A Static Gesture Recognition Method Based on Improved SURF Algorithm and Bayesian Regularization BP Neural Network

Hongji Xu, Haibo Cao

School of Information Science and Engineering, Shandong University, China
hongjixu@sdu.edu.cn, chb831106@126.com

Abstract

Gesture recognition plays an important role in the aspect of human computer interaction (HCI). It has become one of the most challenging tasks in the pattern recognition field. So far, many gesture representations using two-dimensional image have been proposed, but normally they are vulnerable to environmental factors, such as illumination, cluttered backgrounds and so on. In this paper, we propose a static gesture recognition method based on the improved speed up robust feature (SURF) algorithm and the Bayesian regularization back propagation (BP) neural network with the Microsoft Kinect sensor. With the advantages of the Kinect, we can capture the depth data to enhance the robustness of the proposed algorithm. Gesture analysis can be viewed as a two-fold problem, i.e., gesture representation and classification. On the one hand, we implement gesture segmentation by the depth data, and then extract the feature descriptor of the gesture based on the improved SURF algorithm which is optimized through the key point detection and orientation calculation. On the other hand, the method based on the Bayesian regularization BP neural network is used as classifier. Subsequently, in order to further intensify the recognition accuracy, another method of classification of gestures based on maximum angle between fingers is proposed as well. Finally, two kinds of classification results are also combined to get the final classification result. The experimental results show that the proposed method can eliminate the interference of the background, and enhance the robustness and accuracy of the gesture recognition.

Keywords: Depth data, Speed up robust feature, Back propagation neural network, Gesture recognition

1 Introduction

Human computer interaction (HCI) technology has got a great progress with the development of information technology. As an important research field of HCI, gesture recognition has become a research hot-spot in recent years. Traditional gesture recognition methods are mainly based on the mouse and pen, vision and data glove, and they mainly include template matching, neural network and statistical analysis. But these techniques have some limitations in applications. For example, the method based on vision is susceptible to light and complex background conditions, and its recognition accuracy is low. The method based on data glove needs special sensing devices, and users are very inconvenient. Gesture recognition has been further developed after the invention of the Microsoft Kinect sensor.

Lai et al. put forward a hand gesture recognition method which was based on YCrCb color space to separate the hand area, and used convex envelope to detect fingertip [1]. Hussain et al. presented a gesture recognition method which characterized the fingertips and palm centers as gesture features, and used principal component analysis (PCA) to eliminate gesture ambiguity [2]. Wang et al. presented a new superpixel-based hand gesture recognition system which was based on a novel superpixel earth mover’s distance metric, together with Kinect depth camera [3]. The depth and skeleton information from Kinect were effectively utilized to produce markerless hand extraction. Li et al. proposed a fingertip detection method which was called depth-based convex defect detection (DB-CDD) [4]. It segmented the gesture by the depth data and adjacent features, and then extracted the contour of the gesture by Canny algorithm [5] which was based on binary image. Firstly, the method detected all the fingertips area for the convex hull roughly, and then eliminated all the pseudo fingertips points, and it was quick and efficient to detect the fingertips. Zhao et al. presented a gesture recognition algorithm based on the skeleton data captured by the Kinect. This algorithm extracted the features of gesture which were obtained between the starting point and the end point, and then recognized the gesture by the dynamic time wrapping (DTW) algorithm which was improved by distance weighting [6].

The above methods need better image quality for

*Corresponding Author: Hongji Xu; E-mail: hongjixu@sdu.edu.cn
DOI: 10.3966/160792642021052203019

feature extraction, otherwise the accuracy of gesture recognition would be relatively low. In this paper, a method of static gesture recognition based on the improved speed up robust feature (SURF) algorithm and the Bayesian regularization back propagation (BP) neural network is proposed to overcome the above shortcomings. The proposed method firstly segments the gesture using the depth data; and then obtains the feature descriptor of the gesture based on the improved SURF algorithm which is optimized through key point detection and orientation calculation; finally does the classification based on Bayesian regularization BP neural network. In order to further improve the recognition accuracy, another method of classification of gestures based on maximum angle between the fingers is proposed as well in this paper. Finally, two kinds of classification results are combined to get the final classification result.

2 Gesture Segmentation

2.1 Kinect Sensor

With the development of advanced information technology, the gesture recognition has evolved from touchpad to touchless sensors, and it does not need direct contact [7]. Kinect sensor based on somatosensory interaction is a HCI equipment which was invented by the Microsoft in 2010 [8], and now updated as Azure Kinect DK sensor. It combines all kinds of techniques including sound, light, electricity, machinery and so on. Kinect sensor mainly consists of the following components: RGB camera in the middle is used to capture color images; two depth cameras on both sides can capture the depth data; a microphone array is capable of localizing sound sources and recognizing speech; a motor is used to adjust the pitch angle. Kinect sensor supports face recognition, gesture recognition, bones tracking, speech recognition functions and so on, and has been widely studied in the fields of robot, 3D printing and augmented reality, etc.

2.2 Gesture Segmentation

In the proposed algorithm, we firstly segment the human body and background by using the indexes extracted from the depth data of Kinect [9-10], and the indexes are different between the human body and the background. Conventionally, gesture is in front of the human body, so its depth data is different from that of the human body. We segment the gesture from the human body by many threshold values, and then save the gesture as binary images. The upper left corner of Figure 1 is depth map which is captured by Kinect, and the right one is the human body segmentation map, and the bottom left one is the gesture segmentation map. Figure 1 shows that the above method overcomes some kinds of interference of illumination and background, and can segment the gesture from the background and human body, which demonstrates that it is a feasible method.

Figure 1. Gesture segmentation

3 Feature Extraction

The SURF algorithm [11] proposed by Panchal et al. in 2006 is a kind of local feature point detector, which has high robustness and can be used in the computer vision field. This algorithm improves the efficiency of time by the integral image [12], Haar wavelet transformation and approximate Hessian matrix [13], and it consists of key point detection and feature description. Its advantages are obvious for the matching of color images, but large amount of calculation and more interferences reduce the accuracy for the binary images. In this paper, we optimize the method through the key point detection and orientation calculation to overcome the above disadvantages. Figure 2 shows the flow chart of the improved SURF algorithm, and the key steps of the proposed algorithm are as follows:

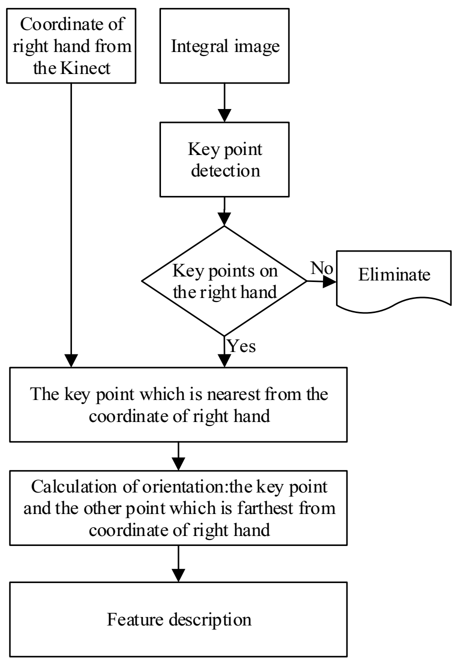


Figure 2. Flow chart of the improved SURF

3.1 Key Point Detection

The SURF algorithm detects the key points based on integral image and Hessian matrix. When the local extremum of Hessian matrix determinant is the maximum value, the opposite point would be the key point. These key points are brighter or darker than the adjacent points. In this paper, we use the right hand for gesture recognition. As shown in Figure 3, the center of each circle is the position of one key point. The green circles (i.e. the circles with the radius lines) represent the key points outside the hand and the red circles represent the key points on the hand.

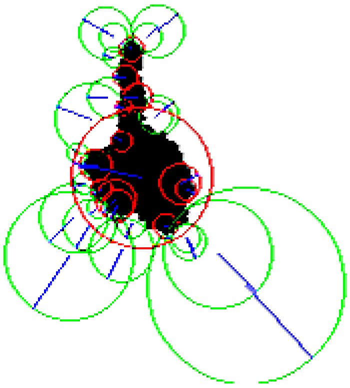


Figure 3. Key points detected by the SURF

Because these images of gestures captured by Kinect are binary images and their features are very simple, we only require few key points on the hand to describe the images. Figure 3 shows that some key points are outside the hand, which would cause interference and increase the amount of calculation for gesture recognition. These key points outside the hand are called noise points in this paper.

In order to overcome the above shortcomings, we propose a new method for the key point detection. The new method extracts the pixel values of the key points which are detected by the SURF algorithm, and eliminates these key points when the pixel value is 0, and then finds the key point which is nearest to the coordinate of the right-hand joint obtained by Kinect.

From Figure 4 we can know that the green circles in Figure 3 which are noise points have been eliminated, and the red circles which are the key points are retained on the hand.

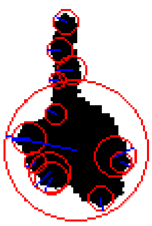


Figure 4. Key points extracted by the SURF on the hand

The red circle shown in Figure 5 is the key point which is the nearest to the right-hand joint of the Kinect.

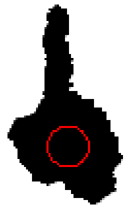


Figure 5. Key point extracted by the improved SURF

3.2 Calculation of Orientation

To keep the rotation invariance, the SURF algorithm sets a main orientation [14-15] for each key point. It sets the key point as the center of the circle, 6s as radius (where s is the scale and automatically selected by the SURF algorithm), and calculates Haar wavelet response at horizontal direction and vertical direction in the circle. Then sets a sector of 60 degrees and rotates, where the apex of the sector is the key point, and calculates Haar wavelet response in this sector area. The orientation which corresponds the maximum of the Haar wavelet response is the main orientation of the key point.

In the above proposed algorithm, we only define one key point, so we propose a new method of orientation calculation. We find the farthest key point to the right-hand joint after eliminating noise points, and set the direction of the vector of the nearest key point to the farthest key point as the main direction. The blue line in Figure 6 shows the main orientation of the key point.

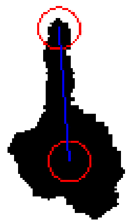


Figure 6. Main orientation

3.3 Feature Description

The SURF algorithm firstly builds a square around the key point. The square sets the key point as the center and its length is 20s. The square is divided into 4×4 sub-square, and we calculate Haar wavelet response at horizontal direction and vertical direction for every sub-square. In each sub-square, we get four features: $\sum d_x$, $\sum d_y$, $\sum |d_x|$, and $\sum |d_y|$. Subsequently, we can get a 64-dimensional feature vector which is the feature descriptor of the gesture [16].

When calculating Haar wavelet response, we need to set one threshold value. The smaller the threshold value is, the more the key points are. On the contrary, the key points are fewer. In this paper, we set the threshold value as 0.0001. Table 1 shows the efficiency comparison of the improved SURF algorithm and the SURF algorithm, and it can be seen that the number of key points for the improved SURF is only 1, which is much smaller than that of the original SURF algorithm. Meanwhile, the computing time is also reduced from 55.07s to 0.9947s, i.e., reduced by about 55 times. It demonstrates that the efficiency of the improved SURF algorithm is much better than that of the SURF algorithm.

Table 1. Efficiency comparison of the improved SURF algorithm and the SURF algorithm

Algorithm	Threshold	The number of key points	Computing time(s)
SURF	0.0001	43	55.07
Improved SURF	0.0001	1	0.9947

4 Bayesian Regularization BP Neural Network

4.1 BP Neural Network

BP neural network [17-18] which is widely used is a forward-feedback neural network, and the main advantages lie in highly self-organizing, self-learning, adaptive ability and nonlinear mapping ability. Generally, there are three layers, i.e. the input layer, output layer and hidden layer. Each layer contains many neurons which are parallel and interconnect between layers.

4.2 Bayesian Regularization Algorithm

Usually BP neural network would set an expected error during training. If the expected error is very small, the network would be over-fitting which can affect its performance [19]. To overcome the shortcoming, we use Bayesian regularization BP neural network in the proposed algorithm.

Bayesian regularization algorithm [20-21] can improve the generalization ability through modifying the training function of BP neural network. The BP neural network usually uses the mean square error (MSE) as a performance function. The modified performance function of Bayesian regularization algorithm can be expressed as

$$F(w) = \beta E_d + \alpha E_w \tag{1}$$

where E_d is the MSE and E_w is the sum square of weights (SSW). α and β are the regularization coefficients, which can affect the network performance. If α is far less than β, the algorithm is to reduce the training error, but it could result in over-fitting. On the contrary, the algorithm can improve the generalization ability, but the training error would be larger. One advantage of Bayesian regularization algorithm is that it can adjust α and β adaptively to make them optimal. Therefore, the regularization coefficients α and β can be obtained by maximizing the posterior probability, which are expressed as

$$\alpha = \frac{\gamma}{2E_w}, \tag{2}$$

$$\beta = \frac{N-\gamma}{2E_d} \tag{3}$$

where γ represents the number of valid network parameters and N represents the total number of network parameters.

4.3 Experiment Analyses

In the experiment, hardware devices consist of Kinect sensor and a computer, and the programming software is C# and Matlab.

In order to show the performance of the proposed method, we capture a dataset containing 400 samples for each of the eight gestures, i.e., numbers 1 to 8 (Chinese sign language) as the training set, and capture 70 samples and 30 samples for each gesture as the test sets, respectively. The performance of the Bayesian regularization (B-R) algorithm is compared with the Levengerg-Marquardt (L-M) optimization algorithm [22].

Figure 7 shows the performance comparison of the B-R algorithm and the L-M optimization algorithm, and it can be seen that the accuracy of the B-R algorithm is better than that of the L-M optimization algorithm in both cases. For the B-R algorithm, the accuracies of eight gestures are relatively similar in the case of 70 samples but the gap becomes larger in the case of 30 samples.

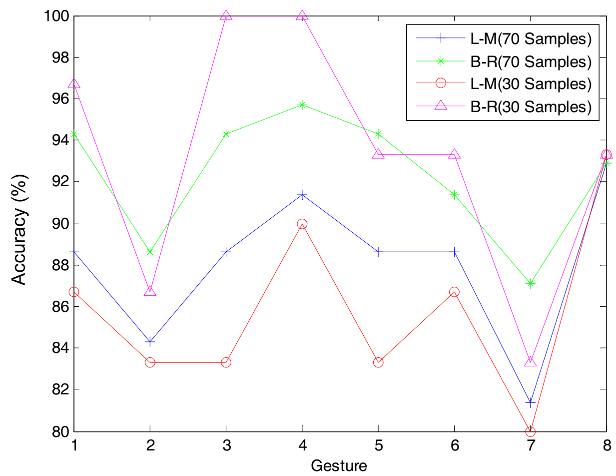


Figure 7. Performance comparison of the B-R algorithm and the L-M optimization algorithm

Table 2 shows the accuracy of every gesture and the average accuracies of the eight gestures in both cases.

Table 2. Accuracy comparison of two test sets

Gesture	70 Samples (%)		30 Samples (%)	
	L-M	B-R	L-M	B-R
1	88.6	94.3	86.7	96.7
2	84.3	88.6	83.3	86.7
3	88.6	94.3	83.3	100
4	91.4	95.7	90	100
5	88.6	94.3	83.3	93.3
6	88.6	91.4	86.7	93.3
7	81.4	87.1	80	83.3
8	92.9	92.9	93.3	93.3
Average accuracy	88.1	92.3	85.2	93.3

From Table 2 we can know that the average accuracy of the B-R algorithm of 30 samples is better than that of 70 samples, and the average accuracies are 93.3% and 92.3%, respectively. Table 2 shows that the accuracies of number 1, 3 and 4 are better in both cases for the B-R algorithm, but for the number 2 and 7, the accuracies are worse. In addition, the gesture 2 is liable to be identified as gesture 3, because the feature descriptors are similar for them. The accuracy of gesture 7 is worse, because the key points of it extracted by the SURF are very concentrated, and the errors of the main orientations are larger. By comparing, we have confirmed the effectiveness of the proposed method.

5 Feature Extraction Based on Maximum Angle Between Fingers

In order to further improve the system performance, this paper proposes another method of feature extraction and classification based on maximum angle between fingers, and fuses the two kinds of classification results for the final recognition result.

5.1 Feature Extraction

The method is also based on the SURF algorithm and Kinect, and the first two steps are the same as the above improved SURF algorithm. Firstly, it gets all the key points of gestures by the SURF algorithm, and then removes all noise points.

The next step is to get the coordinate of the right hand by Kinect, and then calculate the distance of all key points and the right-hand joint after removing noise points, finally, choose the nearest point as the center point of the hand. At the same time, it removes all the key points on the palm by using the threshold. In this paper, we take the distance of the center point and these key points as the threshold.

Figure 8 shows that the red circle is the center point of the hand which is the nearest key point to the right-hand joint.

Figure 8. Center point of the hand

Figure 9 shows the key points of the fingers, and it can be seen that the key points on the palm are removed.

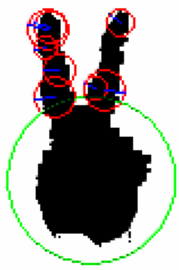


Figure 9. Key points of the fingers

After determining the center point of the hand and key points of the fingers, the follow steps are to calculate the angle of the center point and the key points on the fingers, which are expressed as

$$\omega = \begin{cases} 2\pi - \arctan\dfrac{(Y_2 - Y_1)}{(X_2 - X_1)} & X_1 < X_2, Y_1 < Y_2 \\ \arctan\dfrac{(Y_1 - Y_2)}{(X_2 - X_1)} & X_1 < X_2, Y_1 > Y_2 \\ \pi + \arctan\dfrac{(Y_2 - Y_1)}{(X_1 - X_2)} & X_1 > X_2, Y_1 < Y_2 \\ \pi\text{-}\arctan\dfrac{(Y_1 - Y_2)}{(X_1 - X_2)} & X_1 > X_2, Y_1 > Y_2 \\ \dfrac{3\pi}{2} & X_1 = X_2, Y_1 < Y_2 \\ \dfrac{\pi}{2} & X_1 = X_2, Y_1 > Y_2 \\ 0 & \text{The number of key points is 0} \end{cases} \quad \textbf{(4)}$$

where ω is the angle of the center point and these key points on the fingers, X_1, Y_1 are the coordinates of the center point of hand, and X_2, Y_2 are the coordinates of the key points of the fingers.

After calculating the angle ω, we can calculate the difference φ between the maximum and the minimum values of the angle ω. The difference φ is the feature vector of the gesture, which is expressed as

$$\varphi = \omega_{\max} - \omega_{\min}. \quad \textbf{(5)}$$

Figure 10 shows that the angle of two green lines is the feature vector φ.

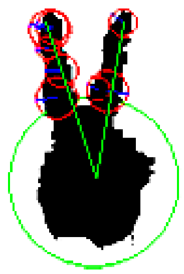


Figure 10. Feature vector φ

5.2 Classification

We calculate the feature vector of each gesture by the above method. Figure 11 shows the feature vectors of the training set.

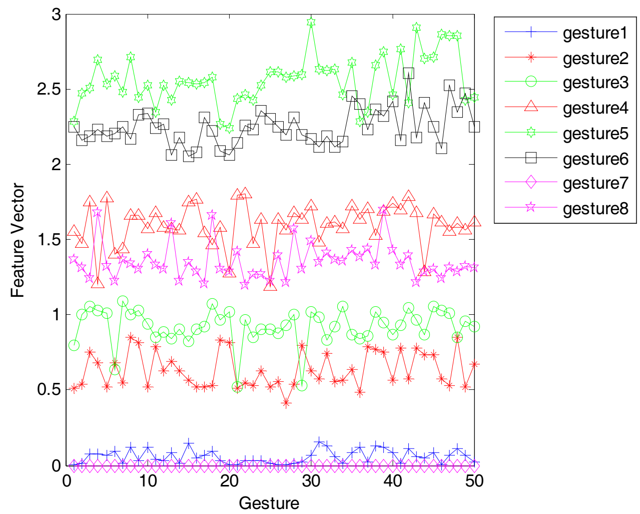


Figure 11. Feature vectors of the training set

From Figure 11 we can know that the feature vectors of the gesture 1 and gesture 7 are similar, and it is the same for that of gesture 2 and gesture 3, gesture 4 and gesture 8, gesture 5 and gesture 6. Therefore, according to the size of the feature vectors, gestures can be divided into four categories: gesture 1 and gesture 7, gesture 2 and gesture 3, gesture 4 and gesture 8, gesture 5 and gesture 6.

6 Decision Level Fusion

Data fusion [23-24] is the process of analyzing and synthesizing multi-source information according to some criteria to achieve a certain objective. The methods of data fusion consist of pixel level fusion, feature level fusion and decision level fusion. This paper adopts the decision level fusion for the final classification.

6.1 The Method of Decision Level Fusion

The first step is to calculate the classification results of the above two kinds of classification methods. We use I to denote the classification result of the B-R algorithm, and use C to denote classification result of the method of the maximum angle between the fingers, which are expressed as

$$I = \{1, 2, 3, 4, 5, 6, 7, 8\}, \quad (6)$$

$$C = \{(1,7)(2,3)(4,8)(5,6)\}. \quad (7)$$

The second step is to calculate the absolute distance between I and C, and then get the absolute distance M, which can be expressed as

$$M = |I - C|. \quad (8)$$

Finally, the minimum value of M would be the final classification result R, i.e.,

$$R = M_{\min}. \quad (9)$$

6.2 Experiment Analyses

Here we capture 70 samples for each gesture as the test set. The performance of the decision level fusion (D-L-F) algorithm is compared with the B-R algorithm.

Table 3 and Figure 12 are the performance comparison of the D-L-F algorithm and the B-R algorithm.

Table 3. Accuracy comparison of the D-L-F algorithm and the B-R algorithm

Gesture	70 Samples (%)	
	D-L-F	B-R
1	95.7	94.3
2	92.9	88.6
3	98.6	94.3
4	98.6	95.7
5	95.7	94.3
6	97.1	91.4
7	95.7	87.1
8	97.1	92.9
Average accuracy	96.4	92.3

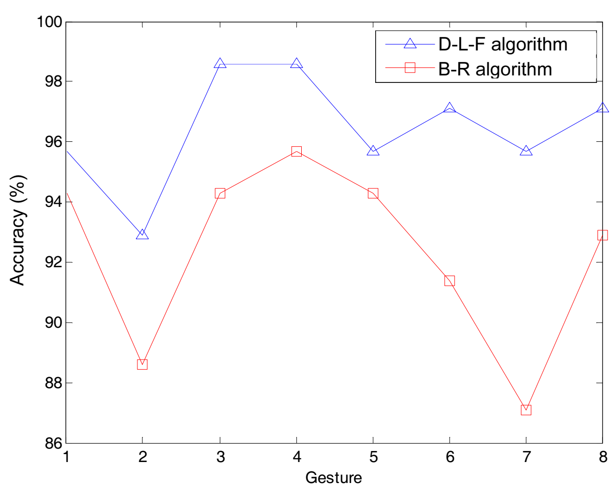


Figure 12. Performance comparison of the D-L-F algorithm and the B-R algorithm

Table 3 shows that the average accuracy of the D-L-F algorithm is 96.4%, which is higher than that of the B-R algorithm. The recognition accuracy of gesture 2 and 7 are the worst for the B-R algorithm, but they are 92.9% and 95.7% in the D-L-F algorithm. It proves that the D-L-F algorithm eliminates the disadvantages of the B-R algorithm and can improve recognition accuracy of the SURF algorithm.

Figure 12 shows that the recognition accuracy of each gesture is improved, and those of gesture 6 and gesture 7 are increased by about 5.7% and 8.6%.

By comparison, we confirm that the system performance is better than that of the single classifier.

7 Conclusion

In this paper, we propose a static gesture recognition method based on the depth data captured by Kinect which can overcome some kinds of interference of illumination and background. To improve the efficiency, we optimize the SURF algorithm through key point detection and calculation of orientation, and these gestures are classified by the Bayesian regularization BP Neural Network and the proposed classification method based on maximum angle between the fingers, and then two kinds of classification results are combined to get the optimal output. The experiments show that the proposed method has a better performance than the compared one and can enhance the robustness and accuracy of gesture recognition. However, for some very similar gestures, the accuracy of recognition is still not very high.

In addition, with the development of the convolutional neural network (CNN), it can bring very good performance and very high accuracy. It normally needs to be supported by certain computing equipment and resources, such as GPU, FPGA and ASIC. For the future work, we plan to combine the feature level fusion and decision level fusion, and also try to use some state-of-the-art neural network techniques to further optimize the system performance.

Acknowledgments

This work was financially supported by the Natural Science Foundation of Shandong Province of China (ZR2020MF139) and the National Key Research and Development Program of China (2018YFC0831001).

References

[1] H. Y. Lai, H. J. Lai, Real-Time Dynamic Hand Gesture Recognition, *2014 2nd International Symposium on Computer, Consumer and Control (IS3C)*, Taichung, Taiwan, 2014, pp. 658-661.

[2] I. Hussain, A. K. Talukdar, K. K. Sarma, Hand Gesture Recognition System with Real-Time Palm Tracking, *2014 11th Annual IEEE India Conference (INDICON)*, Pune, India, 2014, pp. 1-6.

[3] C. Wang, Z. Liu, S.-C. Chan, Superpixel-based Hand Gesture Recognition with Kinect Depth Camera, *IEEE Transactions on Multimedia*, Vol. 17, No. 1, pp. 29-39, January, 2015.

[4] C. Li, *The Research of Gesture Recognition Based on Depth Image from Kinect*, Master's Thesis, University of Xihua, Chengdu, China, 2014.

[5] X. Hou, H. Liu, Welding Image Edge Detection and Identification Research Based on Canny Operator, *2012 2nd International Conference on Computer Science and Service System (CSSS)*, Nanjing, China, 2012, pp. 250-253.

[6] F. Zhao, Z. Liu, Y. Lv, Gesture Recognition Based on Kinect Skeleton Information, *China Sciencepaper*, Vol. 10, No. 20, pp. 2412-2415, October, 2015.

[7] J. Ferri, R. L. Llopis, J. Moreno, J. I. Civera, E. Garcia-Breijo, A Wearable Textile 3D Gesture Recognition Sensor Based on Screen-Printing Technology, *Sensors*, Vol. 19, No. 23, Article No. 5068, December, 2019.

[8] Z. Zhang, Microsoft Kinect Sensor and Its Effect, *IEEE MultiMedia*, Vol. 19, No. 2, pp. 4-10, April, 2012.

[9] O. Lohmann, T. Luhmann, A. Hein, Skeleton Timed Up and Go, *2012 6th IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Philadelphia, PA, USA, 2012, pp. 1-5.

[10] M. Tarabini, D. Marchisotti, R. Sala, P. Marzaroli, H. Giberti, M. Sculati, A Prototype for the Automatic Measurement of the Hand Dimensions Using the Microsoft Kinect V2, *2018 13th IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome, Italy, 2018, pp. 1-6.

[11] P. M. Panchal, S. R. Panchal, S. K. Shah, A Comparison of SIFT and SURF, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 1, No. 2, pp. 323-327, April, 2013.

[12] Sareena, V. Sundaram, Dynamic Selection of Robust Features for Tampering Detection Based on SURF, *International Journal of Applied Engineering Research*, Vol. 11, No. 5, pp. 3591-3595, August, 2016.

[13] L. Tao, X. Jing, S. Sun, H. Huang, N. Chen, Y. Lu, Combining SURF with MSER for Image Matching, *2013 9th IEEE International Conference on Granular Computing (GrC)*, Beijing, China, 2013, pp. 286-290.

[14] L. Zhu, Y. Wang, B. Zhao, X. Zhang, A Fast Image Stitching Algorithm Based on Improved SURF, *2014 10th International Conference on Computational Intelligence and Security (CIS)*, Kunming, China, 2014, pp. 171-175.

[15] W. Wei, C. A, Y. Zhao, G. Zhang, Image Registration Algorithm Based on Super Pixel Segmentation and SURF Feature Points, *2018 8th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Xi'an, China, 2018, pp. 1-5.

[16] J. Chao, R. Huitl, E. Steinbach, D. Schroeder, A Novel Rate Control Framework for SIFT/SURF Feature Preservation in H.264/AVC Video Compression, *IEEE Transactions on*

Circuits and Systems for Video Technology, Vol. 25, No. 6, pp. 958-972, June, 2015.

[17] Y. Shan, Y. Liu, F. Guan, The Application of BP Neural Network Algorithm in Optical Fiber Fault Diagnosis, *2015 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, Guiyang, China, 2015, pp. 509-512.

[18] H. Bing, F. Qiang, Application of BP Neural Network in the Evaluation of Standard Implementation, *2018 3rd IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, China, 2018, pp. 120-125.

[19] H. Zhao, L. Shi, H. Wang, Sunlight Greenhouse Temperature Prediction Model Based on Bayesian Regularization BP Neural Network, *Applied Mechanics and Materials*, Vol. 740, pp. 871-874, March, 2015.

[20] R. Gencay, M. Qi, Pricing and Hedging Derivative Securities with Neural Networks: Bayesian Regularization, Early Stopping, and Bagging, *IEEE Transactions on Neural Networks*, Vol. 12, No. 4, pp. 726-734, July, 2001.

[21] N. Parihar, R-Peak Detection Using Bayesian Regularization Neural Network, *2017 12th IEEE International Conference on Information, Communication, Instrumentation and Control (ICICIC)*, Indore, India, 2017, pp. 1-3.

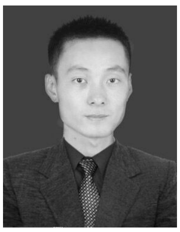
[22] C. Mu, J. Wang, Z. Yuan, X. Zhang, C. Han, The Research of the ATR System Based on Infrared Images and L-MBP Neural Network, *2013 7th International Conference on Image and Graphics (ICIG)*, Qingdao, China, 2013, pp. 801-805.

[23] A. Mumtaz, A. Majid, A. Mumtaz, Genetic Algorithms and Its Application to Image Fusion, *2008 4th International Conference on Emerging Technologies (ICET)*, Rawalpindi, Pakistan, 2008, pp. 6-10.

[24] J. Xu, D. Song, D. Ding, S. Tao, R. Chen, High Resolution 2D-Imaging Based on Data Fusion Technique, *2018 4th IEEE International Conference on Computational Electromagnetics (ICCEM)*, Chengdu, China, 2018, pp. 1-3.

Biographies

Hongji Xu received the Ph.D. degree in communication and information system from Shandong University, China in 2005. He is currently an associate professor in the School of Information Science and Engineering at Shandong University. His research interests include ubiquitous computing, blind signal processing, and human computer interaction.



Haibo Cao received the B.E. degree in electronic and information engineering from Hebei University, Hebei, China, in 2007. He was a M.S. student in the School of Information Science and Engineering at Shandong University, Jinan, China. His research interests include machine learning, computer vision and pattern recognition.