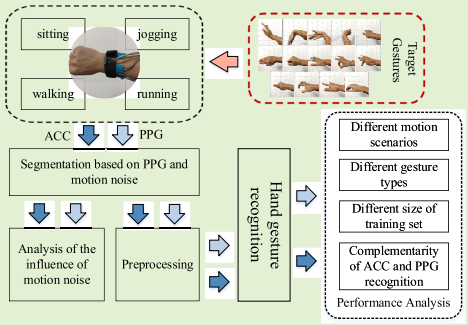
Comparative Study of Gesture Recognition Based on Accelerometer and Photoplethysmography Sensor for Gesture Interactions in Wearable Devices

Yu Ling, Xiang Chen, *Member, IEEE*, Yuwen Ruan, Xu Zhang, *Member, IEEE*, and Xun Chen, *Senior Member, IEEE*

***Abstract*—With the goal of implementing gesture interaction applications on wearable devices, this paper presents a comparative study on hand gesture recognition based on accelerometer (ACC) and photoplethysmography (PPG) sensors. Research is conducted on 14 hand gestures involving wrist and finger movements under four motion scenarios: sitting, walking, jogging, and running. When the participants perform the hand gestures, four channels of PPG signals and 3D ACC signals are collected. To effectively extract gesture repetitions from the continuous movements, a segmentation method based on PPG signal and motion noise is employed. The experimental results verify that the PPG signal is more suitable for implementing gesture interactions on wearable devices. At first, the signal-to-noise ratio analysis demonstrates that the ACC signals are much more affected by motion noise than the PPG signals. Then, the hand gesture recognition results in the four different motion scenarios demonstrate that the accuracies of both the ACC and PPG signals decrease with the intensity of body movement; however, the decrease of the PPG signals is significantly lower than that of the ACC signals, especially for finger-related gestures. A hand gesture recognition experiment with ratios of the samples between the training set and test set from 1:9 to 8:2 demonstrates that the PPG signal also has advantages in regards to reducing user's training burden. Moreover, this research verifies that the green PPG signal is not the best choice for the implementation of gesture interactions. The research results provide a reliable foundation for promoting the development of gesture-based interactions in wearable devices.**



***Index Terms*—Pattern recognition, acceleration, photoplethysmography, wearable sensors.**

I. Introduction

OWING to their advantages in regards to clear language meanings and easy implementation, gestures are suitable for controlling mobile phones, computers, and other terminal devices in human-computer interactions. High-accuracy gesture recognition technology is the key to ensuring practicability and good user experiences in gesture interactions. Generally, gesture recognition technologies can be classified into two main categories, according to the sensing technologies. The first category concerns vision-based techniques, which use imaging equipment (such as cameras) to capture gesture information [1], [2]. The second category concerns motion sensor-based techniques, which use devices such as data gloves [3], [4], electromyographic (EMG) sensors [5]–[7], accelerometers (ACCs), gyroscopes [8], [9], and mechanomyograms [10], [11] to capture gesture information. All of these gesture sensing technologies have obtained good research progress and have already shown significant commercial application value in the fields of rehabilitation engineering, entertainment [12], device control [13], and education [14].

In recent years, with developments in microelectronics and computer technology, wearable devices have become increasingly popular in human daily life. The global shipments of wearable devices reached 336.5 million in 2019, an 89% increase from 2018 [15]. In addition to providing good communication and entertainment experiences, wearable devices, which are embedded with advanced sensors, also provide functions for monitoring health and activities. For instance, the Apple Watch Series 6 [16] is embedded with

Manuscript received April 19, 2021; accepted May 9, 2021. Date of publication May 19, 2021; date of current version July 30, 2021. This work was supported by the National Nature Science Foundation of China under Grant 61871360 and Grant 61671417. The associate editor coordinating the review of this article and approving it for publication was Prof. Kea-Tiong Tang. *(Corresponding author: Xiang Chen.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Ethics Review Committee of First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China, under Application No. PJ 2014-08-04.

The authors are with the Department of Electronic Science and Technology, University of Science and Technology of China (USTC), Hefei 230027, China (e-mail: xch@ustc.edu.cn).

Digital Object Identifier 10.1109/JSEN.2021.3081714

an ACC and gyroscope for fall detection, and a blood oxygen sensor and electric heart sensor for heart rate measurement. The MiBand5 [17] is embedded with an ACC, gyroscope, and photoplethysmography (PPG) sensor, so as to provide motion detections and heart rate estimations in different sports scenarios.

Among the sensors embedded in wearable devices, ACC and PPG sensors are the two main types of potential sensors for hand gesture recognition. The ACC has been widely applied in motion recognition, owing to the ability to capture the trajectory and direction of motion. Xu *et al.* [9] obtained an average accuracy of 95.6% in the classification of seven gestures using a 3D-ACC. Using an inertial measurement unit (IMU) comprising an ACC and gyroscope, Chu *et al.* [8] realized the accurate recognition of 11 gestures, with an average accuracy of 92.36%. Shoaib *et al.* [18] detected activities such as drinking coffee while sitting or smoking while standing, based on using an ACC and gyroscope placed at the wrist and pocket positions, respectively. Wen *et al.* [19] employed IMUs in smart watches to classify hand gestures for interaction purposes, and obtained an average accuracy of 87% for five finger-related gestures. Kang *et al.* [20] recognized nine large-scale spinning gestures with an average accuracy of 88.01% by using IMUs placed at wrist and head positions. In the work of [21], the participants were asked to hold an IMU sensor in one hand and to write 0–9 in the air; an average accuracy of 98.6% was obtained.

The PPG is usually used for health monitoring [22], [23], authentication [24], [25], and emotion recognition [26], [27]. As the motions of hand and wrist can change blood flows and muscle shapes, the PPG signal was shown to be promising for recognizing gestures in recent years [28]–[33]. Zhao *et al.* [31] initially explored the feasibility of using PPG signals for gesture recognition. They used a home-made PPG wrist band to recognize nine finger-based gestures, and obtained an average accuracy of 88%. Subramanian *et al.* [32] conducted a comparative research between PPG and surface EMG (sEMG) signals, and indicated that the PPG signal was a viable alternative modality for gesture recognition applications. Zhang *et al.* [30] conducted the first study to explore the feasibility of using optical sensors from off-the-shelf wearable devices to recognize gestures. They developed a novel mechanism called the gradient-based movement detection unit for detecting hand gestures using the heart rate monitors on commodity smartwatches, and obtained 90.55% accuracy and 90.73% recall for 10 hand gestures.

Although the popularity of wearable devices and mature sensor technologies have made it possible to implement gesture-based interactive applications in wearable devices, current gesture recognition researches based on wearable sensors remain very preliminary. Whether it based on IMU, PPG or other technologies, the number of recognizable gestures and robustness of hand gesture recognition algorithms must be further improved to meet the needs of practical applications. For instance, only five gestures were considered in the work of Wen *et al.* [19]. Only four simple hand gestures from four individuals were explored in the work of Subramanian *et al.* [32]. In the study of Zhang *et al.* [30], participants performed gestures while sitting, and 10 hand gestures were recognized with an accuracy of 90.55% by using PPG sensors embedded in a Samsung Gear 3 smart watch. However, the average accuracy descended rapidly to 73.6% when participants performed a jogging task. Zhao *et al.* [33] indicated that the PPG-based gesture recognition systems show weaknesses in regards to intense body movements.

Inspired by relevant researches, this study conducts a comparative investigation of gesture recognition technologies based on ACC and PPG sensors, and explores the feasibility of using these two types of sensors embedded in wearable devices to realize applications for gesture interactions. Considering the application scenarios for wearable devices in daily life, four different motion scenarios are considered, i.e., sitting, walking, jogging and running. The main innovations and contributions of this research are as follows: 1) a gesture repetition segmentation method is proposed and implemented, based on PPG signal and motion noise; 2) the influence of motion noise on the ACC and PPG gesture signals is explored; 3) the gesture recognition performances of the ACC and PPG signals are compared from the perspectives of motion scenarios, gesture types, and requirement on the sizes of training sets; and 4) the complementarity of the ACC and PPG signals for gesture recognition is explored. The research results lay a solid foundation for the implementation of gesture-based interaction applications in commercial wearable devices.

II. Materials and Method

As shown in Fig. 1., the research route of this study mainly includes data acquisition, gesture repetition segmentation, preprocessing, analysis of the influence of motion noise, hand gesture recognition, and performance analysis. This section introduces these parts in detail.

A. Hand Gesture Data Acquisition

As PPG sensors are designed to detect changes in blood volume, the target gestures must cause changes in the blood flows of blood vessels or muscle tissues. Furthermore, to facilitate the application of human-computer interactions, the gestures for interactions must be easy to perform, and have specific symbolic meanings. Consequently, as illustrated in Table I, 14 gestures with easily understood meanings and involving both finger and wrist movements were defined to compose the target gesture set. As shown in Fig. 2(a), a home-made wrist band embedded with four PPG sensors (MAX30105, Maxim Integrated Inc.) and a nine-axis motion tacking IMU (MPU-9250, InvenSense Inc.) including a three-axis gyroscope, three-axis ACC, three-axis magnetometer, and digital motion processor was used for data collection. Four PPG sensors are integrated into the system, using an adjustable elastic band. To explore the most suitable light for hand gesture recognition, each PPG sensor was designed to synchronously collect red, green, and infrared light.

Fourteen right-handed participants without any neural and musculoskeletal diseases or medical history were recruited. The information of participants is shown in Table II. All participants were informed of the experimental procedures, and signed an informed consent form approved by the Ethics

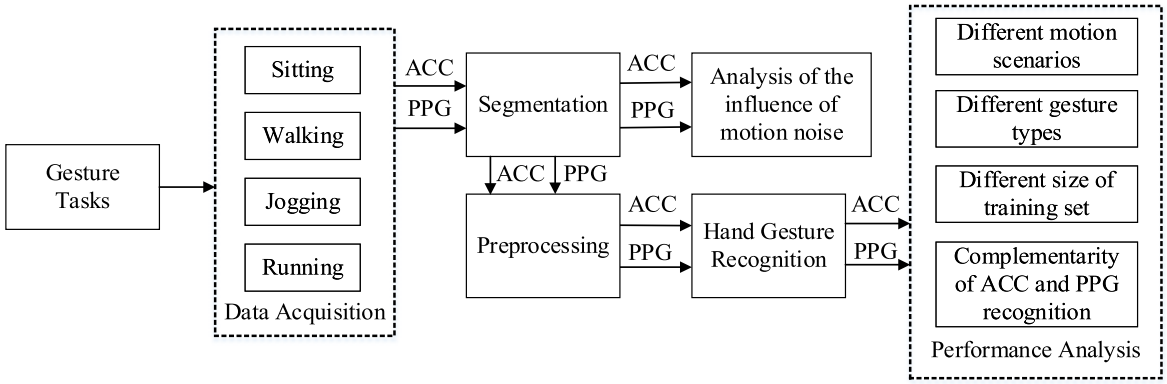
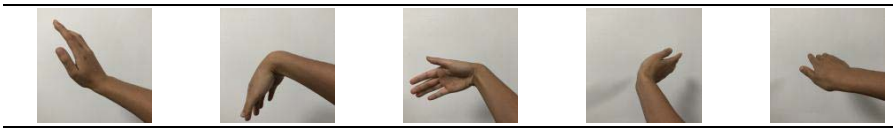


Fig. 1. Block diagram of the research route incorporating accelerometer (ACC) and photoplethysmography (PPG).

TABLE I
LIST OF THE 14 HAND GESTURES



G1. Up	G2. Down	G3. Left	G4. Right	G5. Forward

G6. Back	G7. Hold	G8. Extend	G9. Okay	G10. Click

G11. Great	G12. Grab	G13. Rock&Roll	G14. Repose

TABLE II
INFORMATION OF THE 14 PARTICIPANTS

Gender	Number	Age	Height(cm)	Weight(kg)
Male	8	23-27	170-182	65-90
Female	6	23-24	159-172	44-62

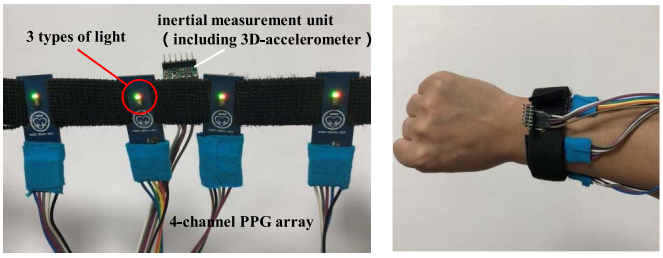


Fig. 2. (a) Home-made wrist band and (b) its placement.

Review Committee of First Affiliated Hospital of Anhui Medical University (No. PJ 2014-08-04). The participants were asked to avoid any strenuous exercise or heavy loads using their right arm for 24 hours before the experiment.

During the data collection experiment, as shown in Fig. 2(b), the participants wore the wrist band with the IMU on the back of their wrists, and the four PPG sensors equally around the wrist. They were asked to finish gesture tasks in four different motion scenarios, including sitting, walking (3.0 km/h), jogging (5.0 km/h), and running (8.0 km/h). In the sitting scenario, the participants sat on a chair. For the other three scenarios, a treadmill (Sole F63, produced by Sole, Inc.) was used to keep the participants walking or running at the same speed. The participants repeated each gesture 50 times in the sitting and walking scenarios, and 25 times in the jogging and running scenarios. There was no requirement regarding the duration of each repetition; however, a two-second interval between repetitions and a five-minute rest between gestures were implemented to avoid muscle fatigue. When a gesture was performed, the nine-axis motion signal and four-channel PPG signal were collected. The sampling rate of each channel for both the PPG and motion signals was 100 Hz. The detection range of accelerometer is 16g. In addition, in each scenario, approximately 10 s of baseline data without any gestures were collected as motion noise, as required for segmenting the gesture repetitions. All of the data were transmitted over Bluetooth, and then were saved to a disk for off-line analysis in Matlab R2018a (Mathworks, Inc) and Python 3.5.3 (Python Software Foundation).

B. Gesture Repetitions Segmentation and Preprocessing

When a gesture is performed, the waveforms and amplitudes of both PPG and ACC signals are significantly different from their baseline signals. Theoretically, the starting and ending

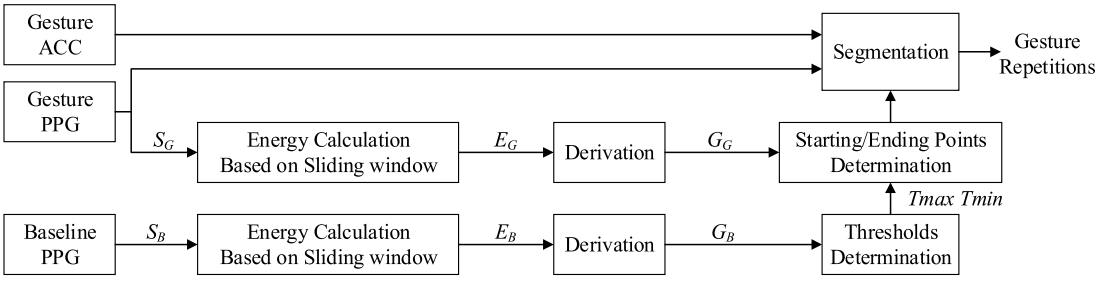


Fig. 3. Gesture repetitions segmentation strategy.

points of a gesture repetition can be determined based on the ACC or PPG signals, owing to the synchronization in signal acquisition. In this study, the PPG signal was chosen for segmenting the gesture repetitions, as the ACC signal is more sensitive to motion noise. Based on a previous reference [30], a data segmentation strategy based on PPG signals and motion noise was proposed. Fig. 3 shows the flowchart of the proposed method, and Fig. 4 illustrates an example of the segmentation of gesture repetitions of G7 in the sitting scenario. As depicted in Fig. 3 and Fig. 4, the gesture segmentation method was mainly conducted according to the following steps.

Step1: Initially, for both the baseline PPG and gesture PPG signals, one channel was selected as the target signal S_B/S_G (Fig. 4(a)) for the gesture repetition segmentation. In this study, the channel placed on the inner side of the wrist and near the ulnar artery was selected, owing to its sensitivity to hand movements. Then, to compute the energy signal E_B/E_G (Fig. 4(b)), S_B/S_G was traversed using a sliding window with a length of 0.3 s and a step size of 0.01 s. Finally, the gradient signal G_B/G_G (Fig. 4(c)) was obtained, based on a derivation of the E_B/E_G signal.

Step2: The thresholds T_{max} and T_{min} were determined based on the gradient signal G_B. Specifically, the average value of the peaks was calculated as T_{max}, and the average value of the troughs was calculated as T_{min}. The following method was adopted to find the peaks and troughs in G_B: for three consecutive points (a_1, b_1), (a_2, b_2), (a_3, b_3), if $(b_1\text{-}b_2)(b_2\text{-}b_3) < 0$ and $b_1\text{-}b_2 < 0$, (a_2, b_2) was selected as a candidate peak. When the amplitude value of the candidate peak was the largest value among the ten points before and after itself, the candidate peak was recorded as a real value. If $(b_1\text{-}b_2)(b_2\text{-}b_3) < 0$ and $b_1\text{-}b_2 > 0$, (a_2, b_2) was considered as a candidate trough. When the amplitude value of one candidate trough was the smallest value among the ten points before and after itself, the candidate trough was recorded as a real value.

Step 3: To determine the starting and ending points of the gestural repetitions, for the gesture gradient signal G_G, a sliding window with a length of 30 and step size of 1 was set. When more than twenty points of the window exceed the threshold range (T_{min}–T_{max}), a gesture repetition was considered to have appeared, and the first point of the window was seen as the starting point. Once twenty points of the window fell into the range of T_{min}–T_{max}, the gesture repetition was considered to end, and the first point falling into the threshold range of the window was seen as the ending point.

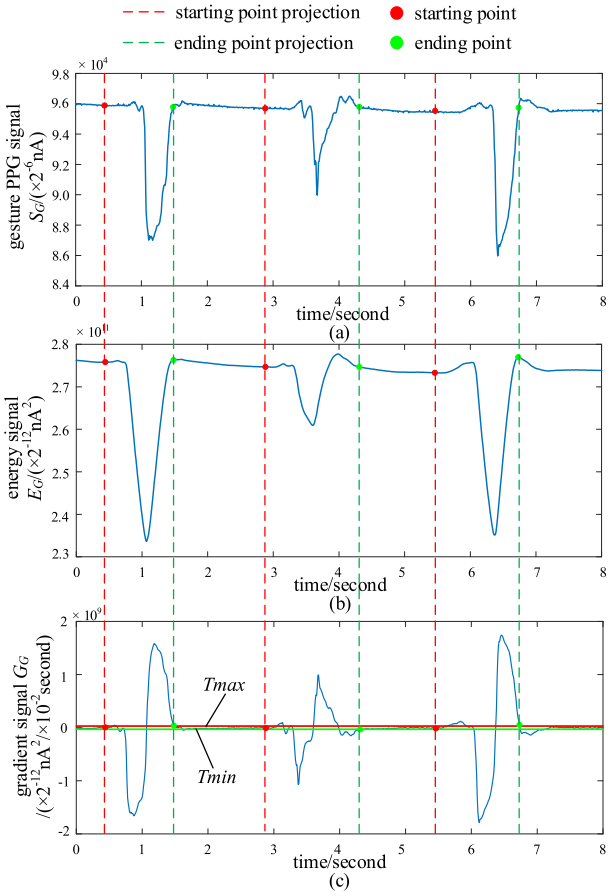


Fig. 4. Example of gesture repetitions segmentation (a) gesture photoplethysmography (PPG) signal S_G (b) energy signal E_G (c) gradient signal G_G.

If there was no window meeting the judgement condition within a length of 200 points, the corresponding ending point was marked as the point that adds 100 to the starting point.

Step 4: For both the PPG and ACC signals, the data between the starting points and corresponding ending points were segmented into gestural samples, usually with distances of 50–200 sampling points.

Assuming that N_{all} is the total number of actual gesture repetitions, $N_{segment}$ represents the number of signal segments obtained by the segmentation method, $N_{correct}$ is the number of those correctly detected, and N_{wrong} represents the number of those wrongly segmented. The correct detection rate and error detection rate were defined according to Equations (1) and (2) to measure the performance of the proposed

segmentation method.

$$Correct\_rate = N_{correct}/N_{all} \quad (1)$$

$$Error\_rate = N_{wrong}/N_{segment} \quad (2)$$

As the main frequency spectrum of a PPG signal is 0–5 Hz, a five-order low-pass Butterworth filter with a cut-off frequency of 5 Hz was used to remove the high-frequency noise in the PPG signal. For the 3-axis ACC signal, a zero-adjustment was made to ensure that the baselines of the three dimensions were the same. For each gesture repetition, the zero adjustment was performed by subtracting the average value of the repetition itself. Then, each channel of PPG or ACC signal was normalized using a min-max normalization method [34].

Because the execution times of different gestures (and even different repetitions) are always different from each other, the gesture repetitions were resampled into a fixed length. In particular, the fixed length L was set to 32. Assuming that the original signal length was l, for each specific point x in the resampled signal, the position projected to the original signal should have been $x*l/L$, and the two adjacent points in original signal corresponding to that position were marked as (x_1, y_1) and (x_2, y_2) $(x_2 > x_1)$. The value y of the position x in the resampled signal was calculated according to Equation (3). As four channels of the PPG signal and the 3-axis ACC signal were used for classification respectively, the sizes of the gesture samples for the PPG and ACC signal were ultimately determined as $4 \times L$ and $3 \times L$, respectively, and one hot label was matched to each sample. Each participant had four gesture datasets in each motion scenario, including *DB_PPG_r* for the red PPG signal, *DB_PPG_i* for the infrared PPG signal, *DB_PPG_g* for the green PPG signal, and *DB_ACC* for the ACC signal.

$$y = y_1 + \frac{x \times l/L - x_1}{x_2 - x_1}(y_2 - y_1) \quad (3)$$

C. Analysis of the Effects of Motion Noise on Gesture Samples

Because walking, jogging, running can cause changes in waveforms and amplitudes of ACC and PPG signals, the baseline signals that do not include gesture information are considered as motion noise in this study. To compare the effects of the motion noise on the ACC and PPG gesture signals, the signal-to-noise ratio (SNR) was calculated and analyzed for the four different motion scenarios. For a baseline signal S_B with a number of channels c (three for ACC and four for PPG), the mean value of the i-th channel $mean(S_B(i))$ was initially computed. For a target signal S_{target} with length L_t and channel c, whether using the baseline signal S_B or gesture segment signal S_G, the power Ps_{target} was calculated according to Equation (4), and the SNR was calculated using Equation (5).

$$P_{s_{target}} = \frac{1}{cL_t}\sum_{i=0}^{c}\sum_{j=0}^{L_t-1}\left|S_{target}(i,j) - mean(S_B(i))\right|^2 \quad (4)$$

$$SNR = 10lg(P_{s_G}/P_{s_B}) \quad (5)$$

D. Hand Gesture Recognition Scheme

The support vector machine (SVM), convolutional neural network (CNN), and long short-term memory (LSTM) were applied to gesture recognition in this study.

1) Support Vector Machine (SVM)-Based Gesture Classifier: The SVM [35] is a typical machine learning classification method. The main function of the SVM is to classify the samples in a data set $\boldsymbol{D}$ into $\boldsymbol{N}$ classes ($\boldsymbol{N}$ is the number of gestures to be recognized), based on the hyperplane $\boldsymbol{\omega^T X + b = 0}$. The ultimate target is to maximize the minimum geometric distance from a sample point (x_i, y_i) in $\boldsymbol{D}$ to the hyperplane as defined in Equation (6), where x_i denotes the feature vector, and y_i is the label of x_i. To further study nonlinear problems, the SVM provides a kernel function for mapping data to a high-dimensional space, i.e., to solve the linear inseparability problem in the original space. In this study, a linear kernel function was used for classification. Before one gesture sample was input into the SVM model, the signals from all channels were formed into a one-dimensional vector.

$$\gamma = \min_{i=1,2,\ldots,N} \gamma_i = \min_{i=1,2,\ldots,N} y_i\left(\frac{\omega}{\|\omega\|}x_i + \frac{b}{\|b\|}\right) \quad (6)$$

2) Convolutional Neural Network (CNN)-Based Gesture Classifier: Generally, a CNN [36] is sensitive to spatial information, and has a superior ability to extract features. As shown in Fig. 5, a 1D-CNN model was constructed for gesture recognition. The first two convolutional layers were set to primarily learn the features. The Relu non-linearity function was selected as activation function after each convolutional layer. Then, a max-pooling layer was adopted to reduce the complexity of the output. The following two convolutional layers were used for learning characteristics of higher levels; a global average pooling layer was applied to avoid over-fitting problem, in which each filter had only one weight; and a dropout layer after the global average pooling layer was used for randomly assigning zero weights to the neurons in the network to improve the accuracy of the model. A softmax function with 14 units was used in the last fully-connected layer to complete the classification. The network parameters that needed to be determined mainly included the number of filters, convolution kernels of each layer, and dropout value. The parameters were adjusted until the model was convergent and the loss function value and recognition accuracy reached better levels.

3) Long Short-Term Memory (LSTM)-Based Gesture Classifier: Compared to a CNN, a sequential neural network such as the LSTM [37] has a stronger ability to capture temporal information. As shown in Fig. 6(a), for each LSTM unit, there were three inputs, i.e., the current input x_t and two inputs from the previous unit cell state c_{t-1} and hidden state h_{t-1}, respectively. A forgetting gate F_t, input gate I_t and output gate O_t were incorporated to allow the LSTM to effectively manage the sequential signals. The two-layer LSTM shown in Fig. 6(b) was constructed for gesture recognition. After each LSTM layer, batch normalization was adopted to reduce the training time, and to avoid the explosion or disappearance of the gradient. The dropout was applied after each batch normalization layer to avoid over-fitting. When the training

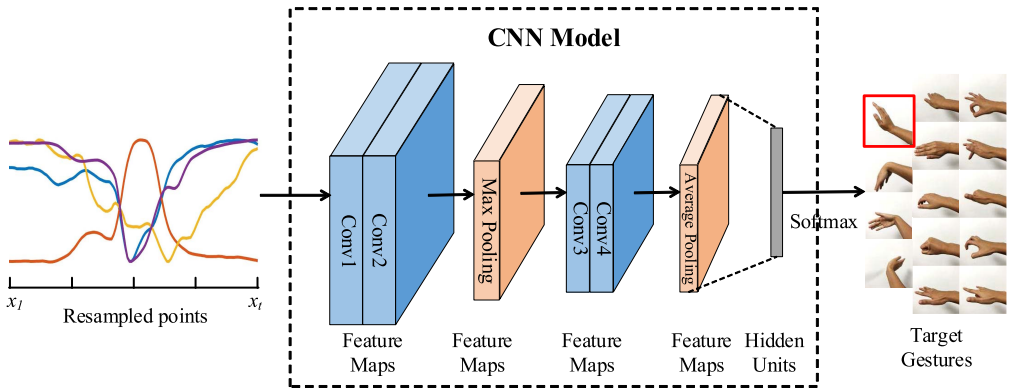


Fig. 5. Convolutional neural network (CNN) model for gesture recognition.

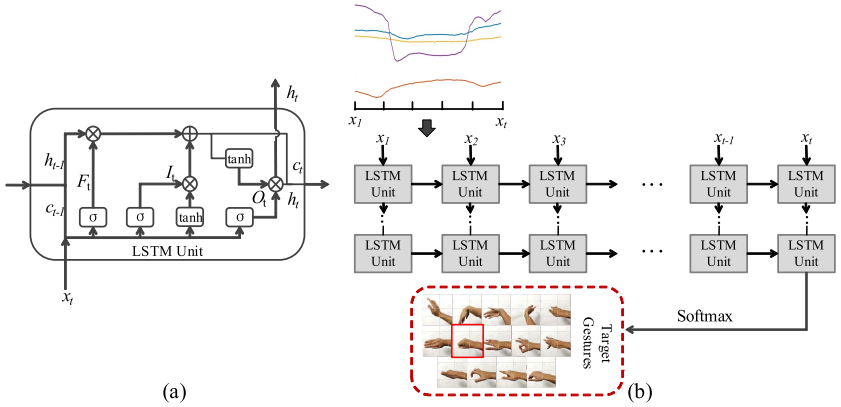


Fig. 6. (a) Long short-term memory (LSTM) unit (b) two-layer LSTM model for gesture recognition.

TABLE III
EFFICIENCY OF THE PROPOSED SEGMENTATION METHOD

Scenarios	Sitting	Walking	Jogging	Running
Correct rate	94.0±3.9%	92.2±6.3%	88.6±5.1%	88.9±4.9%
Error rate	5.8±2.7%	3.3±2.8%	5.2±2.8%	7.7±3.7%

TABLE IV
AVERAGE SIGNAL-TO-NOISE RATIO (SNR) VALUES (dB) FOR 14 PARTICIPANTS IN FOUR MOTION SCENARIOS

	Sitting	Walking	Jogging	Running
DB_ACC	15.93±4.20	8.41±1.66	2.11±2.75	0.06±0.84
DB_PPG_r	10.88±2.88	8.76±2.09	8.40±2.64	6.20±2.24
DB_PPG_i	12.93±3.57	10.17±2.41	9.47±3.13	7.26±2.50
DB_PPG_g	8.80±3.70	9.10±3.22	7.80±1.85	7.37±1.82

error was convergent and the loss function value and gesture recognition accuracy of the test set reached a higher level, the model was considered as an optimal selection.

III. RESULTS

A. Results of Segmentation and Signal-To-Noise Ratio (SNR) Calculations for Gesture Repetitions

To demonstrate the effectiveness of the proposed segmentation method, the correct detection rates and error detection rates for the gesture repetitions for all 14 participants and 14 gestures are given in Table III. The proposed gesture repetition segmentation method based on the PPG signal and motion noise achieves satisfactory performance, with a correct detection rate of over 88% in the four motion scenarios. However, with an increase in the body motion intensity, the correct detection rates decrease, and the error detection rates increase. All of the gesture samples in this section are extracted from the correctly segmented gesture repetitions.

To analyze the influence of motion noise in the four motion scenarios, the SNR values of the 14 gestures among the 14 participants are calculated for gesture samples for the ACC signal and three types of PPG signals, respectively. Table IV shows the SNR values (mean ± standard deviation) for all 14 participants, and Table V shows the SNR values (mean ± standard deviation) from the 14 gestures of one typical participant. From Table IV and Table V, the following phenomena can be observed: 1) the ACC signal is the most affected by the motion noise, and the SNR change range in the four scenarios is approximately 15.87 dB; 2) the SNRs of the red and infrared PPG signals decrease to a certain extent with an increase of the motion intensity, and their SNR change ranges in the four scenarios are 4.68 dB and 5.67 dB, respectively; 3) the green PPG signal is the least sensitive to motion noise ($p > 0.05$), as the SNR change range in the four scenarios is 1.73 dB.

B. Results for Hand Gesture Recognition

The specific network structures are designed for SVM, CNN and LSTM respectively and the parameters are listed in Table VI. The structure of three classifiers are all determined by the data set from participant 1. For SVM, we have tried kernel functions including leaner, polynomial, RBF and sigmoid, and the cost value of 1, 0.1 and 0.01. The best combination is the leaner kernel function and cost value of 0.1. For CNN,

TABLE V
SIGNAL-TO-NOISE RATIO (SNR) (dB) OF THE 14 GESTURES FOR ONE PARTICIPANT

	Sitting	Walking	Jogging	Running
Accelerometer (ACC)	18.81±3.27	9.41±0.63	0.51±0.39	-0.16±0.48
Red photoplethysmography (PPG)	15.26±7.55	11.20±4.09	7.41±2.68	10.55±2.47
Infrared PPG	18.45±5.39	12.71±4.69	7.10±3.37	11.00±3.96
Green PPG	9.77±1.96	8.55±1.19	8.00±3.64	10.16±5.11

TABLE VI
PARAMETERS FOR CNN AND LSTM

	blocks/layers	number of filters	length of filters	dropout
CNN	2	32-64	3-4	0.5
LSTM	2	32-64	-	0.6

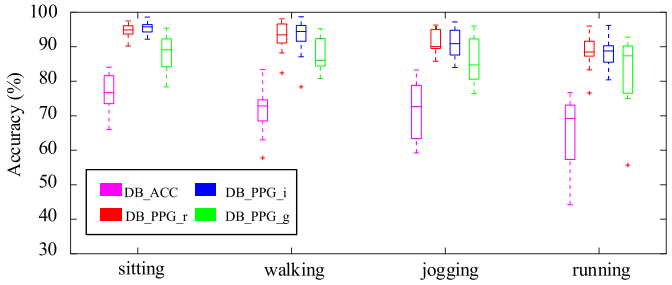


Fig. 7. Gesture recognition results in four motion scenarios.

we have tried one and two blocks (two convolutional layers and one pooling layer), and found that two blocks can well meet the needs of recognition. For the number of filters in each block, we have tried the combination of 32-64, 32-32, 16-32 and 16-16. It is found that the combination of 32-64 is usually an optional selection. The length of filter is set to 3 for the first block and 4 for the second block. As for the dropout, 0.5 is usually a good choice. For LSTM, two-layer LSTM is found to be the optional selection. For the number of LSTM units in each layer, we have tried the combination of 32-64, 32-32, 16-32 and 16-16, and found that 32-64 usually performs better. As for the dropout, 0.6 is usually a good choice in most cases. In addition, for both CNN and LSTM, batch size is usually set to between one and the sample size and categorical cross-entropy is selected as the loss function.

Table VII illustrates the gesture recognition accuracies (mean ± standard deviation) for the 14 gestures and 14 participants, using the ACC, the three types of PPG gesture samples, and the three classifiers. In each scenario, the hand gesture recognition task is conducted based on a five-fold cross-validation method, in a participant-specific way. From a macro point of view, the results are as follows: 1) the three types of PPG samples obtain higher recognition accuracies than the ACC samples for all four scenarios; and 2) the SVM and CNN have similar performance in classifying the ACC samples, and the SVM performs best for the PPG samples. Accordingly, we use the results from the SVM classifier to compare the performances of the PPG and ACC signals in classifying hand gestures from different angles.

1) Comparison in Different Motion Scenarios: Fig. 7 shows boxplots of the results obtained using the SVM classifier. In the four motion scenarios (sitting, walking, jogging, and running), the average recognition accuracies for the 14 gestures and 14 participants are 76.6±5.8%, 71.8±6.4%, 71.8±8.3%, and 65.7±9.4% for *DB_ACC*, 94.6±2.2%, 93.0±4.4%, 91.4±3.7%, and 88.6±5.1% for *DB_PPG_r*, 95.4±1.8%, 92.8±5.3%, 91.2±4.4%, and 88.3±4.1% for *DB_PPG_i*, and 88.8±5.2%, 87.5±4.9%, 86.0±6.5%, 83.3±10.2% for *DB_PPG_g*, respectively. Based on these statistics, it can be determined that: 1) in all four motion scenarios, the performance of the ACC is significantly worse than three types of PPG signals ($p < 0.001$), e.g., when comparing the best performances of the ACC and PPG signals, the accuracy of the ACC signal is approximately 18.8% lower than that of the PPG signals; and 2) for both the ACC and PPG signals, the recognition accuracy decreases with the intensity of body movement, i.e., from comparing the running and sitting scenarios, it can be seen that the accuracy of the ACC signal decreases by approximately 10.9% (from 76.6% to 65.7%), that of the red PPG signal decreases by 6.0%, the infrared PPG signal decreases by 7.1%, and the green PPG signal decreases by 5.5%; and 3) the performance of the green PPG signal is the worst among the three types of PPG signals, and is characterized by a low recognition accuracy and large standard deviation. There is no significant difference between the red PPG and infrared PPG signals ($p > 0.05$). According to above results, it can be concluded that PPG signals are more suitable for gesture interactions than the ACC signal in motion scenarios. However, the green PPG signal is not the best choice for gesture recognition.

2) Comparison in Different Gesture Types: To further explore the advantages of PPG signals, gesture recognition experiments on wrist-related and finger-related gestures are respectively conducted in this study. The 13 gestures (except G14) listed in Table I are divided into a wrist-related gesture set (***W_G***) and finger-related gesture set (***F_G***). ***W_G*** consists of G1–G6, which are mainly related to wrist movement, and ***F_G*** consists of G7–G13, which are mainly related to finger movements. Table VIII illustrates the recognition accuracies (mean ± standard deviation) for the 14 participants for these two gesture sets in the four motion scenarios. Because the infrared and red PPG signals have similarly good performances, the red PPG signal is selected for comparison with the ACC signal.

From Table VIII, the following phenomena can be observed. First, for both ***W_G*** and ***F_G***, the influence of the motion noise on the PPG signal is lower than that on the ACC signal. In the four motion scenarios, the changes of the recognition accuracies are 8.7% for W_G/ACC and 14.1% for F_G/ACC, but 4.2% for W_G/PPG and 6.9% for F_G/PPG. Second, for ***W_G***, both the ACC and PPG signals obtain relatively satisfactory recognition accuracies (over 80%) in

TABLE VII
GESTURE RECOGNITION RESULTS FOR THE THREE CLASSIFIERS (ACCURACY: %)

Scenarios	Classifier	DB_ACC	DB_PPG_r	DB_PPG_i	DB_PPG_g
Sitting	SVM	76.6±5.8	94.6±2.2	95.4±1.8	88.8±5.2
	CNN	76.7±5.7	91.8±2.7	92.6±2.6	87.6±3.7
	LSTM	76.2±8.4	90.5±3.9	92.2±3.2	84.3±4.3
Walking	SVM	71.8±6.4	93.0±4.4	92.8±5.3	87.5±4.9
	CNN	74.8±6.7	90.8±4.1	90.6±4.4	86.5±3.9
	LSTM	71.2±5.7	90.0±4.6	89.2±4.4	85.1±5.0
Jogging	SVM	71.8±8.3	91.4±3.7	91.2±4.4	86.0±6.5
	CNN	73.5±6.9	87.2±4.0	86.9±4.8	83.3±6.4
	LSTM	71.8±9.2	87.5±4.2	86.6±5.2	82.6±5.3
Running	SVM	65.7±9.4	88.6±5.1	88.3±4.1	83.3±10.2
	CNN	68.0±7.2	83.1±5.3	83.3±5.3	79.4±7.8
	LSTM	65.4±8.8	82.2±6.1	83.8±5.5	81.7±7.1

TABLE VIII
RECOGNITION RESULTS FOR WRIST-RELATED AND FINGER-RELATED GESTURE SETS (ACCURACY: %)

	Sitting	Walking	Jogging	Running
W_G/ACC	91.5±5.0	90.6±4.1	88.3±5.7	82.8±7.2
W_G/PPG	97.3±1.8	96.4±1.4	94.3±2.2	93.1±4.2
F_G/ACC	65.2±4.2	55.9±8.3	57.1±4.3	51.1±7.9
F_G/PPG	92.1±3.3	90.8±3.3	89.1±3.8	85.2±4.8

all four motion scenarios. The performance of the PPG signal is slightly better than that of the ACC signal. The average recognition accuracy is 5.8% higher in the sitting scenario, 5.8% higher in the walking scenario, 6% higher in the jogging scenario, and 10.3% higher in the running scenario. Third, for ***F_G***, the performance of the PPG signal is significantly better than that of the ACC signal. The average recognition accuracy is 26.9% higher in the sitting scenario, 34.9% higher in the walking scenario, 32% higher in the jogging scenario, and 34.1% higher in the running scenario. All these results show that the PPG signal is more suitable for the recognition of finger-related gestures in motion scenarios.

3) Comparison Regarding Needed Size of Training Set: The training burden for users is an important factor to consider in the practical applications of gesture-based interactions. To compare the requirements of the ACC-based and PPG-based technologies in regards to training samples, this section discusses a series of hand gesture recognition experiments, with the ratios of the samples between the training set and test set ranging from 1:9 to 8:2. When the ratio is 1:9, the corresponding smallest training set has only two samples for each gesture. Fig. 8(a) shows the average recognition accuracies of four types of gesture samples of 14 gestures among the 14 participants in the sitting scenario; it can be seen that for both the ACC and PPG samples, the accuracy of the test set increases with the size of the training set. When the ratios change from 1:9 to 8:2, the recognition accuracies increase from 44.1±6.6% to 76.6±5.8% for *DB_ACC*, from 68.7±12.2% to 88.8±5.2% for *DB_PPG_g*, from 83.4±6.8% to 95.4±1.8% for *DB_PPG_i*, and from 82.6±6.8% to 94.6±2.2% for *DB_PPG_r*. The ACC samples are relatively more sensitive to the size of the training set than the PPG samples. When the ratio proceeds beyond 3:7, the accuracies of the *DB_ACC* dataset increase greatly with the size of the training set; however, those from the three types of PPG datasets remain relatively stable. From Fig. 8(b), which compares the average recognition accuracies of the ACC and red PPG gesture samples in the four motion scenarios, a similar phenomenon can be observed. The results of this section show that PPG-based technologies have advantages in reducing the user training burden relative to ACC-based technologies.

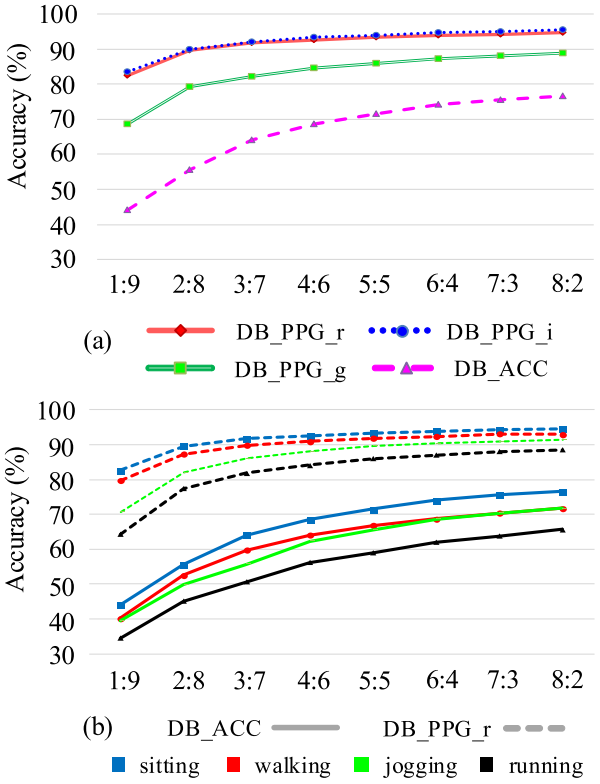


Fig. 8. Gesture recognition results under training sets with different sizes: (a) ACC and the three types of PPG in sitting scenario, (b) ACC and red PPG in the four motion scenarios.

4) Complementarity of Accelerometer (ACC) and Photoplethysmography (PPG) Signals for Gesture Recognition: As both ACC and PPG sensors are commonly embedded in wearable smart devices, the complementarity of the ACC and PPG signals for hand gesture recognition is explored in this section. Fig. 9 shows the recognition accuracies of a single-mode signal (ACC or PPG), and those of their combinations. Overall, combinations of ACC and PPG signals

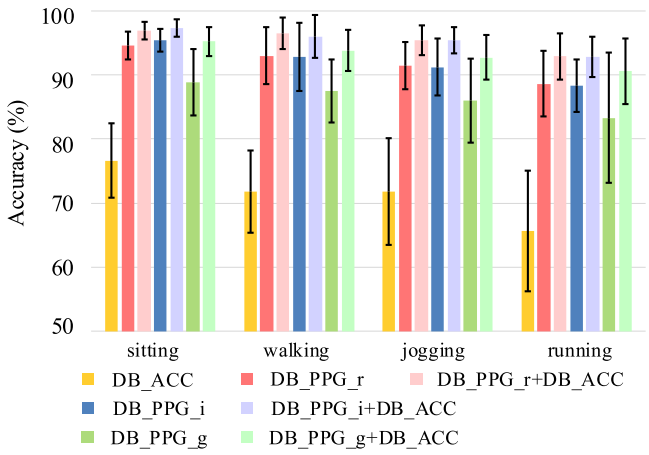


Fig. 9. Comparison of single mode and dual mode.

can significantly improve ($p < 0.05$) the gesture recognition accuracy. The combination of infrared PPG and ACC signals reaches the best average accuracy (97.3±1.3%) in the sitting scenario. The combination of red PPG and ACC signals obtains the best recognition performance in the other three scenarios; the accuracies are 96.5±2.5% in the walking scenario, 95.4±2.3% in the jogging scenario, and 92.9±3.6% in the running scenario. Taking the infrared PPG and ACC signal combination as an example, compared with the ACC signal alone, the combination increases the accuracy by approximately 20.7% in the sitting scenario, 24.2% in the walking scenario, 23.6% in the jogging scenario, and 27.1% in the running scenario. Compared with the infrared PPG signal, the accuracy increases by approximately 1.9% in the sitting scenario, 3.2% in the walking scenario, 4.2% in the jogging scenario, and 4.5% in the running scenario. Notably, the combination of the green PPG and ACC signals brings the greatest improvement in accuracy, which is over 6% higher than using the green PPG signal alone in the four scenarios. According to the above results, we can conclude that the ACC and PPG signals have a certain complementarity for hand gesture recognition.

IV. Discussion

The main problem facing the implementation of gesture interaction technology in wearable devices is that users are often in different life scenarios, such as those for work or fitness. As a result, the accuracy of the gesture recognition will be affected by the motion noise. To explore the feasibility of using sensors embedded in wearable devices to realize practical gesture interaction applications, this study conducts a comparative investigation of gesture recognition technologies based on ACC and PPG sensors. Specifically, the performances of the ACC and PPG signals in regards to gesture recognition are compared in four motion scenarios of daily life. In view of the relevant other frontier research studies, the advantages and limitations of this study can be discussed as follows.

A. Effectiveness of the Proposed Gesture Repetition Segmentation Method

In this study, the hand gesture data are collected in four motion scenarios, i.e., sitting, walking, jogging, and running. To effectively extract the gesture repetitions, a segmentation method based on the PPG signal and motion noise is proposed and implemented. The experimental results verify the effectiveness of the proposed algorithm, with an over 88% correct detection rate in the four motion scenarios. However, the motion noise reduces the correct detection rate and increases the error rate to a certain degree. It is found that irregular movements during the experiments can cause a degradation of segmentation performance. In the walking, jogging, and running scenarios, the participants did not always maintain the same swing amplitude of their arms; as such, the thresholds calculated based on the baseline signals could not reflect the practical motion noise level. Moreover, the participants did not always obey the requirement for a 2-s interval between gesture repetitions in the experiment. When a short time interval occurs, the derivation of the energy of the gesture signal will move outside of the range of the thresholds. As described in the method, the ending points of the gesture repetitions are determined by the length of threshold range (T_{max}–T_{min}); thus, a short time interval can lead to the 'disappearance' of the baseline, and the failure of segmentation.

B. Which is More Suitable for Gesture Interactions in Wearable Devices, PPG or ACC?

For a long time, PPG signals have commonly been used for heart rate measurements and health monitoring [22], [23]. Zhang *et al.* [30] conducted the first study exploring the feasibility of using the optical sensors on the off-the-shelf wearable devices to recognize a set of 10 conflict-free gestures. In their work, the performances of IMU (an ACC with three axes and gyroscope with three axes) and PPG systems in hand gesture recognition were compared for different life scenarios. The IMU-based system was found to outperform the PPG system for the static scenarios, with higher accuracy for sitting down, sitting down while learning on a table, and standing still. However, in the motion scenarios such as walking in a garden, taking public transportation, and jogging, the accuracy of the PPG system was relatively stable, but that of the IMU decreased with the intensity of the movement. In fact, the IMU-based system failed to recognize most of gestures, owing to the significant motion noise.

In this study, a comparative study between ACC and PPG systems is conducted based on 14 gestures, and four motion scenarios of daily life are considered: sitting, walking, jogging, and running. Consistent with reference [30], the experimental results for hand gesture recognition demonstrate that the decrease of the recognition accuracy in the PPG system with the intensity of body movement is significantly lower than that of the ACC system. This study further demonstrates the advantages of PPG systems from different perspectives. The analysis of the SNR values of gesture samples shows that the ACC signal is much more affected by the motion noise than the PPG signals. The gesture recognition experiment on wrist-related gestures and finger-related gestures shows that the performance of the PPG system is significantly better than that of the ACC system for the classification of finger-related

gestures. The hand gesture recognition experiment (with the ratios of the samples between the training set and test set ranging from 1:9 to 8:2) demonstrates that the PPG system has advantages in regards to reducing the user training burden. Based on above results, it is concluded that a PPG system is more suitable for implementing gesture interactions on wearable devices than an ACC system. In particular, the robustness of the PPG system in regards to finger-related gestures makes it possible to use finger gestures to achieve interactions such as keyboard tapping.

Moreover, as the green PPG signal has much greater absorbability for oxyhemoglobin and deoxyhemoglobin [38], most commercial products are embedded with green LEDs to acquire the PPG signal. Correspondingly, most studies have adopted green PPG signals for gesture recognition. In this study, the performances of three types of available PPG signals (including infrared light, red light, and green light) in regards to hand gestures are simultaneously investigated. Although the green PPG signal is the least sensitive to motion noise, it obtains the lowest gesture recognition accuracy. Therefore, the green PPG signal is not the best choice for the implementation of gesture interactions.

At last, this study first explores the complementary effect of PPG and ACC systems, providing a basis for robust gesture recognition based on the fusion of these two types of sensor information.

C. Comparison With Related Works

As described in the Introduction, Zhao *et al.* [31] obtained an average recognition accuracy of 88% by using two PPG sensors for nine finger-based gestures. Zhang *et al.* [30] implemented PPG-based gesture recognition using the heart rate monitor of an off-the-shelf Samsung Gear S3 Frontier smartwatch, and achieved an average accuracy of 73.6%–90.55% for 10 gestures. Subramanian *et al.* [32] realized the recognition of four simple hand gestures over four individuals with above 92.4% accuracy by adopting three PPG sensors. In both [31] and [32], the gesture recognition experiments were carried out in static scenario. Different from the works mentioned above, this study considered four motion scenarios including sitting, walking, jogging and running. From the view of recognition accuracy, in static scenario, this study (88.8%-95.4%) does not perform better than related works. However, the 83.3%-91.4% of recognition accuracy in jogging and running scenarios are significantly higher than those of jogging task in [30].

On the other hand, when a PPG system is applied to implement applications for gesture interactions, the number of sensors should be considered. In this study, a homemade four-channel PPG wristband is applied, and the average accuracy is in the range of 83.3%–95.4% for 14 gestures. In our view, the increase of the PPG channels increases the gesture recognition accuracy. As we know, wearable devices such as the Miband5 [17] are usually embedded with one or two PPG sensors. Therefore, to achieve accurate and robust PPG-based gesture recognition technology, the design of a PPG array (similar to a high-density sEMG array [5]) may be an effective solution.

D. Limitations and Future Work

Initially, we should point out that, different from the IMU-based method in [30] which uses acceleration and gyroscope signals for gesture recognition, this study only uses the acceleration signal for gesture recognition; thus, the recognition rate is relatively low. Taking the sitting scenario, for example, the IMU-based system obtains an average accuracy of 98.27% for 10 gestures; however, the ACC-based scheme only obtained 76.6% recognition accuracy for the 14 gestures. Secondly, only CNN and LSTM networks with simple structures are used for gesture recognition, resulting in poor performance relative to an SVM classifier. In the future, research should be conducted to further improve the accuracy and robustness of gesture recognition by using the powerful learning and generalization abilities of neural networks. Finally, the hand gesture recognition experiments in this study are participant-specific. To achieve universal gesture interaction applications, a follow-up work will explore effective participant-independent recognition schemes.

V. Conclusion

With the goal of exploring the feasibility of using the ACCs and PPG sensors embedded in wearable devices to realize applications for gesture interactions, this paper presents a comparative study of hand gesture recognition based on an ACC signal and three types of PPG signals. A gesture repetition segmentation method is proposed based on PPG and motion noise, and the effectiveness is verified. Second, the performances of ACC and PPG systems are compared from multiple perspectives in four different motion scenarios. A PPG system is found to be more suitable for implementing gesture interactions in wearable devices, owing to its advantages of being insensitive to motion noise, and reducing the user's training burden. Moreover, increasing the number of channels can improve the PPG-based recognition accuracy. However, the green PPG signal is not the best choice for the implementation of gesture interactions. The research results provide a reliable foundation for realizing the application of gesture-based interactions in wearable devices.

References

[1] G. Plouffe and A.-M. Cretu, "Static and dynamic hand gesture recognition in depth data using dynamic time warping," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 2, pp. 305–316, Feb. 2016.

[2] 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.

[3] B. Fang, F. Sun, H. Liu, and C. Liu, "3D human gesture capturing and recognition by the IMMU-based data glove," *Neurocomputing*, vol. 277, pp. 198–207, Feb. 2018.

[4] D. L. Quam, "Gesture recognition with a DataGlove," in *Proc. IEEE Conf. Aerosp. Electron.*, vol 2, May 1990, pp. 755–760.

[5] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 4, pp. 1292–1304, Apr. 2021.

[6] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, no. 1, Dec. 2016, Art. no. 36571.

[7] S. M. Mane, R. A. Kambli, F. S. Kazi, and N. M. Singh, "Hand motion recognition from single channel surface EMG using wavelet & artificial neural network," *Procedia Comput. Sci.*, vol. 49, pp. 58–65, Jan. 2015.

[8] Y.-C. Chu, Y.-J. Jhang, T.-M. Tai, and W.-J. Hwang, "Recognition of hand gesture sequences by accelerometers and gyroscopes," *Appl. Sci.*, vol. 10, no. 18, p. 6507, Sep. 2020.
[9] R. Xu, S. Zhou, and W. J. Li, "MEMS accelerometer based nonspecific-user hand gesture recognition," *IEEE Sensors J.*, vol. 12, no. 5, pp. 1166–1173, May 2012.
[10] M.-K. Liu, Y.-T. Lin, Z.-W. Qiu, C.-K. Kuo, and C.-K. Wu, "Hand gesture recognition by a MMG-based wearable device," *IEEE Sensors J.*, vol. 20, no. 24, pp. 14703–14712, Dec. 2020.
[11] S. Yamakawa and T. Nojima, "A proposal for MMG-based hand gesture recognition," in *Proc. 25th Annu. ACM Symp. User Interface Softw. Technol.*, 2012, pp. 89–90.
[12] K. Fukuchi, T. Sato, H. Mamiya, and H. Koike, "PAC-PAC: Pinching gesture recognition for tabletop entertainment system," in *Proc. Int. Conf. Adv. Vis. Interfaces (AVI)*, 2010, pp. 267–273.
[13] E. Nasr-Esfahani *et al.*, "Hand gesture recognition for contactless device control in operating rooms," 2016, *arXiv:1611.04138*. [Online]. Available: http://arxiv.org/abs/1611.04138
[14] N. Adamo-Villani, J. Heisler, and L. Arns, "Two gesture recognition systems for immersive math education of the deaf," in *Proc. 1st Int. Conf. Immersive Telecommun.*, 2007, pp. 1–6.
[15] (Mar. 2020). *IDC Worldwide Quarterly Wearable Devices Tracker*. [Online]. Available: https://www.idc.com/
[16] (Sep. 2020). *Apple Watch Series 6*. [Online]. Available: https://support.apple.com/
[17] *Mi Smart Band 5*. [Online]. Available: https://www.mi.com/global/mi-smart-band-5
[18] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors*, vol. 16, no. 4, p. 426, Mar. 2016.
[19] H. Wen, J. R. Rojas, and A. K. Dey, "Serendipity: Finger gesture recognition using an off-the-shelf smartwatch," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2016, pp. 3847–3851.
[20] M.-S. Kang, H.-W. Kang, C. Lee, and K. Moon, "The gesture recognition technology based on IMU sensor for personal active spinning," in *Proc. 20th Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2018, pp. 546–552.
[21] M. Kim, J. Cho, S. Lee, and Y. Jung, "IMU sensor-based hand gesture recognition for human-machine interfaces," *Sensors*, vol. 19, no. 18, p. 3827, Sep. 2019.
[22] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1902–1910, Aug. 2015.
[23] C.-C. Zhou, H.-W. Wang, Y.-M. Zhang, and X.-S. Ye, "Study of a ring-type surgical pleth index monitoring system based on flexible PPG sensor," *IEEE Sensors J.*, early access, Nov. 27, 2020, doi: 10.1109/JSEN.2020.3041072.
[24] N. Karimian, M. Tehranipoor, and D. Forte, "Non-fiducial PPG-based authentication for healthcare application," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Feb. 2017, pp. 429–432.
[25] J. Sancho, Á. Alesanco, and J. García, "Biometric authentication using the PPG: A long-term feasibility study," *Sensors*, vol. 18, no. 5, p. 1525, May 2018.
[26] A. Goshvarpour and A. Goshvarpour, "Poincaré's section analysis for PPG-based automatic emotion recognition," *Chaos, Solitons Fractals*, vol. 114, pp. 400–407, Sep. 2018.
[27] M. W. Park, C. J. Kim, M. Hwang, and E. C. Lee, "Individual emotion classification between happiness and sadness by analyzing photoplethysmography and skin temperature," in *Proc. 4th World Congr. Softw. Eng.*, 2013, pp. 190–194.
[28] M. Boukhechba, L. Cai, C. Wu, and L. E. Barnes, "ActiPPG: Using deep neural networks for activity recognition from wrist-worn photoplethysmography (PPG) sensors," *Smart Health*, vol. 14, Dec. 2019, Art. no. 100082.
[29] T. Buddhika, H. Zhang, S. W. T. Chan, V. Dissanayake, S. Nanayakkara, and R. Zimmermann, "FSense: Unlocking the dimension of force for gestural interactions using smartwatch PPG sensor," in *Proc. 10th Augmented Hum. Int. Conf. (AH)*, Mar. 2019, pp. 1–5.
[30] Y. Zhang, T. Gu, C. Luo, V. Kostakos, and A. Seneviratne, "FinDroidHR: Smartwatch gesture input with optical heartrate monitor," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–42, Mar. 2018.
[31] T. Zhao, J. Liu, Y. Wang, H. Liu, and Y. Chen, "PPG-based finger-level gesture recognition leveraging wearables," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2018, pp. 1457–1465.
[32] K. Subramanian, C. Savur, and F. Sahin, "Using photoplethysmography for simple hand gesture recognition," in *Proc. IEEE 15th Int. Conf. Syst. Syst. Eng. (SoSE)*, Jun. 2020, pp. 307–312.
[33] T. Zhao, J. Liu, Y. Wang, H. Liu, and Y. Chen, "Towards low-cost sign language gesture recognition leveraging wearables," *IEEE Trans. Mobile Comput.*, vol. 20, no. 4, pp. 1685–1701, Apr. 2021.
[34] S. G. K. Patro and K. K. Sahu, "Normalization: A preprocessing stage," 2015, *arXiv:1503.06462*. [Online]. Available: http://arxiv.org/abs/1503.06462
[35] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[36] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[37] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[38] T. Tamura, Y. Maeda, M. Sekine, and M. Yoshida, "Wearable photoplethysmographic sensors—Past and present," *Electronics*, vol. 3, no. 2, pp. 282–302, Apr. 2014.

Yu Ling is currently pursuing the M.S. degree in biomedical engineering with the University of Science and Technology of China. His research direction is biomedical signal processing.

Xiang Chen (Member, IEEE) received the M.S. degree in biomedical engineering and the Ph.D. degree in biomedical engineering from the University of Science and Technology of China, Heifei, China, in 2000 and 2004, respectively. Her research interests include biomedical signal processing, human–computer interaction, and sign language recognition.

Yuwen Ruan is currently pursuing the M.S. degree in biomedical engineering with the University of Science and Technology of China. Her research direction is biomedical signal processing.

Xu Zhang (Member, IEEE) received the B.S. degree in electronic science and technology and the Ph.D. degree in biomedical engineering from the University of Science and Technology of China in 2005 and 2010, respectively. His research interests include intelligent myoelectric control and electrophysiological multimodal human–computer interaction.

Xun Chen (Senior Member, IEEE) received the B.S. degree in electronic science and technology from the University of Science and Technology of China in 2009 and the Ph.D. degree in electrical and computer engineering with The University of British Columbia in 2014. His research interests include the broad areas of statistical signal processing and machine learning in biomedical applications.