect sensor," in *Proc. Symp. Image, Signal Process. Artif. Vis.*, Sep. 2012, pp. 158–161. doi: 10.1109/STSIVA.2012.6340575.

[50] W. L. Ding *et al.*, "Intelligent computation in grasping control of dexterous robot hand," *J. Comput. Theor. Nanosci.*, vol. 12, no. 12, pp. 6096–6099, Dec. 2015.

[51] X. L. Sun, X. Yang, X. Zhu, and H. Liu, "Dual-frequency ultrasound transducers for morphological changes of deep-layered muscles," *IEEE Sensors J.*, vol. 18, no. 4, pp. 1373–1383, Feb. 2017.

[52] C. K. Hsia, Y. J. Liou, and J. S. Chiang, "Directional prediction camshift algorithm based on adaptive search pattern for moving object tracking," *J. Real-Time Image Process.*, vol. 12, no. 1, pp. 183–195, Jun. 2016.

[53] J. X. Qi *et al.*, "Surface EMG Hand gesture recognition system based on PCA and GRNN," *Neural Comput. Appl.*, to be published. doi: 10.1007/s00521-018-3885-3.

[54] Z. Li *et al.*, "Computation of grasping and manipulation for multi-fingered robotic hands," *J. Comput. Theor. Nanosci.*, vol. 12, no. 12, pp. 6192–6297, Dec. 2015.

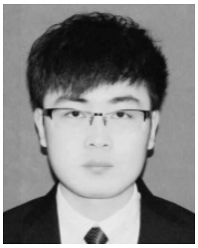
[55] H. H. Liu, "Exploring human hand capabilities into embedded multi-fingered object manipulation," *IEEE Trans. Ind. Inform.*, vol. 7, no. 3, pp. 389–398, Aug. 2011.

[56] R. X. Hu *et al.*, "Hand shape recognition based on coherent distance shape contexts," *Pattern Recognit.*, vol. 45, no. 9, pp. 3348–3359, Sep. 2012.

[57] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using Kinect sensor," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1110–1120, Aug. 2013.

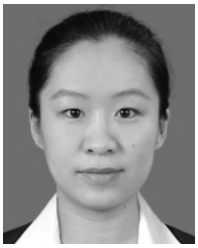
[58] G. Zhu, "Multimodal gesture recognition using 3-D convolution and convolutional LSTM," *IEEE Access*, vol. 5, pp. 4517–4524, 2017.

GONGFA LI received the Ph.D. degree. He is currently a Professor, the Doctoral Supervisor, the Deputy Director of the Key Laboratory of Metallurgical Equipment and Control Ministry of Education, the Director of the Biomechanics and Intelligent Measurement and Control Research Center of Wuhan University of Science and Technology, and a Visiting Scholar with Portsmouth University, U.K. He has published more than 200 papers, including more than 20 papers in SCI and more than 140 papers in EI. He mainly engaged in the optimization and control of steel production processes, production planning and scheduling, dynamic design and simulation of electromechanical systems, intelligent design and control, and intelligent robots.



HAO WU received the B.S. degree from the College of Machine and Electrical Engineering, Anhui Polytechnic University, Wuhu, China. He is currently pursuing the M.S. degree with the Wuhan University of Science and Technology, China. His current research interests include intelligent robot and image processing.

GUOZHANG JIANG received the B.S. degree from Chang'an University, China, and the M.S. degree from the Wuhan University of Technology, China, and the Ph.D. degree from the Wuhan University of Science and Technology, China, where he is currently a Professor, the Doctoral Supervisor, the Deputy Dean of the School of Mechanical Automation, and the famous Teacher. He has published more than 80 academic papers, including more than 30 papers including SCI, EI, and ISTP. He is mainly engaged in teaching and scientific research in the fields of mechanical engineering and industrial engineering. He is also the Vice Chairman of the Industrial Engineering Branch of the Hubei Mechanical Engineering Society, the Tutor of the Youth Entrepreneurship Mentor of Hubei Province, and a Member of the Democratic League Committee of Hubei Province.



SHUANG XU received the Ph.D. degree from the Huazhong University of Science and Technology, China. She is currently a Lecturer with the Wuhan University of Science and Technology. Her research interests include artificial intelligence and image processing.

HONGHAI LIU received the Ph.D. degree in intelligent robotics from the Department of Systems Engineering, Kings College London, University of London, U.K. He is currently the Director of the Department of Intelligent Systems and Biomedical Robotics, University of Portsmouth, U.K., and also the Director and a Professor with the Department of Systems Engineering, U.K. He is also a Professor with the School of Mechanical and Power Engineering, Shanghai Jiao Tong University. He has presided over and completed nearly 30 key research projects of the Engineering and Physical Sciences Research Council, the seventh Framework Program (European Union), the Royal Society, the Royal Academy of Engineering, and the Department of Trade and Industry, and has achieved a series of internationally influential innovations. He has published more than 300 papers in international journals and conferences in the field of robotics and biomechanics integration, including the IEEE Transactions on Industrial Informatics, the IEEE Transactions on Fuzzy Systems, and the IEEE Transactions on Mechatronics, and more than 100 famous journal articles.